OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 01 - Issue 12 (2024)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

UTILIZING BIG DATA ANALYTICS TO OPTIMIZE INFORMATION COMMUNICATION TECHNOLOGY STRATEGIES

Aafia Latif
Department of Computer Science & Information Technology, Govt Graduate College, Burewala, Pakistan.
Sadia Latif
Department of Computer Science Bahauddin Zakaria University, Multan, Pakistan.
Assad Latif
School of management and engineering North China University of water resources and electric power Zhengzhou Henan, China
Furqan Jamil
Department of Computer Science, National College of Business Administration and Economics, Multan, Pakistan
Salahuddin
Department of Computer Science, NFC Institute of Engineering and technology, Multan, Pakistan.
Corresponding Author : Aafia Latif (aafialatif@gmail.com)

DOI: https://doi.org/10.71146/kjmr158

Article Info

Abstract
In today's interconnected world, ICT-Information and Communication Technology is a pivotal economic development driver. This transformative force fosters global technological advancements, facilitates seamless communication, revolutionizes manufacturing systems, and propels economic growth and progress. As international trade and foreign direct investment take center stage in the global economy, the role of modern information and communication technologies becomes increasingly indispensable. This report explores the profound impact of ICT on economic activities and presents policy proposals aimed at enhancing ICT standards, with a particular focus on informed investment decision-making. Additionally, it delves into the positive correlation between decision-making informed by big data analytics and economic development, utilizing advanced econometric models and the GMM-Generalized Method of Moments model for an in-depth analysis. By embracing these recommendations, nations can leverage ICT to its fullest extent, fostering sustainable economic development and prosperity in the digital age.

Keywords: *Information and communication technology, GMM model and Big Data Analytics*

Introduction

ICT-Information and Communication Technology is a cornerstone of modern society, wielding an immense influence on economic growth and development. Its impact extends across multiple dimensions, from enhancing productivity through automation and data-driven decision-making to generating new employment opportunities in technology-related fields. Moreover, *ICT* catalyzes economic growth by nurturing innovation and entrepreneurship, fostering an environment conducive to prosperity. Notably, *ICT* also plays a vital role in poverty reduction by creating avenues for income generation, particularly in the ever-expanding tech sector. Simultaneously, it acts as an enabler of equitable access to essential services such as healthcare and education. Integrating technology in these sectors, exemplified by telemedicine and e-learning platforms, improves accessibility, cost-effectiveness, and overall quality. *ICT's* comprehensive components, encompassing hardware, software, networks, and multimedia, enable its multifunctional nature, making it a driving force behind the evolving landscape of economies and societies worldwide.

Existing research indicates that *ICT* plays a pivotal role in driving economic growth, exerting a substantial influence on productivity, employment, and overall development. According to findings by the *OECD*, the integration of ICT not only creates fresh avenues for employment but also facilitates affordable access to vital healthcare and educational services. Technological advancements hold the potential to alleviate poverty by generating new employment opportunities and ensuring cost-effective access to essential healthcare and education resources. The components of *ICT* encompass a wide spectrum, encompassing functions such as data collection, storage, processing, transmission, hardware and software infrastructure, networks, multimedia applications, and data visualization, including audio, text, images, and other forms of information (World Bank, *2017*).

Big data is broadly recognized as a wellspring of innovative offerings, services, and business prospects (Davenport et al., *2018*; Davenport & Kudyba, *2016*; McAfee & Brynjolfsson, *2018*). Additionally, it is perceived as a catalyst for enhancing operational efficiency and efficacy. This includes streamlining supply chain operations, optimizing pricing strategies, judiciously assigning personnel to tasks, minimizing errors and quality issues, and fostering improved customer relations (Chen et al., *2018*; Davenport, *2017*; McAfee & Brynjolfsson, *2018*).

Furthermore, big data offers the potential for both economic and social benefits, including improved decision-making (Sharma et al., *2018*) and more informed strategic planning (Constantiou & Kallinikos, *2018*). This has led to a strong emphasis in academic and practical literature on the opportunities big data presents for organizations (Clarke, *2016*). However, it is essential to note that while there is significant optimism surrounding big data, the mere presence of vast amounts of data does not guarantee the realization of actual value. Organizations sometimes may overestimate their ability to extract value from big data (Ransbotham et al., *2016*; Ross et al., *2018*).

As the initial discussions about big data tend to be characterized by unfounded optimism (Arnott & Pervan, *2018*), there is a critical need to examine how organizations translate the potential benefits into tangible social and economic value (Markus & Topi, *2018*). More specifically, research is required to understand the strategies organizations develop to harness the value inherent in big data and identify instances where these strategies may fall short in practice.

While compelling success stories are associated with Big Data Analytics, it is crucial to acknowledge the significant costs and challenges of implementing Big Data initiatives. Many organizations need help to undertake such endeavors. Furthermore, despite numerous anecdotes, extensive empirical evidence still needs to be provided to support the overall net benefits of Big Data Analytics.

Therefore, the primary research question addressed in this paper is whether adopting Big Data Analytics leads to enhanced performance and, consequently, a competitive edge within the manufacturing industry of Pakistan, particularly when it comes to making data-informed decisions. Successful business narratives are often the result of informed and innovative decision-making processes founded on reliable and precise information.

In recent years, *ICT* has been recognized as a pivotal driver of economic growth. This recognition stems from various factors, including *ICT's* capacity to provide rapid access to information and expertise for various economic stakeholders. It is worth noting that many firms have effectively increased their productivity by leveraging *ICT* to reduce manufacturing costs through improved communication channels (Chen et al., *2018*).

Moreover, internet connectivity has played a crucial role in fostering the long-term development of businesses and *SME*-Small and Medium-sized Enterprises. It has accomplished this by lowering financial barriers, reducing information imbalances, and minimizing agency costs. Several scholars have established a substantial connection between *ICT* adoption and economic growth (Bahrini & Qaffas, *2019*).

