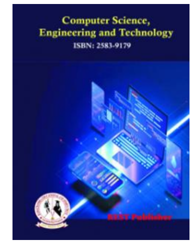
Computer Science, Engineering and Technology

Vol: 3(2), June 2025

REST Publisher; ISSN: 2583-9179

Website: https://restpublisher.com/journals/cset/

DOI: https://doi.org/10.46632/cset/3/2/3



Future Trends in AI: Data Management and Analysis Using SPSS Methodology

Lal Kishore Kumar, Vikash Kumar Singh, Rishi Sharma, *Ranjan Kumar Mishra

Netaji Subhas University Jamshedpur, Jharkhand, India.

*Corresponding Author Email: ranjanlnmi4u@gmail.com

Abstract: *Introduction: This study offers a thorough analysis of deep learning and artificial intelligence from 1961 to 2018, providing information about the underlying mechanics, industrial applications, and future developments. The study aims to help researchers and practitioners understand the evolution, challenges, and opportunities associated with AI-driven innovations. Research Significance: This study's importance stems from its comprehensive exploration of the impact of AI on various industrial domains. AI-driven data analytics, predictive maintenance, and decision-making systems are reshaping industrial landscapes by improving operational efficiency and reducing costs. By analyzing past trends and current developments, this study offers insightful information for upcoming AI applications. It also draws attention to difficulties associated with to AI adoption, such as data security, algorithmic transparency, and ethical considerations, ensuring a holistic perspective for stakeholders in academia and industry. Methodology: SPSS Statistics is a powerful software tool utilized for to analyze data in a number of domains, like as social sciences, healthcare, marketing, and education. It offers an extensive collection of statistical tools for organizing, evaluating, and interpreting data. SPSS allows users to perform a wide range of analyses, such as descriptive statistics, regression, ANOVA, factor analysis, and hypothesis testing. Its sophisticated data manipulation features and user-friendly interface make it popular among researchers and analysts. SPSS also supports the creation of charts and reports, aiding in the presentation of data-driven insights. Input Parameters: Industry Sector, Data Management Approach, AI Integration Level, Primary AI Use Case, Data Privacy Strategy, Adoption of AI Ethics Framework. Evaluation Parameters: Data Processing Efficiency, AI Model Accuracy, Scalability & Flexibility, Regulatory Compliance, Business Value Impact. Reliability Statistics measure the consistency of a dataset or survey instrument. One important Cronbach's Alpha is a measure of internal reliability. The Alpha of Cronbach's value of 0.967 in this instance indicates a high degree of dependability, indicating that the dataset's items consistently yield reliable findings. This high The standardized Cronbach's Alpha further supports the degree of intrinsic coherence of 0.974. The data can be regarded as extremely dependable for research purposes after five items have been examined.*

Key Words: *Big data analytics, predictive analytics, artificial intelligence, machine learning and Industry 4.0.*

