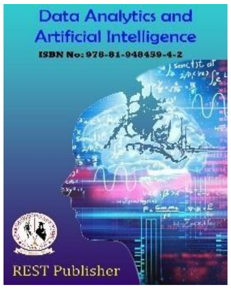
Data Analytics and Artificial Intelligence

Vol: 5(1), 2025

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/

DOI: https://doi.org/10.46632/daai/5/1/1



AI-Driven Automation in Power Utility Management: A Scalable Approach Using SPSS

Dhivya Karunya S

S.E.A College of Engineering & Technology, Bangalore, Karnataka, India.

**Corresponding Author Email: dhivyaskarunya@gmail.com*

Abstract: *Making solution decisions in the power industry is a complex and multifaceted task that requires careful management of scalability, service reliability, and security vulnerabilities. Traditional manual decision-making approaches are inefficient, time-consuming, and prone to errors, thereby increasing costs and security risks. To address these challenges, this research paper introduces an automated framework based on machine learning that enhances decision analysis for electrical applications. Through rigorous testing with real operational datasets, the framework demonstrates its effectiveness in improving decision-making processes, improving security, and reducing operational overhead. Research Significance: This research is significant as it addresses critical inefficiencies in power utility decision-making by leveraging machine learning automation. The proposed framework improves scalability, mitigates security risks, and reduces human intervention, ultimately leading to cost-effective and efficient power management. This research offers factual understanding of the US OP AI-driven solutions in infrastructure management, providing a replicable model for other industries facing similar challenges. The results demonstrate the potential for widespread adoption in improving operational resilience and decision-making accuracy. SPSS Statistics is a powerful software tool used to analyze data in many fields such as social sciences, healthcare, marketing, and education. It provides a comprehensive set 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, helping to provide data-driven insights. Input Parameters taken as industry, Automation Tool Used, Threat Type Addressed, Incident Response Framework, Remediation Strategy, Organization Size. Evaluation Parameters taken as Efficiency of Automation, Reduction in Response Time, Accuracy of Threat Detection, Compliance Adherence, Cost-effectiveness. The reliability statistics show a 0.533 Cronbach's Alpha and 0.535 based on standardized items, with a total of 5 items analyzed.*

Keywords: *SPSS Statistics, Machine Learning, Power Utilities, Decision Automation, Security Vulnerabilities.*