The rapid expansion of the *ICT* sector has significantly augmented overall economic productivity by enhancing the efficiency and effectiveness of various industries. This research makes a dual contribution to the existing body of knowledge. Firstly, it delves into the connection between *ICT* development and economic growth by considering five distinct *ICT* development indicators, a departure from the limited use of one or two indicators in most prior studies. Secondly, this study emphasizes developing countries, a noteworthy departure from the paucity of recent research examining the interplay between *ICT* and economic growth within these nations.

1. Related Work

Pakistan has taken a significant step forward in data science and cloud computing by establishing the *NCBC*-National Center in Big Data and Cloud Computing. This pioneering initiative unites 12 cutting-edge laboratories from 11 prestigious universities across the country, spanning diverse fields such as medicine, agriculture, energy management, and distribution. The *PKR* 1.5 billion center, headquartered at the *LUMS*-Lahore University of Management Sciences, was officially inaugurated.

Participating laboratories draw from universities such as LUMS, NUST-National University of Science and Technology; the UET-University of Engineering and Technology Lahore and Peshawar campuses; ITU-Information Technology University; NUCES-National University of Emerging Sciences-FAST; SZABIST-Shaheed Zulfikar Ali Bhutto Institute of Science and Technology; CECOS-University of Information Technology and Emerging Sciences-Peshawar; the University of Agriculture-Faisalabad; Ziauddin University; and NED-University of Engineering and Technology.

Amidst economic challenges that forced some businesses to shut down operations and others to downsize their workforces, the Government of Pakistan is actively allocating resources to fortify its defense against internal vandalism and street crimes. This endeavor places a heavier tax burden on the industrial sector, affecting the import of raw materials, energy procurement, and fuel purchases for manufacturing units.

Furthermore, the appreciation of the US *Dollar ($)* to the Pakistani Rupee has intensified the financial challenges, making it even more critical for the manufacturing industry in Pakistan to adopt Big Data Analytics. This transformation is not merely a choice but a necessity, enabling the industry to refine its operations and make more informed and efficient decisions in these testing times.

Opting for business continuation is one prevailing choice in Pakistan's manufacturing industry cash burning within seconds. This situation has caused the demand for huge volumes of data with accurate and reliable information to become even stricter. If it is not made or bought, downsize or sustain, and shut down or remain open and operational, it would require intellectual insight. Big Data with Big Cost itself is very huge in making a decision, especially for the prevalent conditions of the manufacturing sector in Pakistan, where profits are diminishing. Research into the empirical need of intelligent decision-making system to convert problems into opportunities requires immense patience.

Research Gap

In order to examine the relationship between the key characteristics of Big Data, namely Volume, Variety, Velocity, Veracity, and Value, and their impact on decision-making within the manufacturing industry of Pakistan, this study aimed to address the following inquiry: What demographic attributes can be ascertained regarding the respondents, including Executives/Board of Directors and Managers, concerning the following factors: *Name, Age, Educational attainment, Designation, Civil status, Gender, Number of years of experience in the industry, Number of years of tenure within the company*

RQ1: How significantly do the following characteristics of Big Data Analytics impact decision-making processes?

RQ2: How Does Information-Driven Decision Making Enhance Business Performance

RQ3: What are the results of integrating Big Data Analytics into an organization's decision-making?

RQ4: Is there a noteworthy distinction between the attributes of big data analytics and their impact on decision-making

RQ5: Does the Influence of Big Data Analytics (*BDA*) Significantly Impact Foreign Direct Investment (*FDI*) Decision-Making

1.1.1. Hypotheses

Building upon the conceptual framework and pertinent theories in the realm of Big Data analytics and decision-making, this study formulated the subsequent null hypotheses:

- *H01:* No substantial correlation exists between the volume of data in Big Data Analytics and its impact on investment decision-making.
- *H02:* There exists no notable correlation between the speed of information processing in Big Data Analytics and its impact on investment decision-making.
- *H03:* There is no notable correlation between the diversity of information in Big Data Analytics and its impact on investment decision-making.
- *H04:* There is no significant relationship between veracity of information inBig Data Analytics and decision making for the investment decision.
- *H05:* There is no noteworthy correlation between the value of information in the realm of Big Data Analytics and its impact on investment decision-making.

- *H06:* There exists no substantial correlation between the variability of information within the realm of Big Data Analytics and its impact on investment decision-making.
- *H0 7*: Visualizing Information in Big Data Analytics Does Not Significantly Impact Investment Decision-Making
- *H08:* ICT Positively Correlates with Investment Decision-Making

2. Proposed Methodology

2.1. Technical Framework

The technological and digital revolution era" refers to the period characterized by the rapid advancement of technology and the widespread adoption of digital tools and processes. During this era, there has been a fundamental shift in how information is stored, managed, and utilized. This has led to significant changes in various aspects of society, including business, communication, and daily life.

The phrase "changed the way we store and handle information" points to the transformative effects technology has on information handling techniques. Instead of storing records in paper, digital data stores, cloud storage, and other kinds of data analytics have predominantly replaced the traditional information storage and handling techniques. Thus, the change has smoothed information access and retrieval processes, ensuring an efficient process of data handling.

The statement "*fostering a new paradigm on how to process all this information*" highlights the emergence of a fresh approach to dealing with the abundance of data generated in the digital age. This new paradigm involves leveraging technology, machine learning, and data analytics to extract meaningful insights from large datasets, facilitating data-driven decision-making.

The reference to a "*simple search on the word 'innovation' in Google*" illustrates the vastness of information available on the internet. It is an example of the overwhelming volume of data that individuals and organizations encounter daily. This abundance of data presents opportunities and challenges, emphasizing the need for effective data collection, analysis, and interpretation.