1. INTRODUCTION

An important factor in the advancement of industry is artificial intelligence (AI), which also acts as a catalyst for the incorporation of new technologies like cloud computing, block chain, the Internet of Things, and graphics processing units into the big data and Industry 4.0 framework. This essay provides an in-depth analysis of AI and deep learning over the period from 1961 to 2018. Through multi-faceted systematic analysis, it provides valuable insights for researchers and practitioners, covering aspects from foundational algorithms to real-world applications, from basic algorithms to industrial innovations, and from the current landscape to future developments. [2] Upcoming Developments in Big Data Analytics and SQL Databases, it looks at how AI-driven automation affects database administration, emphasizing how it can boost efficiency and cut expenses. [3] The main data sources and motivating reasons for data analytics adoption are examined in this study, along with how AI and machine learning may help networks become proactive, self-aware, adaptable, and prescriptive. Furthermore, a variety of data analytics-based network design and optimization techniques are offered. An examination of the difficulties and advantages of incorporating Big data analytics, AI, and machine learning into future communication systems is presented in the paper's conclusion. [4] Industry 4.0, often known as the Fourth Industrial Revolution, has ushered in a wave of disruptive technologies that have significantly reshaped existing systems across various industries and led to innovative business models and processes. At the same time, decision-making is increasingly being automated and delegated to computer systems, driven by the rise of big data analytics (AI) and artificial intelligence (AI). [5] By uncovering hidden patterns and forecasting future trends, Big data, artificial intelligence, and machine learning techniques have demonstrated encouraging outcomes in a variety of corporate and industrial applications. The ability to extract intricate nonlinear structures from large datasets has been demonstrated by recent sophisticated models, like artificial neural networks for deep learning. [6] The conference on leading authorities and scholars from a range of disciplines, including data science, are brought together by big data and artificial intelligence in education, cognitive neuroscience, psychology, education, and artificial intelligence, exchanged ideas and expertise, and served as the basis for this work. The paper's structure is as follows: An overview of recent developments in artificial intelligence and big data in education is given at the outset, and then important issues and new trends are examined. Lastly, conclusions and suggestions for future research are made in light of the conversations. [7] Large amounts of operational data that are difficult for humans to analyze are being produced by the expansion and development of Big Data infrastructures in various industries. Data analytics that incorporate algorithms that can make correct and autonomous decisions is becoming more and more popular. Indeed, the future industry is expected to have self-learning, intelligent robots and seamless communication between cyber-physical systems. [8] In many disciplines, including science, engineering, and business, mathematics is essential. Statistics in particular are becoming more and more popular. The increasing availability of data creates new opportunities in a variety of fields, from fundamental ones like descriptive statistics and data visualization to more complex ones like data analysis. [9] Their research offers a graphic representation of the development of business intelligence (BI), emphasizing significant shifts and patterns over the previous several decades. The study, which takes a social network viewpoint, reveals a plateau in recent years' BI-related papers, suggesting that the subject has developed with solid core ideas and tools. A move toward more complex and sophisticated methods of data analysis and interpretation within BI is also shown by the study's identification of big data and machine learning as new key points. [10] Public acceptance, security standards, and regulatory criteria must all be met by a cybersecurity framework for impact assessment that uses machine learning (ML) in cyber risk studies. In light of these considerations, the study incorporates managerial, policy, and impact analysis recommendations while proactively taking into account computer science advancements to create and enhance frameworks for machine learning-driven supply chain cyber risk assessments. [11] John McCarthy, a computer scientist, originally used the term "artificial intelligence" in the middle of the 1950s. "Application of Artificial Intelligence to the Pattern-Cutter Problem" by Bardin Burkholder, which was published in Operations Research in 1963, is the earliest known study that connects AI to business. The Web of Science (WoS) indexes this article under the headings of Management Science, Operations Research, and Business and Economics. [12] The same basic component of information systems (IS) design is described by a number of terminology, including conceptual modeling, data modeling, and information modeling, depending on the viewpoint and degree of abstraction. We refer to "data modeling" in this study as a general term that includes modeling approaches applied at every stage of the IS development life cycle. This covers data-centric elements at all abstraction levels as well as conceptual modeling. The model's development aligns with the selected development paradigm's tenets. [13] By increasing diagnostic precision, optimizing workflows, enhancing clinical judgment, and improving patient outcomes, Artificial intelligence and machine learning developments are revolutionizing pathology and medicine. With their contributions to automated image processing, biomarker discovery, drug development, clinical trials, and predictive analytics, these technologies are also becoming more and more important in pathology research. Key trends also include the use of multimodal and multivalent AI to use a variety of data sources, the growth of virtual education for training and simulation, the acceleration of translational research, and the integration of machine learning (ML) functions for managing models in clinical contexts. [14] To guarantee the quality and applicability of the sources chosen for study, we meticulously crafted our inclusion and exclusion criteria. Peer-reviewed scholarly publications, conference proceedings, and official reports that particularly address AI legitimacy in data management and analytics met our requirements. We methodically grouped research during the synthesis process according to their methodology, thematic relevance, and the particular legitimacy issues they encountered. This methodical technique makes it possible to integrate various viewpoints from the literature in a logical and orderly manner. [15] In the twenty-first century, predictive analytics has become an essential part of decision-making based on data. By applying machine learning techniques, statistical models, and historical data, it enables firms to make exceptionally accurate predictions about the future. Despite having its roots in early statistical analysis, its capabilities and uses have greatly grown with the rise of large data and sophisticated computerization. [16] Due to their broad potential These new technologies have drawn increasing attention in healthcare research and practice as a means of enhancing healthcare quality. The application of artificial intelligence and big data analytics in healthcare has been the subject of numerous studies, however the literature is still dispersed. To obtain a thorough comprehension of the potential of these technologies, a systematic mapping study was conducted. This approach was considered appropriate given the abundance and diversity of existing research in this field. [17] Combining machine learning (ML) with Another Important Utilizing innovative technologies such as the Internet of Things (IoT) and big data analytics is popular. All-inclusive water management technologies capable of addressing the challenges of modern wastewater treatment could be produced by this integration. The interaction between these technologies should be investigated in future studies, and best methods for using both of them together in wastewater management should be established. [18] Predictive maintenance (PdM) state-of-the-art (SOTA) includes a number of crucial strategies: (i) data-driven methods like big data analytics, data mining, data visualization, and predictive analytics; (ii) AI-driven PdM techniques like statistical control, machine learning, deep learning, and statistical modeling; (iii) vibration and thermal analysis; (iv) augmented reality, virtual reality, mixed reality, and their advanced variants with digital intelligent assistants; (v) prescriptive maintenance strategies; (vi) technologies like edge and cloud computing, IoT, federated learning, and block chain; (vii) energy-driven approaches; and (viii) predictive and health management (PHM) methods. [19] Large volumes of digital data are generated by the pharmaceutical and medical sectors, including data from genetic and medical datasets, wearable sensors, and electronic medical records. Pharmaceutical businesses are better able to create individualized medicine regimens and treatment programs for the best results when they have more patient insights. At the same time, businesses are implementing cutting-edge artificial intelligence, which is utilizing machine learning and deep learning techniques to collect and analyze data thanks to advancements in processing power.

2. MATERIALS & METHODS

Input Parameters:

Industry Sector: A broad category of businesses and organizations operating within a specific field, such as healthcare, finance, retail, manufacturing, education, or government. Each sector has unique challenges, regulations, and technological needs.

Data Management Approach: The strategy used by organizations to collect, store, process, and manage data. Common approaches include cloud-based solutions for scalability, on- premise storage for security, and hybrid models that balance both.

AI Integration Level: The extent to which artificial intelligence is incorporated into business operations, from basic automation to advanced machine learning models that optimize decision-making and predictive analysis.

Primary AI Use Case: The main application of AI in an organization, such as predictive analytics for forecasting trends, chat bots using natural language processing and image identification using computer vision, fraud detection in finance, or process automation.

Data Privacy Strategy: A company's policies and practices designed to safeguard sensitive data, ensure user confidentiality, and comply with data protection laws like GDPR, HIPAA, or CCPA to prevent breaches and misuse.

Adoption of AI Ethics Framework: The degree to which an organization follows ethical guidelines when developing and deploying AI, ensuring fairness, transparency, accountability, and bias mitigation in AI-driven decision-making.

Evaluation Parameters:

Data Processing Efficiency: The ability of a system to handle, transform, and analyze large datasets quickly and accurately, optimizing computing power, minimizing processing time, and ensuring timely insights for business operations.

AI Model Accuracy: A measure of how precisely an AI system predicts, classifies, or analyzes data. High accuracy means the model makes correct decisions consistently, while lower accuracy may require optimization and retraining.