1. INTRODUCTION

Making solution decisions in the power industry is a complex task due to multiple factors, the need to balance scalability with There are several vulnerabilities that must be handled, and service reliability is one of the biggest obstacles. These choices are now made by hand, which takes time, raises security concerns, and raises vulnerability management expenses. In order to enhance decision analysis for power applications, this research presents an automated system based on machine learning. Two actual operational datasets are used to thoroughly test the framework after it is integrated within a power application. The outcomes demonstrate the solution's strong performance. [2] De carbonization, a process that reduces carbon dioxide (CO2) emissions, plays a key role in environmental management and climate change mitigation. Conventional methods Data collection and analysis in contaminated soil remediation (CSR), wastewater treatment (WWT), and solid waste management (SWM) are frequently done by hand, which can lead to inaccuracies and inefficiencies. This study examines how decision support systems (DSS) and data automation can transform existing procedures by offering precise, real-time insights that enhance operational effectiveness and decision-making. demonstrate the solution's reliable performance efficiency. [3] These businesses find it difficult to implement best practices because of a lack of understanding mitigate attack risks, have no way to monitor their infrastructure for unusual activity, and lack the ability to detect active threats or determine the most effective mitigation strategies. As a result, small businesses often rely on external security outsourcing to protect their software, filling their knowledge and skill gaps while managing costs. [4] As a result, Institutions of higher learning frequently lack the know-how required to improve their cybersecurity posture. This report identifies vulnerabilities that lack appropriate repair solutions by doing a comprehensive vulnerability assessment of 272 institutions. In order to solve this, we create automated reporting systems that produce succinct, useful reports for effective vulnerability management, as well as reproduce and fix a few chosen vulnerabilities in a virtual environment. 27.80% of the vulnerabilities identified in the evaluated institutions are successfully addressed by our improved reports. [5] As software development and computing continue to evolve, the need to automate tasks that are traditionally performed manually has increased significantly. Ensuring continuous operation with minimal downtime underscores the importance of automatically detecting and fixing runtime anomalies. Recognized for its scalability, Ansible is a reliable option for safely managing complicated systems because of its high-level abstraction and modularity. But the difficulty is in building an adaptive Ansible-based auto-fixing solution that requires a substantial dataset to fine-tune large language models (LLMs). [6] The fundamentals of hole cleaning in drilling are widely known with established methods for accurate diagnosis during operations and effective mitigation or remediation strategies. Although significant technological advances have improved the detection of hole cleaning failures through a variety of approaches, the most reliable identification still relies on human expertise to synthesize individual results. This research paper explores how a complete approach to hole cleaning analysis can be automated using an innovative digital solution. [7] Static analysis uses automated tools to analyze compiled binaries or source code without actually running the program. It assesses the code structure, finds any errors, and finds common vulnerabilities including injection attacks, buffer overflows, and improper data manipulation. By applying methods like pattern matching, Static analysis, data flow analysis, and control flow analysis tools effectively pinpoint security risks. These tools improve scalability in vulnerability detection by enabling analysis of large code bases with high performance. [8] Effective administration and oversight of contemporary and next-generation networks, in line with growing traffic demands and evolving applications, are crucial for maintaining high service reliability. Mitigating network service degradations quickly is essential to ensure excellent service quality. Automation is essential to accomplishing this, particularly at the network edge, where diversity and scalability are critical factors. This entails quickly identifying, locating, and, if feasible, fixing performance problems and malfunctions that affect services. The most difficult part, though, is figuring out which events to look for, how to link them to the location of the issue, and what kind of mitigation is best measures. [9] The automatic localization of mitigation information for security vulnerabilities is the subject of this study for the first time. In order to find mitigating details, we offer three techniques for pulling pertinent text from reference pages. The first method extracts pertinent paragraphs using a keyword-based system and presumes prior knowledge of a predetermined set of reference sources. The second method is designed to specifically target sections related to mitigation strategies. [10] This provides empirical evidence for possible correlations between CVSS metrics, which could improve the automation of security vulnerability assessments. It suggests that the assessment approach should exploit these relationships and predict the entire vector string as a unified entity, rather than combining predictions for each metric independently. [11] Using AI models for automated incident detection and routing, operators continuously collect metrics and logs. These models identify incidents in real-time, use predictive systems to alert operators, and route them to the appropriate teams. These models detect and effectively notify incidents, while operators traditionally rely on fault resolution to determine root causes. [12] The integration Combining cloud computing with machine learning (ML) presents a substantial opportunity to improve cloud security and compliance. This study investigates how machine learning (ML) may enhance cloud compliance through automated incident response, adaptive security measures, and proactive threat detection. ML-based strategies are particularly valuable for processing vast datasets, detecting anomalies, and mitigating risks in ever-evolving cloud environments. [13] Inspired by natural processes, autonomous Computing is a computational model that manages itself. This paper proposes an integrated architecture based on autonomous computing to automatically detect, raise an alarm, assess, classify, rank, mitigate, and manage software vulnerabilities. By means of self-optimization, self-prevention, self-healing, and self-configuration, the framework automatically performs repair activities for future security vulnerabilities using an inference engine and knowledge base. By providing an intelligent, cross-domain, and integrated approach, this framework improves security and provides a self-managed environment that benefits both industry and society. [14] Once a threshold breach is detected, automated warnings and responses can be sent off within a minute. Enhanced spatial and temporal resolution further improves solution design optimization and ensures the effectiveness of the mitigation system. While regulatory bodies have recognized continuous monitoring as an effective solution to address spatial and temporal dynamics, cost constraints and instrument limitations previously made it impractical. However, recent advances in Rapid, continuous evaluation and response from multiple locations with a single tool is now possible because to automation and multiplexing. [15] Specifically, we leverage the OASIS Collaborative Automated Actions (CACAO) standard to build highly automated workflows that support cybersecurity operations within Advanced Metering Infrastructure (AMI). The proposed approach is validated using the AMI testbed, where large-scale cyberattacks are simulated, and playbooks are used to enable rapid threat containment and mitigation. This ensures business continuity for smart meter data exchange services, while also ensuring compliance with incident reporting requirements. [16] Automated tools are used in vulnerability scanning to find known security flaws in the dependencies, infrastructure, or code base of an application. These tools analyze software setups and components in a methodical manner and identify potential vulnerabilities that could be exploited. By quickly assessing security risks, vulnerability scanning helps prioritize patches and mitigation efforts. The technique of manually or automatically reviewing Code review is the process of examining source code to identify security vulnerabilities at the code level. Typical issues include cross-site scripting (XSS), SQL injection, and inadequate authentication techniques are found using standard analysis tools. This method promotes secure development techniques and aids in the early detection of coding errors. [17] Because of the ever-evolving threat landscape, modern software development must incorporate shift-left security measures and automated vulnerability detection in container images. As soon as the integration with container environments such as Docker and Kubernetes began, our method places a high priority on security. This study analyzes the main security issues with containerization, such as supply chain risks, runtime threats, image vulnerabilities, and misconfigurations, while focusing on regulatory compliance. Additionally, it looks at how vulnerability databases, policy enforcement tools, and static and dynamic analysis are used to include automated vulnerability detection into pipelines for continuous integration and deployment (CI/CD).