The phrase "*the challenge of collecting the right information from many possible inputs*" underscores the difficulty of sifting through a myriad of data sources to identify and extract relevant information. Organizations must employ strategies and technologies to filter and prioritize data based on their specific needs and objectives in a world inundated with information.

The mention of "*Chesbrough (2018)*" alludes to the work of Henry Chesbrough, a prominent scholar in innovation. *Chesbrough's* research emphasizes the value that companies can derive from innovation, which can take various forms, including the development of new ideas, technologies, products, and services or the transformation of business models. This concept underscores the importance of innovation in staying competitive and creating value in today's rapidly changing business landscape.

Big data analytics has only recently been scratched according to Kusiak (2017). It is a very powerful tool for companies to shape, evolve, and improve the designs of products, systems, and services by embedding them with intelligence, better connections, effectiveness, and access. Innovation framework wherein choice for an alternative is made which best fulfills the expectations and makes data positive influence on innovation as Kusiak (2018) mentioned. In contrast, Cooper (2018) says that data has to be improved during the innovation development process. That is emphasized in many sections of his book "Winning New Products." As companies enter into the development phase, 4/5 companies have no useful information about the price sensitivity of customers, 3/4 lack data from customers, and 2/3 have no dependable data on market size and sales forecasting (Cooper, 2018). Therefore, the largest problem faced while choosing a project or in portfolio management is better data integrity and not tool selection (Cooper, 2018). Another problem identified by Cooper (2018) was weak front-end homework, which included market studies, technical assessments, and financial analysis.

For instance, 84% do not value the product for their customers, 82% rely on bad market research, 78% are weak in doing assessment for operations or source of supply, and last but not least, 74% are weak on business and financial analysis (Cooper, 2018). It is those companies that create a customer data-obsessed culture and make changes based on results, who will see success. The benefits of big data are there for any company in any industry to grasp and, upon implementation, yield insights and reveal opportunities that may not be recognized in any other way (Glass and Callahan, 2018). Big Data is more than a flash-in-the-pan technology because it focuses on single, actionable insights of customers, products, and operations as an excellent business transformation opportunity to rethink the process of value creation, streamline key business activities, and reveal new monetization opportunities (Schmarzo, 2018). According to a number of reports, big data analytics has enormous advantages; among them is BDVA-Big Data Value Association (2016) and CDI-Center for Data Innovation (MacDonnell & Castro, 2016). No doubt, its potential is unquestionable, but the business understanding and implementation of big data are astonishingly haphazard (Glass and Callahan, 2018).

Most people need it explained to them, and very few people know what to do with it (Marr, 2018). According to Marr (2018), the actual worth is not in the large quantities of data but in the ability to analyze massive and complex sets of data to help in decision-making. Cooper (2018) states that there is a need for more innovation because of a lack of data; however, today, data is mass-produced. Kusiak (2017) comments that big data analytics can impact innovation, but that option has not been well developed yet. Big data is a form of business transformation, and the firms that take the big data leap will lead the way (Schmarzo, 2018; Glass & Callahan, 2018). The majority of reports list big data analytics as a strategic enabler for economic prosperity and social development (Big et al. Association, 2016; MacDonnell & Castro, 2016). Data is perceived as the new raw material equal to capital and labor (The Economist, 2018). Big data analytics can find patterns and relations that humans cannot see (The Economist, 2018a). Despite all these benefits, the business world hardly integrates big data and there should be much more understanding of these opportunities

(Glass and Callahan, 2018; Marr, 2018). So, the role of big data analytics in the innovation process needs to be further explored and explained.

There were several global management consulting firms within the world's leading information technology research and advisory companies that forecasted big data analytics as the next logical step for industries to succeed. Gobble (2018) points out big data as the next big thing, while Gartner shows big data and demonstrates its significance through 2018 to 2018 analysis reports of the top 10 strategy technologies forecasted (2018, 2018b, 2018). According to McKinsey Global Report (McKinsey Global et al., 2018), big data represent the next frontier for innovation, competition, and productivity, transforming innovation in products and services. Columbus reports that 89 percent business leaders assert that big data will transform operations in a similar way the internet did and 83 percent have initiated big data projects to achieve a competitive advantage. According to Forrester, Brian Hopkins (2018), in 2016, customer-obsessed leaders will outshine their competitors as companies will want to transform customer experiences and increase revenue. Because firms move beyond big data and solve issues through data thinking, insight will become a sharp competitive weapon according to Hopkins (2018). According to Deloitte, the benefits of big data analytics are very well described and disclose it is "about making better business decisions faster and with more confidence" (Cheng et al., 2018).

In their broad data analytics report, Cheng et al. (2018) of Deloitte Australia mention, "For industries whose success depends more on customer insights than on operational intelligence, it makes sense to maintain a sense of opportunity with big data." Porter and Heppelmann (2018) explain how intelligent and connected products can better provide for customers.

Insights about companies making those products and potentially selling that data to a third-party company. As mentioned by Columbus (2018), big data analytics enables Asian manufacturing companies to trim costs, gain new customers, enhance productivity, as well as optimize facilities, as presented in Appendix B. According to Gartner in 2018, big data would spawn $28 billion in IT spending in 2018, $34 billion in 2018, and $232 billion in 2016 (2018c, 2018a). In 2018, Gartner (2018a) launched a survey which showed an increase in significant data investments by having more than 75 percent of the companies planning to invest (Gartner, 2018a). Fig 1. Columbus (2018) illustrated the upward forecast trend of the big data analytics market from $7.6B in 2018 to $84.69B in 2026. IDC1 study, conducted in 2018, predicts that there will be a shortage of skilled staff in 2020 due to the rising demand for big data. According to The Economist (2018), it should worry about how to train the next generation-not only scientists but people in government and industry-to be prepared for the "industrial revolution of data" and its impacts on society. According to the Center for Data Innovation, data-driven innovation can influence the industry towards increasing economic productivity and peoples' quality of life through the facilitation of customizable essential social services such as education (MacDonnell & Castro, 2016).