Scalability & Flexibility: The capability of an AI or data system to expand and adapt based on business needs, handling growing workloads, integrating new data sources, and adjusting to evolving market conditions.

Regulatory Compliance: Adherence to industry standards and legal frameworks governing AI and data management, ensuring organizations meet data security, privacy, and ethical requirements to avoid legal penalties.

Business Value Impact: The overall effect AI has on an organization's efficiency, profitability, innovation, and competitive advantage, demonstrating tangible benefits such as cost reduction, revenue growth, and improved decision-making.

SPSS Method: To deal with missing values, the SPSS factor procedure offers solutions such mean replacement, pairwise deletion, and list wise deletion; nevertheless, these techniques have well-established drawbacks. According to Graham (2009), it is more practical to employ expectation-maximization (EM) covariance or correlation matrices as the analysis's input. Users of SPSS can create EM-based means, standard deviations, correlation matrices, and covariance matrices by using the Missing Values Analysis (MVA) module. The /MATRIX subcommand, which is absent from the MVA procedure, prohibits the direct export of EM correlations into a matrix dataset that may be utilized in the FACTOR and RELIABILITY procedures. [21] While the SAS program would need to be modified for older versions, the SPSS syntax is backward compatible with earlier versions. Opening the dataset for analysis and then carrying out the regression analyses are the first steps in the comparable logical structure of both syntax files. One significant distinction is that the SPSS software generates three related datasets with results from Each of the three regression analyses reflects a distinct model. [22] The current study's goal is to illustrate the value of Individual Growth Curve (IGC) modeling in this setting utilizing SPSS software. The methods for developing, testing, and comparing models are described in this work. Through the analysis of a five-year longitudinal dataset, we looked at how high school pupils changed after taking part in a constructive youth development program. [23] In clinical research, single-case experimental designs are useful instruments for evaluating the progress of particular clients. However, methodological issues, such as the inability to use standard statistical procedures, may limit their broad acceptance. Using the publicly accessible SPSS program, this article offers a data analysis method for assessing single-case (single-symptom) data. [24] One of the main obstacles to the broad usage of propensity score approaches is the dependence on specialized software, as many social scientists rely on SPSS for data analysis. In this study, different propensity score matching approaches are used in SPSS. In particular, researchers can employ these techniques with an intuitive point-and-click interface in a customized SPSS dialog. In addition to providing customizable nearest-neighbor matching options, such as calipers, common support regions, matching with or without substitution, and one-to-many matching, the software provides propensity score calculation through logistic regression. It also generates graphical displays and comprehensive balancing statistics. [25] Through benchmarking applications, this paper illustrates the efficacy of two essential components of the LINEAR method in SPSS Statistics. It encourages the use of the LINEAR program as a substitute for the conventional REGRESSION technique, particularly when working with very big datasets where time and effort constraints make manual data processing impracticable. The article also examines the theoretical aspects of this novel process, possible areas for development, and its combinatorial potential as a line of inquiry. [26] The fundamental idea of principal component regression, multi collinearity detection codes, and the process for identifying the "best" equation are all well covered in this paper. It covers all relevant computations and operations, such as linear regression, factor analysis, descriptive statistics, imputed variables, and bivariate correlations, and provides an example of How to use SPSS 10.0 for main component regression analysis. When used in SPSS, principal component regression analysis is emphasized as a successful strategy for reducing the impact of multi collinearity, offering a simplified, accurate, and efficient statistical solution. [27] This study used five synthetic datasets with preset subgroups to evaluate two latent class analysis (LCA) approaches (Latent Gold and ALMO) with a distance-based clustering method (SPSS Two Step). Two Step is a hybrid strategy that selects the best subgroup model by classifying individuals based on a distance metric using a technique akin to LCA. Two Step performs better than conventional hierarchical clustering methods, according to earlier studies. discovered that Latent Gold was the most effective in determining the appropriate number of subgroups, whereas Two Step was the least effective of the three approaches. Specifically, Two Step had trouble with datasets that had both nominal and interval data. [28] Since SPSS Basic program already has the first two techniques list wise elimination and pairwise elimination, also known as complete-case and available-case analysis they are not the primary benefits of the MVA module. Actually, a far more complete implementation is offered by SPSS Basic, which offers list wise and pairwise estimates for a range of models, such as factor analysis and regression. [29] SPSS ALSCAL generates a participant space and a fixed stimulus space for this kind of analysis, reflecting the various weights that each participant gives to the dimensions in the shared stimulus space and the models that were applied to their data. A multi-matrix comprising either matrix-conditioned or non-conditioned data is used in this method. The individual distances between any two participants in weighted multidimensional scaling (WMDS) do not have to be coupled by a monotonic or linear function. [30] It is more difficult to handle more complicated analyses with Excel or Calc, like multiple linear regression. The techniques described for SPSS and R, on the other hand, apply naturally and intuitively to these circumstances. Reliable methods for R and SPSS also perform well for multiple regression, enhancing the results by adding more predictors and pertinent data. [31] By choosing "Analysis ➔ General Linear Model ➔ Univariate..." A factorial design ANOVA should be conducted in SPSS using the general linear model. The findings show that the kind of illness and treatment have a statistically significant interaction ($p < 0.001$) indicating that both factors work together to influence the dependent variable (pain score). Consequently, each factor requires a basic main effects test. Since these factors include independent samples, the necessary samples for an independent sample one-way ANOVA can be filtered out using SPSS's "Section File" option located under the "Data" menu.