2. MATERIALS & METHODS

Input Parameters:

Industry: The sector in which the organization operates, such as cybersecurity, healthcare, finance, manufacturing, energy, or retail.

Automation tool used: The specific software or platform used to automate remediation and mitigation processes, including tools such as Splunk, IBM Resilient, Palo Alto XSOAR, Service Now, Ansible, or Microsoft Sentinel.

Threat type: The type of cybersecurity threat, such as phishing, ransomware, DDoS, insider threat, data breach, or malware.

Incident response framework: A structured approach or standard used to manage and mitigate security incidents, for example, NIST, ISO 27001, Mitre ADTR ADTR compliance, SOC 2, or CIS controls.

Remediation Strategy: The method or practice implemented to mitigate and resolve security incidents, including endpoint detection and response (EDR), network segmentation, patch management, identity and access management (IAM), automated threat hunting, or zero trust.

Organization Size: The size of the organization, categorized as small business, medium, enterprise, government, non-profit, or startup based on structure and employee count.

Evaluation Parameters:

Automation Effectiveness: A measure of how effectively the automation tool streamlines the remediation and mitigation processes, rated from 1 (worst) to 5 (best).

Response Time Reduction: The extent to which the automation tool reduces the time required to detect, respond to, and mitigate security incidents, rated from 1 (no improvement) to 5 (significant improvement).

Threat Detection Accuracy: The reliability of the automation tool in identifying real threats while minimizing false positives and false negatives, rated on a scale of 1 (low) to 5 (high).

Compliance Adherence: The extent to which the automation tool helps the organization meet regulatory and security compliance requirements, evaluated from 1 (non-compliant) to 5 (completely compliant).

Cost-Effectiveness: The balance between the financial investment in automation tools and the benefits gained in terms of improved security and operational efficiency, rated from 1 (not cost-effective) to 5 (very cost-effective), on a scale of 1.