2.2. Conceptual Framework

The study conceptualizes a decision-making model and framework based on the characteristics of big data: volume, value, velocity, variety, veracity, value, variability, and visualization. Decision making for the foreign direct investment by the multinational specially for the manufacturing industry of Pakistan has remained difficult for the years. Current competitive market-based decisions in the manufacturing-based industry of Pakistan need to be innovative and timely, as market is saturated with rivals and competing agencies waiting for leverage to gain customer base and loyalties. So, study intends to design a framework of decision making based on questionnaire with analysis. The study in the second part also provides the time series data from the Pakistan and *GMM* model for understanding the nature of *FDI* decisions and their trends**.**

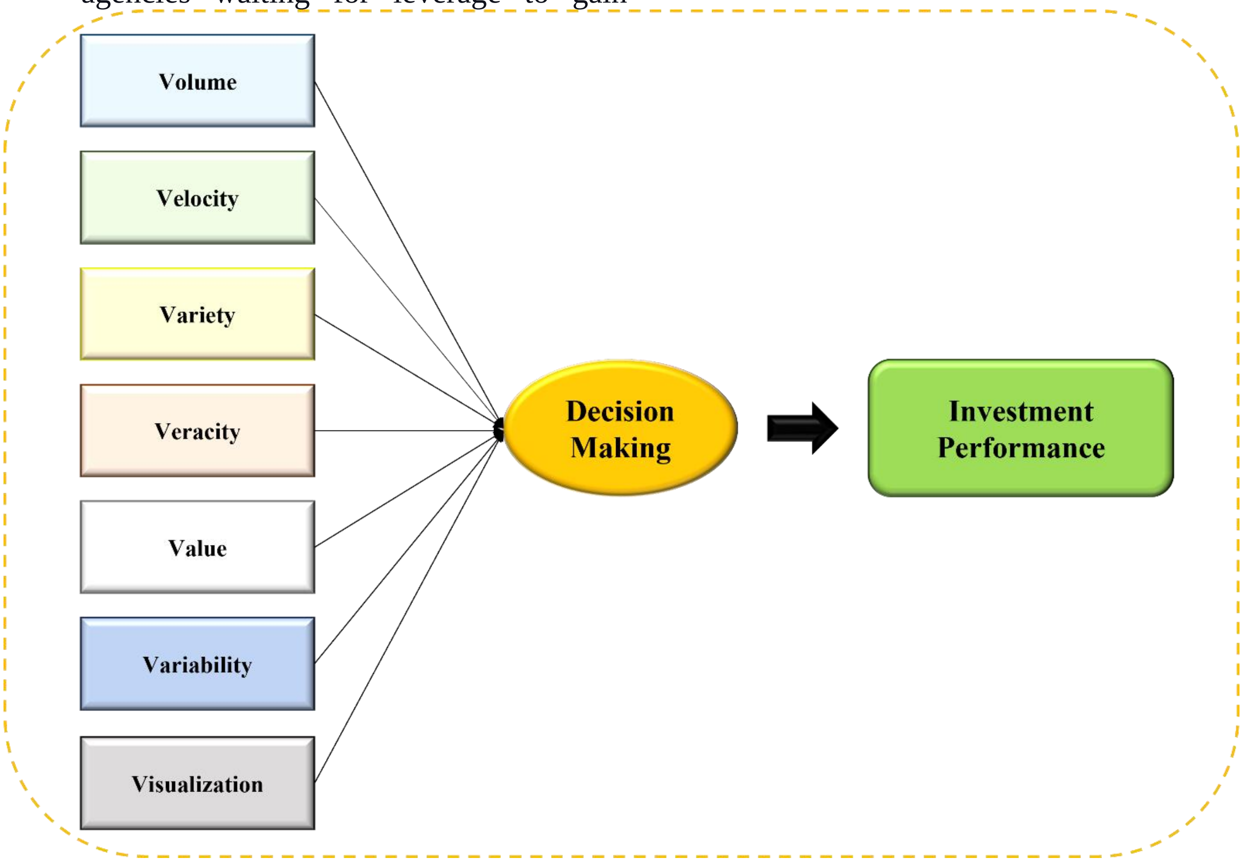

Fig 1 Conceptual Framework

3.3 Research Design

The research plan is to indicate the strategies and methods for data collecting and data analysis (Saunders, 2018). Partial Least Square SEM (PLS–SEM) is best suited for exploratory studies with less theoretical backing to the concepts and hypo paper, and the sample size is small. Since there is mixed information, the chances of multi-collinearity between independent or predictor variables could be high about the constructs. Due to non-membership, measurement error variables in PLS-SEM are not correlated. Another bonus with the SEM approach is that single variables have been modified step-by-step with all the rest in the model and their measurement model fit indices also evaluated in the measurement part of the model. Since the error terms are not dealt in the model in PLS-SEM unlike the

CBSEM, it makes use of the proxies of the latent variables. Since the sample was relatively small and an exploratory research study we have opted for PLS-SEM technique rather than SEM technique. R2 dependent construct each for decision-making models in economy. Since it is an exploratory study and the adoption of IOT is still in its nascent stage in developing countries, the study was done on the sample of respondents. For smaller sample sizes, PLS-SEM is preferred. Applications of GMM model have been used to explain the relationship. Quantitative research is to evaluate the data as well as the result or conclusive proof. In other words, quantity is used for numerical data in the data collection procedure and investigation method (Saunders, 2018). Generally, quantitative research uses some statistical analysis. In this study, the purpose for, we are using quantitative research to test the theory from the surveys. A few common methods of collecting quantitative data are questionnaires, interviews, perceptions, or using existing materials. Quantitative research method employs one quantitative data collection technique "with quantitative data analysis methods (Saunders, 2018). At that point, this research study is a cross-sectional study, which means this study is the only study of a phenomenon at a" specific time.