3. RESULT AND DISCUSSION

TABLE 1. Reliability Statistics

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.967	.974	5

The reliability statistics presented in Table 1 indicate the internal consistency of a measurement scale using Cronbach's Alpha. This statistical measure assesses how closely related a set of items is within a test or survey. In this case, the Cronbach's Alpha value is 0.967, suggesting an excellent level of reliability. A reliability coefficient above 0.9 is generally considered high, meaning the items in the scale are very consistent with each other. Additionally, when computed based on standardized items, the Cronbach's Alpha increases slightly to 0.974, reinforcing the strong reliability of the scale. The table also notes that the scale consists of five items. This indicates that all five items contribute effectively to measuring the intended construct, with minimal measurement error. Such high. reliability suggests that the instrument used is dependable for research or assessment purposes. However, while a high Cronbach's Alpha is desirable, values that are too high (close to 1) may suggest redundancy among items, meaning some questions might be too similar. Researchers should consider whether all five items are necessary or if some could be removed to maintain efficiency while preserving validity.

TABLE 2. Reliability Statistic individual

Item-Total Statistics	
	Cronbach's Alpha if Item Deleted
Data Processing Efficiency	.958
AI Model Accuracy	.961
Scalability & Flexibility	.960
Regulatory Compliance	.953
Business Value Impact	.959

Table 2 presents the Item-Total Statistics, specifically showing the Cronbach's Alpha if each item is deleted. This analysis helps determine whether removing a particular item would significantly affect the overall reliability of the scale. The reported Cronbach's Alpha values range from 0.953 to 0.961, indicating that the scale maintains a high level of reliability even if any single item is removed. Since the overall Cronbach's Alpha (from Table 1) is 0.967, none of the items appear to drastically weaken or strengthen the internal consistency of the scale. Regulatory Compliance has the lowest Alpha if deleted (0.953), meaning that removing this item would slightly decrease reliability. The other four items (Data Processing Efficiency, AI Model Accuracy, Scalability & Flexibility, and Business Value Impact) have values between 0.958 and 0.961, indicating that they all contribute similarly to the overall consistency of the scale. Since none of the items significantly reduce the reliability score when deleted, it suggests that all five items are well-constructed and contribute meaningfully to the measurement scale. The instrument appears to be highly reliable, making it suitable for research or assessment purposes without requiring item removal.

Copyright@ REST Publisher

TABLE 3. Descriptive Statistics

Descriptive Statistics												
	N	Range	Minimum	Maximum	Sum	Mean	Mean	Std. Deviation	Variance	Skewness		Kurtosis
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic
Data Processing Efficiency	30	3	2	5	109	3.63	0.2	1.098	1.206	-0.197	0.427	-1.239
AI Model Accuracy	30	2	3	5	126	4.2	0.139	0.761	0.579	-0.362	0.427	-1.141
Scalability & Flexibility	30	2	3	5	127	4.23	0.141	0.774	0.599	-0.441	0.427	-1.16
Regulatory Compliance	30	3	2	5	110	3.67	0.2	1.093	1.195	-0.289	0.427	-1.174
Business Value Impact	30	2	3	5	123	4.1	0.147	0.803	0.645	-0.188	0.427	-1.406
Valid N (listwise)	30											

Table 3 presents descriptive statistics for five measured variables: Data Processing Efficiency, AI Model Accuracy, Scalability & Flexibility, Regulatory Compliance, and Business Value Impact, based on 30 responses. The statistics provide insights into the distribution, central tendency, and variability of the data. Range & Min-Max: Most variables have a range of 2 or 3, indicating that responses vary moderately. Minimum values range from 2 to 3, while maximum values are consistently 5, suggesting that most participants rated the items positively. Mean & Standard Deviation (SD): The highest mean is for Scalability & Flexibility (4.23), suggesting it was rated most favorably. The lowest mean is for Data Processing Efficiency (3.63), indicating slightly lower ratings. Standard deviations range from 0.761 to 1.098, meaning responses have moderate variability. Skewness & Kurtosis: All skewness values are negative, showing a leftward tilt, meaning responses were generally high. Kurtosis values are negative, indicating a flatter distribution than a normal curve, suggesting less clustering around the mean.

TABLE 4. Frequencies Statistics

Statistics						
		Data Processing Efficiency	AI Model Accuracy	Scalability & Flexibility	Regulatory Compliance	Business Value Impact
N	Valid	30	30	30	30	30
	Missing	0	0	0	0	0
Median		4	4	4	4	4
Mode		4	4[a]	5	4	4[a]
Percentiles	25	3	4	4	3	3
	50	4	4	4	4	4
	75	5	5	5	5	5

The frequency statistics in Table 4 present an overview of key performance indicators across five critical dimensions: Data Processing Efficiency, AI Model Accuracy, Scalability & Flexibility, Regulatory Compliance, and Business Value Impact. The dataset consists of 30 valid observations for each metric, with no missing data. The median value for all five categories is 4, indicating that the central tendency of responses is consistent across dimensions. The mode, or most frequently occurring value, is also 4 in most cases, except for AI Model Accuracy and Business Value Impact, where there are variations (4a and 5, respectively). This suggests that while most responses cluster around 4, some factors exhibit minor variations in their dominant rating. Examining percentiles provides further insight into the distribution of responses. At the 25th percentile, Data Processing Efficiency, Regulatory Compliance, and Business Value Impact are rated 3, while AI Model Accuracy and Scalability & Flexibility score 4. The 50th percentile aligns with the median (4) for all categories. At the 75th percentile, all metrics rise to a score of 5, indicating that a significant portion of responses lean toward higher ratings.