SPSS Method: For researchers performing reliability analysis or exploratory factor analysis (EFA), missing data is a frequent problem. The process of SPSS Factor offers list wise imputation, pairwise imputation, and mean imputation as options for handling missing data, although these methods have well-documented limitations. Graham (2009) suggests that the most useful A in this instance, the analysis's input is a matrix of Expectation Maximization (EM) covariance or correlations. Covariance matrices, correlation matrices, EM means, and standard deviations can be produced by SPSS users using the Missing Values Analysis (MVA) module. However, since MVA lacks a /MATRIX subcommand, it is not possible to directly output EM correlations into a matrix dataset suitable for use in In this instance, the analysis's input is a matrix of Expectation Maximization (EM) covariance or correlations. Covariance matrices, correlation matrices, EM means, and standard deviations covariance. [19] Mediation modeling illustrates how one variable mediates the relationship between an intervention and its outcome and helps to understand relationships between three or more variables. Introduced in 1990, the Sobel test offers a statistical method for evaluating a mediator's influence on an intervention or result. The Sobel test can be computed using standalone and web-based technologies, and nursing researchers can now more easily access mediation modeling thanks to SPSS and SAS software that automates the required regression analyses and computations. [20] The purpose of the current study is to demonstrate the utility of Individual Growth Curve (IGC) modeling using SPSS. This study outlines methods for creating, evaluating, and contrasting models. We investigate patterns of change when Students in high school take part in a constructive youth development program by examining a five-year longitudinal dataset. [21] Single-case experimental designs are valuable tools in clinical research for assessing individual client progress. However, their widespread acceptance may be limited by methodological challenges, including difficulties in using existing statistical techniques. This article presents a data analysis method for examining single-case (single-symptom) data using the widely available SPSS software. [22] A comprehensive implementation in SPSS has been lacking, and this paper aims to address that gap. This addition has proven valuable, as many applied researchers in psychology, education, and other social sciences use SPSS primarily for data analysis and may not be familiar with R. The program "PS matching" is designed as a custom dialog for SPSS, compatible with versions 18, 19, and 20. While it generates SPSS syntax, it provides a familiar point-and-click interface that users can modify if necessary. The program "PS matching" implements all analyses in R through the SPSS R-Plugin, using newly created R code by the author with existing R packages from other researchers. [23] Before performing linear modeling, data must be cleaned and prepared, which includes: (1) replacing missing values, (2) converting date, month, and hour data to a time format, (3) specifying categorical predictors, and (4) identifying and addressing outliers. This study evaluates its effectiveness in identifying and managing outliers. Within the LINEAR procedure, continuous predictor values that exceed a cutoff of three standard deviations from the mean are classified as outliers. [24] In order to guarantee that each component is uncorrelated with the others, principal component analysis (PCA) divides highly correlated independent variables into primary components. A new set of correlated principal components is created from the set of correlated variables as a result of this transformation. These elements are then used to create regression equations, and the best equation is chosen by taking into account the lowest standard error of estimation and the largest corrected $R2$. Finally, the optimal equation is transformed back into a general linear regression model. This paper demonstrates how SPSS 10.0 can be used to solve multi collinearity problems through principal component regression. [25] One limitation of the Two Step clustering method is that it is not specifically designed for analyzing ordinal data. Although dummy variables can be used to add ordinal data, this approach disproportionately affects the distance measure, leading to unpredictable subgroup modeling results. As part of the IBM SPSS Basic package, Two Step is available on Linux, Apple Mac, and IBM PC systems, with continuous, fee-based support offered. [26] The first two methods, list wise imputation and pairwise imputation - also known as complete-case and available-case analysis - are not unique advantages of the Missing Values Analysis (MVA) module, as they are already available in SPSS Basic software. In fact, SPSS Basic provides a much more comprehensive implementation, providing list wise and pairwise estimation for a variety of models, including regression and factor analysis. [27] In this analysis, SPSS ALSCAL creates a fixed Stimulus space with participant space representing the different priorities that each person assigns on dimensions within the common stimulus space. It also provides fitted models for each participant’s data. This approach uses multiple matrices of matrix-conditioned or unconditional data. In weighted multidimensional scaling (WMDS), the individual distances between any two participants do not need to be connected by a linear or monotonic function. [28] The various methods for extracting regression slopes can be summarized as follows. In SPSS, this process requires two distinct steps: conducting separate regression analyses and then converting the results into a new dataset using the Output Management System (OMS). In R, a similar approach is used, rotating the dataset while saving the regression output at every iteration. Regression slopes and intercepts for basic linear regressions are readily available in Excel or Calc thanks to built-in tools. [29] In SPSS, choose "Analysis ➔ "General Linear Model ➔ Univariate" to use the general linear model to perform a factorial ANOVA. The results indicate that both factors are effective because there is a statistically significant interaction between treatment and illness type (p < 0.001) together to influence the dependent variable (pain score). Consequently, each factor requires a basic main effects test. Given that both factors involve independent samples, the models appropriate for an independent sample one-way ANOVA can be filtered out using SPSS's "Section File" feature located under the "Data" menu.