3. Implementation and Results

4.1. Sample size

Besides, the sample size of this research is 720 in Pakistan, with a 99% response rate. In any case, the sample size has been chosen based on the rule of thumb prescribed by (Hair et al., 2018), which says that the sample size ought to be 5 to 10 multiplied by the number of items. The current study has satisfied the most minor threshold and utilized the sample according to the rule of thumb.

Sample size calculation, by putting the values in formula we get

$$\frac{Z^{2} * P\,(1\text{-}P) / e^{2}}{1 + (Z^{2} * P\,(1\text{-}P) / e^{2} N)}$$

$$\frac{2.58^{2} * .5(1\text{-}.5) / .05^{2}}{1 + (2.58^{2} * .5\,(1\text{-}.5) /.05^{2} * 2016)}$$

$$\frac{6.6564 * .25 / .0025}{1 + (6.6564 * .25 / 1.55)}$$

$$\frac{940.64}{1.5.0736}$$

Sample size = 720

Table 2: Checking Reliability and Validity

		Reliability	
Indicator Reliability	Outer loadings numbers	PLS → Calculation Results → Outer Loading	Square each of the outer loading to find the indicator reliability value. *0.70* or higher is preferred. If it is exploratory research, *0.4* or higher is acceptable (Hulland, *2016*).
Internal Consistency Reliability	Reliability numbers	PLS → Quality Criteria → Overview	Composite reliability should be 0.7 or higher. If it is an exploratory research, *0.6* or higher is acceptable.**(Bagozzi and Yi, *1988*)**
		Validity	
Convergent validity	AVE numbers	PLS → Quality Criteria → Overview	It should be 0.5 or higher(Bagozzi and Yi, 1988)
Discriminant validity	AVE numbers and Latent Variable Correlation	PLS → Quality Criteria → Overview (for the AVE number as shown above) PLS → Quality Criteria → Latent Variable Correlations"	**(*Fornell and Larcker, 1981*)** suggest that the square root of AVE of each latent variable should be greater than the correlations among the latentvariables

Missing data treatment is vital in the analysis, with Smart PLS being one of the statistical tools used. Data will only run if values are included (Hair, 2018). Alternatively, SPSS can be used to treat missing data by replacing missing values with a mean or median of nearby points or via linear interpolation. However, as indicated in the case processing summary in Table 3, all data of the study are available and no missing data has been detected.

Table 3: Case processing summary

		Cases			
		Valid		Missing	
	Factors	**N**	**%**	**N**	**%**
V1	Volume	720	100%	0	0%
V2	Velocity	720	100%	0	0%
V3	Value	720	100%	0	0%
V4	Variety	720	100%	0	0%
V5	Veracity	720	100%	0	0%
V6	Variability	720	100%	0	0%
V7	Visualization	720	100%	0	0%

Reliability is a measure that ensures the measure is reliable. The most frequently used reliability coefficient is Cronbach's alpha, which has been used in this study to measure the reliability of variables in a summated scale. Table 4 below shows Cronbach's alpha values computed for all the variables, which revealed that the highest value generated was that of organizational cynicism with a value of 0.927. All the variables were above the minimum level of 0.70 as recommended by Nunnally Jr. (1970) and Nunnally & Bernstein, 2018 for good reliability.

Reliability Analysis and Inter Item Variances and Means

Table 4: Reliability Statistics BDA variables

	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Items
Overall	.802	.798	20

Table 5: Summary Item Statistics BDA variables

	Particulars	Mean	Range	Maximum/ Minimum	Variance	N of Items
Participants	Item Means	4.692	.090	1.019	.003	20
	Item Variances	1.664	.226	1.143	.015	20
	Inter-Item Co-variances	.210	.782	-5.508	.131	20

4.2. Correlation Analysis

The Pearson correlation technique was done to determine how variables are related with one another and whether an observed variable has perfect covariance with any other variable that would be observed in this study. A summary of the correlation is represented in the table below. This two-variable relationship was significant at the p-value of 0.01. It, therefore, depicts a positive and moderate to strong variables.

Table 6: Correlation Matrix

			V1	V2	V3	V4	V5	V6	V7
Factors	V1	Pearson Correlation	1	.867**	.785	.458**	.465	.628**	.875
	V2	Pearson Correlation	.872**	1	.785	.638**	.665	.638**	.875
	V3	Pearson Correlation	-.045	.866**	1	.658**	.785	.758**	.475
	V4	Pearson Correlation	.398**	.865**	-.172**	1	.546	.673	.774
	V5	Pearson Correlation	.767**	.615	.628**	.675	1	.777	.981
	V6	Pearson Correlation	.767**	.585	.638**	.675	.567	1	.567
	V7	Pearson Correlation	.967**	.785	.758**	.675	.789	.878	1

4.3. Regression Analysis

A regression analysis was performed to examine the connection between model predictors and adaptability intentions. The study involved hypothesis testing and a reiteration of the analysis assumptions. The aim was to forecast predictor values for the adaptability of smart device users, and dependent variable.