Histogram Plot:

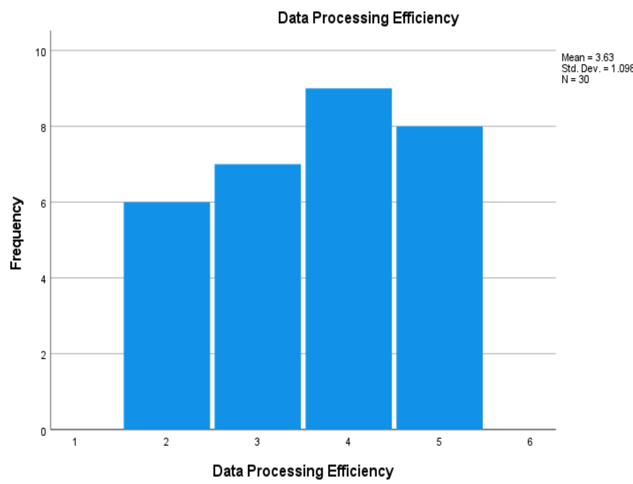


FIGURE 1. Data Processing Efficiency

Figure 1 illustrates the frequency distribution of Data Processing Efficiency ratings. The x-axis represents the rating scale, ranging from 1 to 5, while the y-axis indicates the frequency of responses for each rating. The most frequent rating is 4, with the highest number of responses, followed by 5 and 3. Rating 2 appears slightly less frequently, while rating 1 has the lowest frequency. The mean score is 3.63, which suggests a generally positive perception of data processing efficiency. The standard deviation of 1.098 indicates some variability in responses but suggests that most values are relatively close to the mean. With an overall sample size of 30, the distribution highlights that a majority of respondents rated efficiency at 4 or 5, indicating a favorable assessment. However, the presence of lower ratings suggests that there may be some inconsistencies or areas for improvement in efficiency performance.

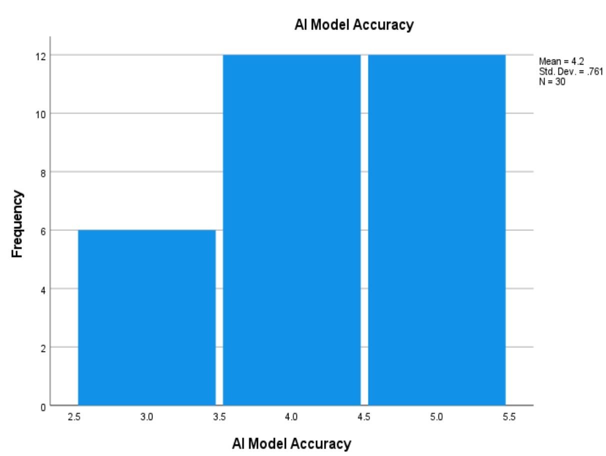


FIGURE 2. AI Model Accuracy

Figure 2 represents the frequency distribution of AI Model Accuracy ratings. The x-axis displays the rating scale, ranging from approximately 3 to 5, while the y-axis shows the number of respondents selecting each rating. The most frequent ratings are 4 and 5, each appearing 12 times, indicating a strong positive perception of AI Model Accuracy. A smaller number of responses fall at rating 3, suggesting that while some users perceive moderate accuracy, the majority view the model as highly accurate. The mean rating is 4.2, signifying a generally favorable evaluation of accuracy. The standard deviation of 0.761 indicates relatively low variability, meaning responses are fairly consistent. With a sample size of 30, the data suggests that most users are satisfied with the model's accuracy, though a minority sees room for improvement.

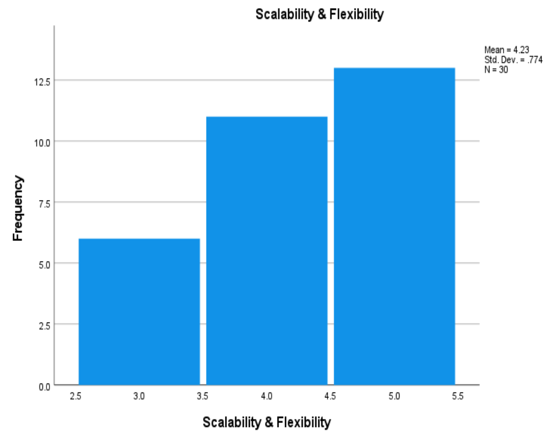


FIGURE 3. Scalability & Flexibility

Figure 3 illustrates the frequency distribution of Scalability & Flexibility ratings. The x-axis represents the rating scale, ranging from approximately 3 to 5, while the y-axis shows the number of respondents selecting each rating. The most common rating is 5, with the highest frequency, followed by 4, which also has a significant number of responses. Rating 3 appears less frequently, indicating that fewer respondents perceive scalability and flexibility as moderate. The absence of lower ratings (1 and 2) suggests that respondents generally view this aspect positively. The mean score of 4.23 reflects a strong overall assessment of scalability and flexibility, while the standard deviation of 0.774 indicates relatively low variability, meaning most responses are clustered around the mean. With a sample size of 30, the data suggests that the majority of respondents find the system's scalability and flexibility effective, with minimal concerns about limitations in this area.

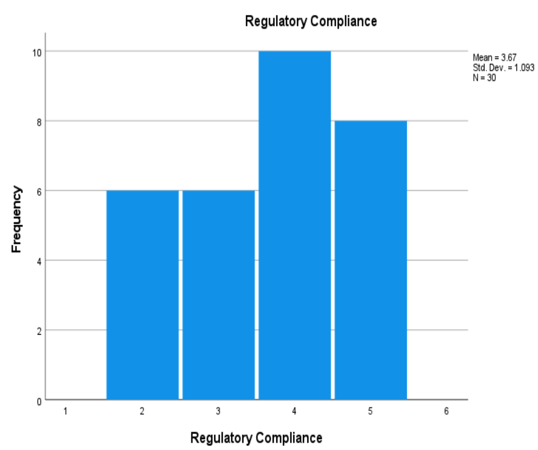


FIGURE 4. Regulatory Compliance

Figure 4 illustrates the frequency distribution of Regulatory Compliance ratings. The x-axis represents the rating scale from 1 to 5, while the y-axis shows the number of respondents for each rating. The most frequent rating is 4, with 10 responses, followed by 5, which has 8 responses. Ratings of 2 and 3 have equal frequency, with 6 responses each, while the lowest rating (1) is absent from the dataset. This distribution suggests that most respondents perceive regulatory compliance favorably, though some report lower ratings, indicating areas for improvement. The mean score of 3.67 suggests a generally positive perception of compliance, albeit with some variation. The standard deviation of

1.093 shows a moderate spread in responses, reflecting diverse opinions. With a sample size of 30, the data suggests that while compliance is generally well-regarded, some respondents see potential gaps that may need addressing.