3. RESULT AND DISCUSSION

TABLE 1. Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.533	.535	5

Table 1 presents the reliability statistics for the given dataset, including Cronbach’s alpha, the Cronbach's alpha according to the rated things, and the number of items analyzed. Cronbach’s alpha is a measure of internal consistency, indicating how well the items on a scale correlate with each other. In this case, Cronbach’s alpha is 0.533, indicating poor reliability. Similarly, the standardized Cronbach’s alpha is 0.535, indicating only minimal improvement when the items are rated. The number of items included in the analysis is five (N = 5). A Cronbach’s alpha value above 0.7 is generally considered acceptable, values above 0.8 indicate good reliability, and those above 0.9 are considered excellent. However, a value of 0.533 falls below the acceptable range, indicating that the scale may not be reliably measuring the intended construct. This low reliability suggests that the items included in the scale may not be highly correlated, meaning they do not consistently measure the same underlying concept. To improve reliability, researchers may need to examine individual items through item-total statistics, remove weak items, revise ambiguous questions, or consider adding more items that align with the construct. Conducting an exploratory factor analysis (EFA) may also help identify underlying dimensions affecting reliability.

TABLE 2. Reliability Statistic individual

	Cronbach's Alpha if Item Deleted
Efficiency of Automation	.629
Reduction in Response Time	.392
Accuracy of Threat Detection	.446
Compliance Adherence	.217
Cost-effectiveness	.584

The Cronbach’s Alpha if Item Deleted values indicate how the overall reliability of the scale would change if each individual item were removed. The current Cronbach’s Alpha is 0.533, which is considered poor reliability. By examining these values, we can identify problematic items that may be lowering the overall internal consistency. If the Compliance Adherence item were removed, the Cronbach’s Alpha would drop to 0.217, suggesting that this item contributes positively to reliability. However, removing the Efficiency of Automation would increase reliability to 0.629, indicating that this item may not align well with the other items. Similarly, eliminating Cost-effectiveness would raise Alpha to 0.584, suggesting it has a weaker correlation with the scale. To improve reliability, researchers may consider removing or revising weak items, such as Efficiency of Automation and Cost-effectiveness, while refining the scale to ensure a more consistent measurement of the construct.

TABLE 3. Descriptive Statistics

	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance	Skewness	Kurtosis
efficiency of Automation	20	4	1	5	65	3.25	1.409554	1.987	0.008	-1.336
Reduction in Response Time	20	4	1	5	49	2.45	1.316894	1.734	0.581	-0.578
Accuracy of Threat Detection	20	4	1	5	59	2.95	1.503505	2.261	0.197	-1.381
Compliance Adherence	20	4	1	5	62	3.1	1.552587	2.411	0.004	-1.511
Cost-effectiveness	20	4	1	5	61	3.05	1.605091	2.576	0.08	-1.72

Table 3 presents the descriptive statistics for five measured variables: Efficiency of Automation, Reduction in Response Time, Accuracy of Threat Detection, Compliance Adherence, and Cost-effectiveness. Each variable was measured on a five-point scale (1–5) with a sample size of 20 participants. The mean scores suggest that Efficiency of Automation (3.25), Compliance Adherence (3.1), and Cost-effectiveness (3.05) received relatively higher ratings, while Reduction in Response Time (2.45) had the lowest mean, indicating lower perceived effectiveness in this area. The standard deviations range from 1.31 to 1.60, suggesting moderate variability in responses. The variance values further confirm this spread. Skewness values range between 0.004 and 0.581, indicating a relatively symmetrical distribution of responses. However, kurtosis values are negative, showing that the distributions are slightly platykurtic (flatter than a normal distribution).