Table 7: Model Summary

					Change Statistics					
Model	R	R Squaree	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin - Watson
1	.923	.573	.261	1.11982	.273	22.503	6	35 9	.000	1.113

Table 8: ANOVA

Model		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
1	Regression	169.312	6	28.219	22.503	.000b
	Residual	450.182	359	1.254		
	Total	619.494	365			

a. Dependent Variable: investment Decisions

b. Predictors: (Constant), Predictors of BDA characteristics

Table 9: Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	.404	.493		41.820	.013		
	Volume	.019	.079	.012	61.245	.006	.794	1.259
	Value	.120	.044	.140	34.725	.007	.371	1.298
	Velocity	.018	.098	.010	22.185	.003	.677	1.478
	Variety	.019	.079	.012	21.245	.006	.794	1.259
	Veracity	.120	.044	.140	11.725	.007	.771	1.298
	Variability	.120	.044	.140	24.725	.007	.671	1.298
	Visualization	.018	.098	.010	26.185	.003	.677	1.478

a. Dependent Variable: FDI Decisions

4.4. Hypotheses Testing

This part of the report contains the structural model and the test result of hypotheses. In this section, the main point is the testing of hypotheses concerning the central mediating effect. For that purpose, a multiple regression approach of PLS path modeling was employed to test main effects while mediating effects were assessed by the help of the bootstrapping technique in PLS. Chin et al. (2016) used PLS bootstrapping techniques in this study to conduct path modeling. It used 720 cases and 5000 bootstrapped samples to test the hypotheses of the current study. The 5000 bootstrapped samples were utilized to ensure that all the model parameters had empirical sampling distribution and standard error was obtained. Using the same method stated above, the path coefficients were estimated using t-statistics. The t-value was subjected to one-tailed distribution in order to find the level of significance (Chin et al., 2016). In a one-tailed statistical test, any t-value that is greater or equal to 2.326 at a 1% level of significance whereas at 5% it is greater or equal to 1.645 whereas at 10% level of significance it is greater or equal to 1.282 any lesser value are considered as insignificant (Churchill Jr, 1979).

Table 10 shows the structural model results for direct relationships after path analysis and bootstrapping with resamples *5000* to ensure more stability and consistency of results (F. et al. et al., *2018*). The following results have been observed for each direct relationship separately.

Table 10: Results of Path Coefficients (Direct Relationship)

Direct Hypotheses	Beta	SD	T-Statistics	P-Values	Decision
H01: No substantial correlation exists between the volume of data in Big Data Analytics and its impact on investment decision-making.	0.049	2.33	22.384	0.043	Significant
H02: There exists no notable correlation between the speed of information processing in Big Data Analytics and its impact on investment decision-making.	0.071	2.33	21.582	0.021	Significant
H03: There is no notable correlation between the diversity of information in Big Data Analytics and its impact on investment decision-making.	0.049	2.33	34.386	0.001	Significant
H04: There is no significant relationship between veracity of information in Big Data Analytics and decision making for the investment decision.	0.096	2.33	45.688	0.000	Significant
H05: There is no noteworthy correlation between the value of information in the realm of Big Data Analytics and its impact on investment decision-making.	0.088	2.33	23.882	0.031	Significant
H06: There exists no substantial correlation between the variability of information within the realm of Big Data Analytics and its impact on investment decision-making.	0.076	2.33	11.587	0.021	Significant
H0 7: Visualizing Information in Big Data Analytics Does Not Significantly Impact Investment Decision-Making	0.036	2.33	21.587	0.011	Significant
H08: ICT Positively Correlates with Investment Decision-Making	0.056	2.33	33.587	0.002	Significant

4.5 GMM Results

Our variables' robustness is checked based on the GMM estimator, and the data has been presented in Table 11. The two-step GMM results are normally considered to be very accurate compared to one-step system GMM. The GMM system structures instruments through degree and first difference equations. Sample size information is quite detailed. However, it increases the instruments available. A two-step GMM technique has effective power and provides for informative and acceptable initial conditions even in the presence of endogenous regressors. Lastly, the Hansen J-test and second-order autocorrelation tests are used to assess the quality of the instrument based on equations of level and difference and validity of over-identification constraint. Two-step system GMM depicts that FDI has a negative correlation with GDP per capita. This means that Decision Making lowers BDA in the manufacturing sector of Pakistan. Assuming all other variables are constant, a one percent rise in Decision-making reduces the influence of BDA by 0.376 percent. Except for V1 and V2, all other variables have a positive correlation with Decision making.

Table 11 Results of Unit Root Test

Variables	Im, Pesa ran and Shin W-stat				A.D.F. Fisher Chi-square				Results
	Level		First Difference		Level		First Difference		
	T-stat	P-value	T-stat	P-value	T-stat	P-value	T-stat	P-value	
LNGDP (1)	2.941	0.998	–2.481	0.007	2.826	0.985	21.335	0.019**	I (1)
LN(FDI) (2)	–1.301	0.097	–6.122	0.000	15.890	0.103	51.534	0.000***	I (1)
LN(FTS) (3)	–0.401	0.344	–4.556	0.000	20.886	0.022	40.447	0.000***	I (1)
LN(FBS) (4)	–12.009	0.000	–13.122	0.000	58.695	0.000	52.453	0.000***	I (0)
LN(ISI) (5)	1.620	0.947	–3.295	0.001	20.631	0.024	30.821	0.001***	I (1)
LN(SIS) (6)	3.942	1.000	–3.894	0.000	2.141	0.995	30.748	0.001***	I (1)
LN(MCS) (7)	0.704	0.759	46.767	0.000	17.185	0.070	45.484	0.000***	I (1)
LN(TR) (8)	1.217	0.888	–4.452	0.000	7.945	0.634	35.588	0.000***	I (1)

Note: *** means significant at 10%, ** at 5%, and * at 1% level of significance.

Table 11: GMM Estimations Results

LnGDP	Coef.	Std. Err.	z	P > z	95% Conf.	Interval
LNV1	–0.103	0.035	–2.940	0.003**	–0.172	–0.035
LNV2	–0.376	0.033	–11.490	0.000***	–0.441	–0.312
LNV3	0.289	0.134	2.150	0.031**	0.026	0.553
LNV4	0.517	0.766	6.700	0.000***	0.367	0.668
LNV5	0.034	0.303	1.110	0.266	–0.026	0.093
LNV6	0.568	0.360	15.760	0.000***	0.497	0.638
LNV7	–0.151	0.703	–2.150	0.031**	–0.289	–0.014
_cons	2.303	0.149	15.490	0.000***	2.011	2.594

Note: *** means significant at 10%, ** at 5%, and * at 1% level of significance.