FIGURE 5. Business Value Impact

Figure 5 presents the frequency distribution of Business Value Impact ratings. The x-axis represents the rating scale, ranging from approximately 3 to 5, while the y-axis shows the frequency of each rating. The most frequent ratings are 4 and 5, each appearing around 11 times, indicating that most respondents perceive a high business value impact. Rating 3 appears less frequently but still has a notable presence, suggesting that some respondents view the impa ct as moderate. The absence of ratings below 3 suggests an overall positive perception. With a mean score of 4.1, the data indicates a strong consensus on the positive business value impact. The standard deviation of 0.803 suggests some variation in responses, but the overall distribution remains consistent. With a total of 30 respondents, the findings highlight that most users recognize significant business value, though a minority see areas for improvement.

TABLE 5. Correlations

Correlations					
	Data Processing Efficiency	AI Model Accuracy	Scalability & Flexibility	Regulatory Compliance	Business Value Impact
Data Processing Efficiency	1	.833**	.835**	.957**	.903**
AI Model Accuracy	.833**	1	.972**	.870**	.812**
Scalability & Flexibility	.835**	.972**	1	.869**	.849**
Regulatory Compliance	.957**	.870**	.869**	1	.903**
Business Value Impact	.903**	.812**	.849**	.903**	1

Table 5 presents the correlation matrix among five key performance indicators: Data Processing Efficiency, AI Model Accuracy, Scalability & Flexibility, Regulatory Compliance, and Business Value Impact. The values indicate the strength and direction of relationships between these variables, with denoting statistically significant correlations. All correlations are positive and strong, suggesting that improvements in one area are associated with improvements in others. The highest correlation is between Data Processing Efficiency and Regulatory Compliance (0.957), indicating that efficient data processing strongly aligns with regulatory adherence. AI Model Accuracy and Scalability & Flexibility also show a very high correlation (0.972), suggesting that more a ccurate AI models tend to be highly scalable. Business Value Impact is significantly correlated with all variables, particularly with Data Processing Efficiency (0.903) and Regulatory Compliance (0.903), emphasizing their importance in driving business suc cess. These findings suggest that optimizing one factor can yield broad improvements across multiple areas.

TABLE 6. Model Summary for regression analyses

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics		Change Statistics			
					R Square Change	F Change	df1	df2	Sig. F Change	
Data Processing Efficiency	.962[a]	0.926	0.914	0.322	0.926	77.973	4	25	0	2.423
AI Model Accuracy	.976[a]	0.953	0.945	0.178	0.953	125.919	4	25	0	1.982
Scalability & Flexibility	.978[a]	0.956	0.949	0.174	0.956	136.846	4	25	0	2.116
Regulatory Compliance	.968[a]	0.937	0.927	0.295	0.937	93.234	4	25	0	2.855
Business Value Impact	.930[a]	0.866	0.844	0.317	0.866	40.309	4	25	0	1.671

Table 6 presents the model summary for regression analyses assessing the relationships between key performance indicators. The R values, which indicate the strength of relationships, are all above 0.93, suggesting a strong correlation between the independent variables and each dependent variable. R Square values range from 0.866 (Business Value Impact) to 0.956 (Scalability & Flexibility), indicating that the models explain a high percentage (86.6% to 95.6%) of the variance in these outcomes. Adjusted R Square values, which account for the number of predictors, remain strong, confirming the reliability of the models. The Standard Error of the Estimate is relatively low across models, suggesting minimal deviation from predicted values. F Change statistics are all significant ($p = 0.000$), confirming that the predictor variables significantly improve the models. Overall, the data highlights robust predictive relationships, with the highest explanatory power observed for Scalability & Flexibility and AI Model Accuracy.

TABLE 7. ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Data Processing Efficiency	32.372	4	8.093	77.973	.000[b]
	2.595	25	0.104		
	34.967	29			
AI Model Accuracy	16.006	4	4.001	125.919	.000[b]
	0.794	25	0.032		
	16.8	29			
Scalability & Flexibility	16.608	4	4.152	136.846	.000[b]
	0.759	25	0.03		
	17.367	29			
Regulatory Compliance	32.489	4	8.122	93.234	.000[b]
	2.178	25	0.087		
	34.667	29			
Business Value Impact	16.19	4	4.047	40.309	.000[b]
	2.51	25	0.1		
	18.7	29			

Table 7 presents the ANOVA (Analysis of Variance) results, which assess the statistical significance of the regression models for the five key performance indicators. For all models, the F-values are high (ranging from 40.309 for Business Value Impact to 136.846 for Scalability & Flexibility), and the p-values (Sig.) are all .000, indicating that the predictor variables significantly contribute to explaining the variance in each dependent variable. The Sum of Squares values show how much variation is explained by the model versus the residual error. The Mean Square values for the regression components are notably larger than the residuals, confirming strong model performance. The lowest error variance is observed in AI Model Accuracy (Mean Square Error = 0.032), indicating high precision in predictions.