TABLE 4. Frequencies Statistics

Statistics						
		Efficiency of Automation	Reduction in Response Time	Accuracy of Threat Detection	Compliance Adherence	Cost-effectiveness
N	Valid	20	20	20	20	20
	Missing	0	0	0	0	0
Median		3	2	3	3	2.5
Mode		5	1	2.000[a]	5	2.000[a]
Percentiles	25	2	1	2	2	2
	50	3	2	3	3	2.5
	75	5	3	4.75	5	5

Table 4 presents the frequency statistics for the five measured variables, including median, mode, and percentiles. The sample size (N = 20) is consistent across all variables, with no missing values. The median values suggest that most responses are centered around moderate ratings, with Efficiency of Automation (3), Accuracy of Threat Detection (3), and Compliance Adherence (3) scoring higher than Reduction in Response Time (2) and Cost-effectiveness (2.5). The mode values, which indicate the most frequently chosen response, show that Efficiency of Automation and Compliance Adherence had the highest mode (5), while Reduction in Response Time and Cost-effectiveness had the lowest mode (2). The percentile distribution shows that 25% of participants rated each variable at 2 or lower, while 75% rated Efficiency of Automation and Compliance Adherence at 5, indicating some variability in responses. These results suggest that some aspects, like Reduction in Response Time and Cost-effectiveness, were rated lower, signaling potential areas for improvement.

Histogram Plot

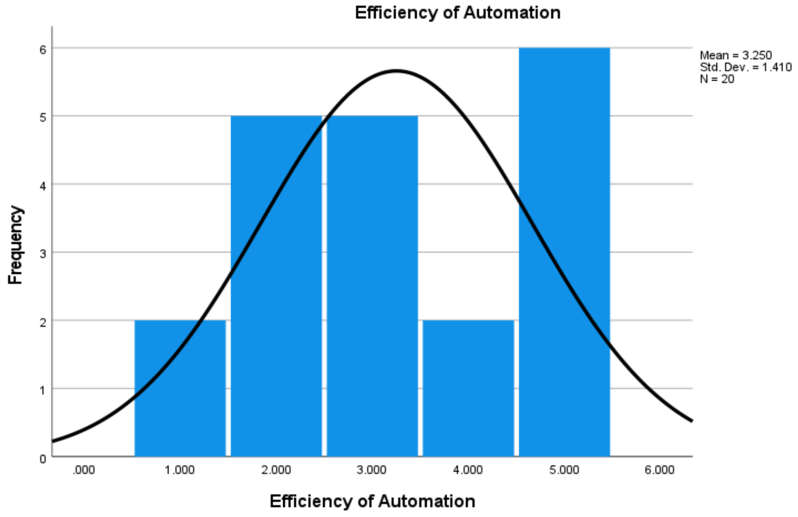


FIGURE 1. Efficiency of Automation

Figure 1 presents a histogram illustrating the frequency distribution of Efficiency of Automation scores. The mean value is 3.25, with a standard deviation of 1.41 based on N = 20 responses. The histogram shows a moderately symmetrical distribution, with scores spread across the range from 1 to 5. The black curve represents the normal distribution fit. The most frequent responses are 3 and 5, suggesting neutral to high ratings. However, the spread of responses indicates some variability. This suggests that while automation is seen as efficient, opinions vary among participants.