4. Conclusion

All null hypotheses within the study have been meticulously scrutinized and subsequently rejected, while the study's hypotheses have been robustly affirmed, as evident in the provided tables. In the context of the first section of the study, it is noteworthy that the *t*-values corresponding to various factors such as volume, velocity, variety, veracity, variability, visualization, and value significantly exceed the critical t-table values. This compelling finding underscores a profound and statistically significant relationship between these factors and the investment decision-making process.

Furthermore, examining skewness and kurtosis within the data reveals a noteworthy characteristic: the data distribution closely adheres to the normal distribution. This statistical insight bolsters the rational basis for rejecting the null hypotheses and, in turn, substantiates the reliability of the study's findings.

Delving into the descriptive statistics in the tables, a succinct summary emerges, encapsulating the mean ratings' characteristics. These characteristics pertain to the pivotal role of Big Data Analytics in influencing the decision-making processes within the manufacturing industry of Pakistan. The implications of these statistics are profound, shedding light on the transformative potential of Big Data Analytics in a critical industrial sector.

The results of the regression analysis and change statistics underscore the efficacy of the study's predictive model. Notably, the model stands out as suitable for prediction and significantly distinct, with p-values well below the 0.50 threshold. This finding reinforces the notion of a linear relationship characterizing the decision-making process, with the '*7Vs*' factors exerting a consistent influence.

The scatter plots provide invaluable insight for a visual representation of the study's findings. They vividly portray the dataset values concerning decision-making and Big Data, revealing a remarkable alignment with a firm agreement with the study's hypotheses. This visual confirmation bolsters the study's credibility and underscores the robustness of the findings.

Lastly, the study ventures beyond the immediate scope of decision-making and extends its exploration into organizational performance within the manufacturing industry, drawing insightful correlations with investment decisions. This multifaceted approach enriches the study's depth and offers valuable insights into the interplay of these critical aspects within the industrial landscape.

References

1. A. Choudhury and D. Gupta, *A Survey on Medical Diagnosis of Diabetes Using Machine Learning*. Springer Singapore. doi: 10.1007/978-981-13-1280-9.
2. Khan, S. R., Raza, A., Shahzad, I., & Ijaz, H. M. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).
3. Bilal, Omair, Asif Raza, and Ghazanfar Ali. "A Contemporary Secure Microservices Discovery Architecture with Service Tags for Smart City Infrastructures." VFAST Transactions on Software Engineering 12, no. 1 (2024): 79-92.
4. S. Abhari, S. R. N. Kalhori, M. Ebrahimi, and H. Hasannejadasl, “Artificial Intelligence Applications in Type 2 Diabetes Mellitus Care : Focus on Machine Learning Methods,” vol. 25, no. 4, pp. 248–261, 2019.
5. H. Kaur and V. Kumari, “Predictive Modelling and Analytics for Diabetes using a Machine Learning Approach,” *Appl. Comput. Informatics*, no. December, 2018, doi: 10.1016/j.aci.2018.12.004.
6. Raza, A., Salahuddin, & Inzamam Shahzad. (2024). Residual Learning Model-Based Classification of COVID-19 Using Chest Radiographs. Spectrum of Engineering Sciences, 2(3), 367–396.
7. A. Mujumdar and V. Vaidehi, “ScienceDirect ScienceDirect ScienceDirect ScienceDirect Diabetes Prediction using Machine Learning Aishwarya Mujumdar Diabetes Prediction using Machine Learning Aishwarya Mujumdar Aishwarya,” *Procedia Comput. Sci.*, vol. 165, pp. 292–299, 2019, doi: 10.1016/j.procs.2020.01.047.
8. S. Brian and R. R. B. Pharmd, “Prediction of Nephropathy in Type 2 Diabetes: An Analysis of the ACCORD Trial applying Machine Learning Techniques,” no. 317, doi: 10.1111/cts.12647.
9. Faruque, “Performance Analysis of Machine Learning Techniques to Predict Diabetes Mellitus,” *2019 Int. Conf. Electr. Comput. Commun. Eng.*, pp. 1–4, 2019.
10. Asif, S., Wenhui, Y., ur-Rehman, S., ul-ain, Q., Amjad, K., Yueyang, Y., ... & Awais, M. (2024). Advancements and Prospects of Machine Learning in Medical Diagnostics: Unveiling the Future of Diagnostic Precision. Archives of Computational Methods in Engineering, 1-31.
11. Asif, S., Zhao, M., Li, Y., Tang, F., Ur Rehman Khan, S., & Zhu, Y. (2024). AI-Based Approaches for the Diagnosis of Mpox: Challenges and Future Prospects. Archives of Computational Methods in Engineering, 1-33.
12. J. Li *et al.*, “International Journal of Medical Informatics Establishment of noninvasive diabetes risk prediction model based on tongue features and machine learning techniques,” *Int. J. Med. Inform.*, vol. 149, no. August 2020, p. 104429, 2021, doi: 10.1016/j.ijmedinf.2021.104429.
13. S. G. Azevedo *et al.*, “System-Independent Characterization of Materials Using Dual-Energy Computed Tomography,” *IEEE Trans. Nucl. Sci.*, vol. 63, no. 2, pp. 341–350, 2016, doi: 10.1109/TNS.2016.2514364.
14. B. Pranto, S. M. Mehnaz, E. B. Mahid, and I. M. Sadman, “Evaluating Machine Learning Methods for Predicting Diabetes among Female Patients in Bangladesh,” 2020, doi: 10.3390/info11080374.
15. Shahzad, I., Khan, S. U. R., Waseem, A., Abideen, Z. U., & Liu, J. (2024). Enhancing ASD classification through hybrid attention-based learning of facial features. Signal, Image and Video Processing, 1-14.
16. Khan, S. U. R., Raza, A., Shahzad, I., & Ali, G. (2024, October). Enhancing Concrete and Pavement Crack Prediction through Hierarchical Feature Integration with VGG16 and Triple Classifier Ensemble. In 2024 Horizons of Information Technology and Engineering (HITE) (pp. 1-6). IEEE.