TABLE 8. Communalities

Communalities		
	Initial	Extraction
Data Processing Efficiency	1.000	.908
AI Model Accuracy	1.000	.890
Scalability & Flexibility	1.000	.905
Regulatory Compliance	1.000	.937
Business Value Impact	1.000	.883

Table 8 presents the communalities for the five key performance indicators, showing how much variance in each variable is explained by the extracted components. The Initial column lists all values as 1.000, indicating that before extraction, each variable contains 100% of its variance. The Extraction column represents the proportion of variance retained after factor analysis. Regulatory Compliance has the highest extracted communality (0.937), suggesting that most of its variance is well explained by the underlying factors. Data Processing Efficiency (0.908) and Scalability & Flexibility (0.905) also have high communalities, indicating strong representation. AI Model Accuracy (0.890) and Business Value Impact (0.883) have slightly lower values but remain well-explained. Since all extraction values are above 0.88, the factors effectively capture the variance in each variable, meaning these indicators contribute significantly to the underlying structure of the data. This confirms the robustness of the factor analysis model.

TABLE 9. Total Variance Explained

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		Extraction Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.522	90.45	90.45	4.522	90.45	90.45
2	0.296	5.92	96.37			
3	0.119	2.382	98.752			
4	0.039	0.788	99.541			
5	0.023	0.459	100			
Extraction Method: Principal Component Analysis						

Table 9 presents the total variance explained by the principal component analysis (PCA), showing how much of the data's variance is captured by each component. The first component has an eigenvalue of 4.522 and explains 90.45% of the total variance, indicating that a single factor accounts for most of the variation across the five performance indicators. The remaining components contribute minimal variance, with the second component explaining 5.92%, and the third, fourth, and fifth components contributing less than 3% combined. Since only the first component has a significant eigenvalue (>1), it is the only retained factor in the analysis. The high cumulative variance (90.45%) suggests that a single principal component effectively summarizes the dataset, meaning the five indicators are highly correlated and can be grouped into a single underlying factor. This confirms that a simplified model can represent the relationships among these variables with minimal information loss.

4. CONCLUSION

Artificial intelligence has emerged as the fourth industrial revolution's main force, significantly transforming industries through automation, predictive analytics, and improvements when making decisions. The development of artificial intelligence from its inception in the 1960s to its current state demonstrates a continuous trajectory of innovation that is shaping businesses, healthcare, finance, and manufacturing. Combining Big data analytics, cloud computing, IoT, and artificial intelligence have all combined to create an ecosystem where intelligent technologies continuously optimize processes, boost productivity, and cut expenses. Predictive analytics is one of AI's primary contributions, where Large datasets are analyzed by machine learning algorithms to find trends and detect hazards, and improve strategic decision-making. For example, in financial services, AI-powered fraud detection systems are reducing risks, while in healthcare, predictive models support early diagnosis and personalized treatments. Similarly, in industrial automation, AI-powered robotics are improving productivity by reducing downtime and improving supply chain management. Despite these advances, challenges remain in AI implementation, particularly around data security, privacy, and ethical implications. The growing reliance on AI-powered decision-making necessitates robust governance structures to ensure transparency, accountability, and compliance with regulatory standards. In addition, workforce adaptation to AI-powered technologies requires significant investment in reskilling and development to bridge the skills gap in industries that rely heavily on automation. In the future, artificial intelligence is expected to grow even more because to new developments including explainable AI, federated learning, and AI-driven cybersecurity draw traction. These advances will enable more transparent, secure, and collaborative AI applications, addressing concerns about data privacy and bias in machine learning models. As industries continue to embrace AI innovations, interdisciplinary research and collaboration among academics, industry leaders, and policymakers are essential to shaping ethical and sustainable AI ecosystems. By comprehensively analyzing the historical evolution, current landscape, and future trajectory of AI, for academics, professionals, and lawmakers, this study offers insightful information. The results highlight both the necessity for responsible and strategic AI as well as the revolutionary potential of AI integration to maximize its benefits across various industry sectors.