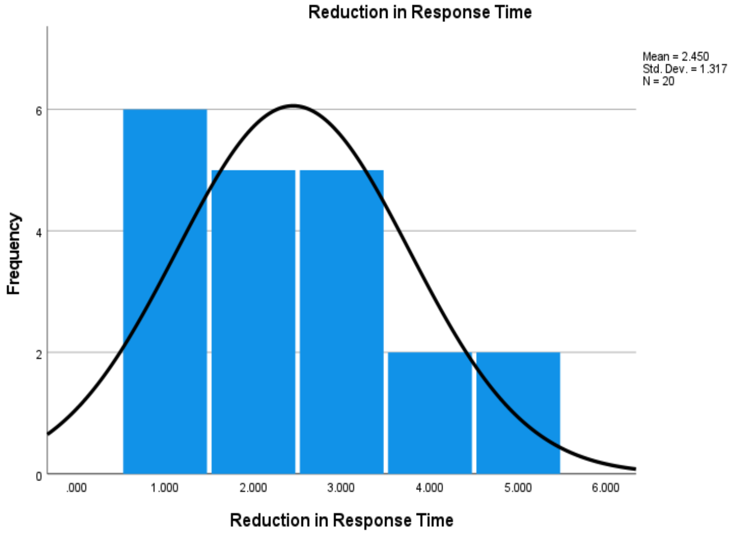


FIGURE 2. Reduction in Response Time

Figure 2 displays a histogram showing the frequency distribution of Reduction in Response Time scores. The mean value is 2.45, with a standard deviation of 1.317, based on N = 20 responses. The distribution is left-skewed, indicating that many participants rated response time reduction lower on the scale. The most frequent responses are 1 and 2, suggesting that many respondents perceived automation as less effective in reducing response time. While some ratings are higher, the overall trend indicates concerns about efficiency in this area, requiring potential improvements in automation performance.

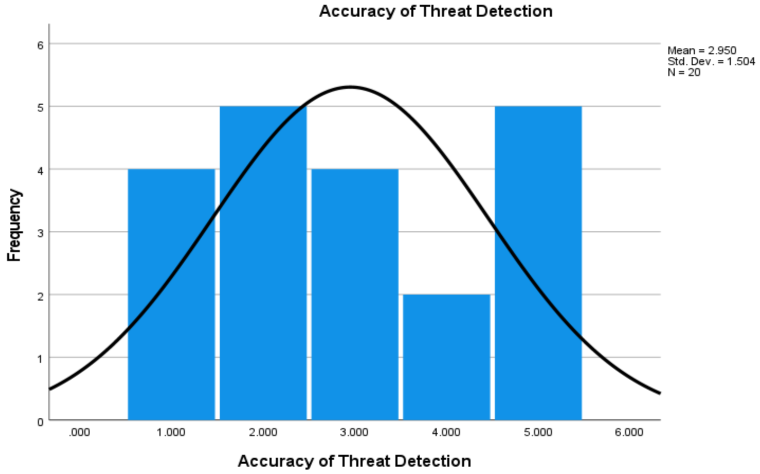


FIGURE 3. Accuracy of Threat Detection

Figure 3 presents a histogram depicting the frequency distribution of Accuracy of Threat Detection scores. The mean value is 2.95, with a standard deviation of 1.504, based on N = 20 responses. The distribution appears moderately symmetrical, indicating a balanced spread of ratings. The most frequent ratings are 2 and 5, suggesting a mixed perception of accuracy. While some respondents rated it highly, others expressed concerns about its reliability. The variability in responses indicates that improvements may be needed to enhance the consistency and effectiveness of threat detection accuracy.