17. Khan, S. R., Asif Raza, Inzamam Shahzad, & Hafiz Muhammad Ijaz. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).
18. Y. Sun and D. Zhang, "Machine Learning Techniques for Screening and Diagnosis of Diabetes : a Survey," vol. 3651, pp. 872–880, 2019.
19. B. G. Choi, S. Rha, S. W. Kim, J. H. Kang, J. Y. Park, and Y. Noh, "Machine Learning for the Prediction of New-Onset Diabetes Mellitus during 5-Year Follow-up in Non-Diabetic Patients with Cardiovascular Risks," vol. 60, no. 2, pp. 191–199, 2019.
20. Khan, S.U.R.; Raza, A.;Waqas, M.; Zia, M.A.R. Efficient and Accurate Image Classification Via Spatial Pyramid Matching and SURF Sparse Coding. Lahore Garrison Univ. Res. J. Comput. Sci. Inf. Technol. 2023, 7, 10–23.
21. Farooq, M.U.; Beg, M.O. Bigdata analysis of stack overflow for energy consumption of android framework. In Proceedings of the 2019 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 1–2 November 2019; pp. 1–9.
22. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Li, X. (2024). Optimized Deep Learning Model for Comprehensive Medical Image Analysis Across Multiple Modalities. Neurocomputing, 129182.
23. F. Farrokhi *et al.*, "Investigating risk factors and predicting complications in deep brain stimulation surgery with machine learning algorithms," *World Neurosurg.*, 2019, doi: 10.1016/j.wneu.2019.10.063.
24. M. Waqas, Z. Khan, S. U. Ahmed and Asif. Raza, "MIL-Mixer: A Robust Bag Encoding Strategy for Multiple Instance Learning (MIL) using MLP-Mixer," 2023 18th International Conference on Emerging Technologies (ICET), Peshawar, Pakistan, 2023, pp. 22-26
25. Farooq, M. U., Khan, S. U. R., & Beg, M. O. (2019, November). Melta: A method level energy estimation technique for android development. In 2019 International Conference on Innovative Computing (ICIC) (pp. 1-10). IEEE.
26. Khan, S. U. R., & Asif, S. (2024). Oral cancer detection using feature-level fusion and novel self-attention mechanisms. Biomedical Signal Processing and Control, 95, 106437.
27. Raza, A.; Meeran, M.T.; Bilhaj, U. Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers. VFAST Trans. Softw. Eng. 2023, 11, 80–92.
28. Dai, Q., Ishfaque, M., Khan, S. U. R., Luo, Y. L., Lei, Y., Zhang, B., & Zhou, W. (2024). Image classification for sub-surface crack identification in concrete dam based on borehole CCTV images using deep dense hybrid model. Stochastic Environmental Research and Risk Assessment, 1-18.
29. Raza, A., & Meeran, M. T. (2019). Routine of Encryption in Cognitive Radio Network. Mehran University Research Journal of Engineering and Technology [p-ISSN: 0254-7821, e-ISSN: 2413-7219], 38(3), 609-618.
30. Khan, S.U.R.; Asif, S.; Bilal, O.; Ali, S. Deep hybrid model for Mpox disease diagnosis from skin lesion images. Int. J. Imaging Syst. Technol. 2024, 34, e23044.
31. Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X.; Zhu, Y. GLNET: Global–local CNN's-based informed model for detection of breast cancer categories from histopathological slides. J. Supercomput. 2023, 80, 7316–7348.
32. Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X. Hybrid-NET: A fusion of DenseNet169 and advanced machine learning classifiers for enhanced brain tumor diagnosis. Int. J. Imaging Syst. Technol. 2024, 34, e22975.
33. Salahuddin, Syed Shahid Abbas, Prince Hamza Shafique, Abdul Manan Razzaq, & Mohsin Ikhlaq. (2024). Enhancing Reliability and Sustainability of Green

Communication in Next-Generation Wireless Systems through Energy Harvesting. Journal of Computing & Biomedical Informatics.

34. IoannisKavakiotis, Olga Tsave, Athanasios Salifoglou, and NicosMaglaveras, “Machine Learning and Data Mining Methods in Diabetes Research”, Computational and Structural Biotechnology Journal, vol. 15, pp. 104–116, 2017
35. Khan, U. S., & Khan, S. U. R. (2024). Boost diagnostic performance in retinal disease classification utilizing deep ensemble classifiers based on OCT. Multimedia Tools and Applications, 1-21.
36. Khan, M. A., Khan, S. U. R., Haider, S. Z. Q., Khan, S. A., & Bilal, O. (2024). Evolving knowledge representation learning with the dynamic asymmetric embedding model. Evolving Systems, 1-16.
37. Raza, A., & Meeran, M. T. (2019). Routine of encryption in cognitive radio network. Mehran University Research Journal of Engineering & Technology, 38(3), 609-618.
38. Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. Computational Economics, 1-32.
39. Khan, U. S., Ishfaque, M., Khan, S. U. R., Xu, F., Chen, L., & Lei, Y. (2024). Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images. Frontiers of Structural and Civil Engineering, 1-17.
40. Hekmat, A., Zhang, Z., Ur Rehman Khan, S., Shad, I., & Bilal, O. (2025). An attention-fused architecture for brain tumor diagnosis. Biomedical Signal Processing and Control, 101, 107221. https://doi.org/https://doi.org/10.1016/j.bspc.2024.107221

I.