REFERENCES

1. Lu, Yang. "Artificial intelligence: a survey on evolution, models, applications and future trends." Journal of Management Analytics 6, no. 1 (2019): 1-29.
2. Islam, Siful. "Future Trends in SQL Databases and Big Data Analytics: Impact of Machine Learning and Artificial Intelligence." Available at SSRN 5064781 (2024).
3. Kibria, Mirza Golam, Kien Nguyen, Gabriel Porto Villardi, Ou Zhao, Kentaro Ishizu, and Fumihide Kojima. "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks." IEEE access 6 (2018): 32328-32338.
4. Thayyib, P. V., Rajesh Mamilla, Mohsin Khan, Humaira Fatima, Mohd Asim, Imran Anwar, M. K. Shamsudheen, and Mohd Asif Khan. "State-of-the-art of artificial intelligence and big data analytics reviews in five different domains: A bibliometric summary." Sustainability 15, no. 5 (2023): 4026.
5. Wong, Zoie SY, Jiaqi Zhou, and Qingpeng Zhang. "Artificial intelligence for infectious disease big data analytics." Infection, disease & health 24, no. 1 (2019): 44-48.
6. Luan, Hui, Peter Geczy, Hollis Lai, Janice Gobert, Stephen JH Yang, Hiroaki Ogata, Jacky Baltes, Rodrigo Guerra, Ping Li, and Chin-Chung Tsai. "Challenges and future directions of big data and artificial intelligence in education." Frontiers in psychology 11 (2020): 580820.
7. Jagatheesaperumal, S. K. "The Duo of Arti cial Intelligence and Big Data for Industry 4.0: Review of Applications, Techniques, Challenges, and Future Research Directions." (2021).
8. Raschka, Sebastian, Joshua Patterson, and Corey Nolet. "Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence." Information 11, no. 4 (2020): 193.
9. Eboigbe, Emmanuel Osamuyimen, Oluwatoyin Ajoke Farayola, Funmilola Olatundun Olatoye, Obiageli Chinwe Nnabugwu, and Chibuike Daraojimba. "Business intelligence transformation through AI and data analytics." Engineering Science & Technology Journal 4, no. 5 (2023): 285-307.
10. Radanliev, Petar, David De Roure, Kevin Page, Jason RC Nurse, Rafael Mantilla Montalvo, Omar Santos, La'Treall Maddox, and Pete Burnap. "Cyber risk at the edge: current and future trends on cyber risk analytics and artificial intelligence in the industrial internet of things and industry 4.0 supply chains." Cybersecurity 3 (2020): 1-21.
11. Ruiz-Real, José Luis, Juan Uribe-Toril, José Antonio Torres, and Jaime De Pablo. "Artificial intelligence in business and economics research: Trends and future." Journal of Business Economics and Management 22, no. 1 (2021): 98-117.
12. Naeem, Muhammad, Tauseef Jamal, Jorge Diaz-Martinez, Shariq Aziz Butt, Nicolo Montesano, Muhammad Imran Tariq, Emiro De-la-Hoz-Franco, and Ethel De-La-Hoz-Valdiris. "Trends and future perspective challenges in big data." In Advances in intelligent data analysis and applications: Proceeding of the sixth euro-China conference on intelligent data analysis and applications, 15–18 October 2019, Arad, Romania, pp. 309-325. Springer Singapore, 2022.
13. Hanna, Matthew G., Liron Pantanowitz, Rajesh Dash, James H. Harrison, Mustafa Deebajah, Joshua Pantanowitz, and Hooman H. Rashidi. "Future of Artificial Intelligence (AI)-Machine Learning (ML) Trends in Pathology and Medicine." Modern Pathology (2025): 100705.
14. Chen, Pu, Linna Wu, and Lei Wang. "AI fairness in data management and analytics: A review on challenges, methodologies and applications." Applied sciences 13, no. 18 (2023): 10258.
15. Panda, Khirod, and Shobhit Agrawal. "Predictive analytics: an overview of evolving trends and methodologies." The Journal of Scientific and Engineering Research. 8 (2024): 175-180.
16. Mehta, Nishita, Anil Pandit, and Sharvari Shukla. "Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study." Journal of biomedical informatics 100 (2019): 103311.
17. Baarimah, Abdullah O., Mahmood A. Bazel, Wesam Salah Alaloul, Motasem YD Alazaiza, Tharaa M. Al-Zghoul, Basheer Almuhaya, Arsalaan Khan, and Ahmed W. Mushtaha. "Artificial intelligence in wastewater treatment: Research trends and future perspectives through bibliometric analysis." Case Studies in Chemical and Environmental Engineering (2024): 100926.

18. Ucar, Aysegul, Mehmet Karakose, and Necim Kırımça. "Artificial intelligence for predictive maintenance applications: key components, trustworthiness, and future trends." Applied Sciences 14, no. 2 (2024): 898.
19. Kulkov, Ignat. "The role of artificial intelligence in business transformation: A case of pharmaceutical companies." Technology in Society 66 (2021): 101629.
20. Weaver, Bruce, and Hillary Maxwell. "Exploratory factor analysis and reliability analysis with missing data: A simple method for SPSS users." The Quantitative Methods for Psychology 10, no. 2 (2014): 143-152.
21. Dudley, William N., Jose G. Benuzillo, and Mineh S. Carrico. "SPSS and SAS programming for the testing of mediation models." Nursing research 53, no. 1 (2004): 59-62.
22. Shek, Daniel TL, and Cecilia MS Ma. "Application of SPSS linear mixed methods to adolescent development research: basic concepts and steps." International Journal on Disability and Human Development 13, no. 2 (2014): 169-182.
23. Maric, Marija, Else de Haan, Sanne M. Hogendoorn, Lidewij H. Wolters, and Hilde M. Huizenga. "Evaluating statistical and clinical significance of intervention effects in single-case experimental designs: An SPSS method to analyze univariate data." Behavior Therapy 46, no. 2 (2015): 230-241.
24. Chinnasamy, Sathiyaraj, M. Ramachandran, Prabakaran Nanjundan, and Vidhya Prasanth. "Two Phase Retrial Queue Evaluation using DEMATEL (Decision Making Trial and Evaluation Laboratory)." *Journal of Communication Engineering & Systems* 12, no. 3 (2022): 6-19p.
25. Thoemmes, Felix. "Propensity score matching in SPSS." arXiv preprint arXiv:1201.6385 (2012).
26. Yang, Hongwei. "The case for being automatic: introducing the automatic linear modeling (LINEAR) procedure in SPSS statistics." Multiple Linear Regression Viewpoints 39, no. 2 (2013): 27-37.
27. Liu, R. X., J. Kuang, Qiong Gong, and X. L. Hou. "Principal component regression analysis with SPSS." Computer methods and programs in biomedicine 71, no. 2 (2003): 141-147.
28. Kent, Peter, Rikke K. Jensen, and Alice Kongsted. "A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep Cluster analysis, Latent Gold and SNOB." BMC medical research methodology 14 (2014): 1-14.
29. Von Hippel, Paul T. "Biases in SPSS 12.0 missing value analysis." The American Statistician 58, no. 2 (2004): 160-164.
30. Giguère, Gyslain. "Collecting and analyzing data in multidimensional scaling experiments: A guide for psychologists using SPSS." Tutorials in Quantitative Methods for Psychology 2, no. 1 (2006): 27-38.
31. Pfister, Roland, Katharina Schwarz, Robyn Carson, and Markus Jancyzk. "Easy methods for extracting individual regression slopes: Comparing SPSS, R, and Excel." Tutorials in Quantitative Methods for Psychology 9, no. 2 (2013): 72-78.
32. Liang, Guangping, Wenliang Fu, and Kaifa Wang. "Analysis of t-test misuses and SPSS operations in medical research papers." Burns & trauma 7 (2019).