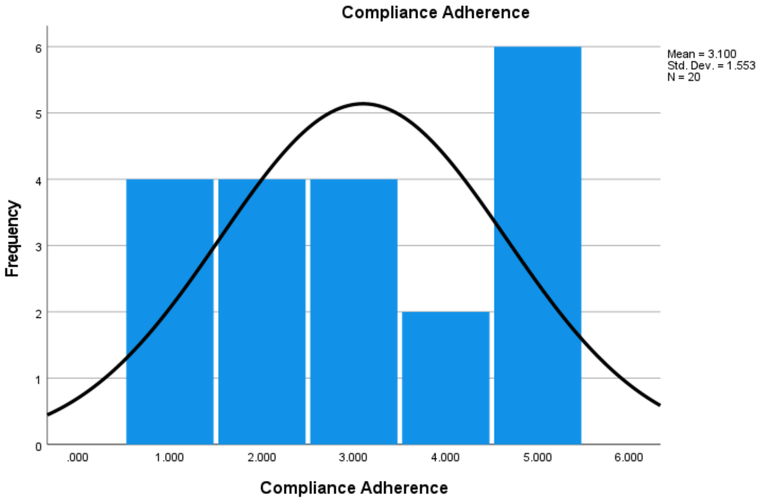


FIGURE 4. Compliance Adherence

Figure 4 presents a histogram illustrating the frequency distribution of Compliance Adherence scores. The mean value is 3.10, with a standard deviation of 1.553, based on N = 20 responses. The distribution appears somewhat symmetrical, with responses spread across the scale. The most frequent rating is 5, indicating that many respondents perceive high compliance adherence. However, there is noticeable variation, with some ratings as low as 1 or 2, suggesting mixed opinions. While compliance adherence is generally rated positively, the variability suggests room for improvement in ensuring consistent regulatory alignment across all users.

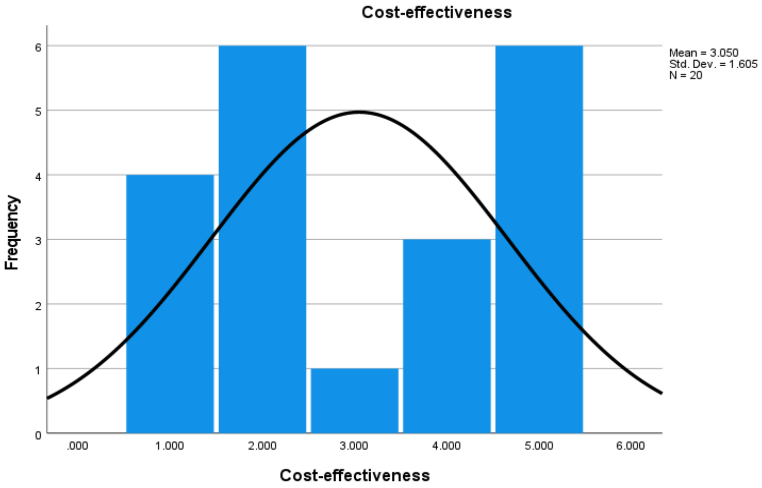


FIGURE 5. Cost-effectiveness

Figure 5 presents a histogram displaying the frequency distribution of Cost-effectiveness scores. The mean value is 3.05, with a standard deviation of 1.605, based on N = 20 responses. The distribution is moderately symmetrical, with ratings spread across the scale. The most frequent rating is 5, suggesting that many respondents perceive the system as highly cost-effective. However, there is notable variability, with some ratings as low as 1 or 2, indicating divergent opinions. This suggests that while cost-effectiveness is generally viewed positively, some users may feel the costs outweigh the benefits.

TABLE 5. Correlations

Correlations					
	Efficiency of Automation	Reduction in Response Time	Accuracy of Threat Detection	Compliance Adherence	Cost-effectiveness
Efficiency of Automation	1	0.191	0.006	0.18	-0.308
Reduction in Response Time	0.191	1	0.437	0.311	0.138
Accuracy of Threat Detection	0.006	0.437	1	0.431	0.001
Compliance Adherence	0.18	0.311	0.431	1	.484*
Cost-effectiveness	-0.308	0.138	0.001	.484*	1

Table 5 presents the correlation coefficients between the five measured variables: Efficiency of Automation, Reduction in Response Time, Accuracy of Threat Detection, Compliance Adherence, and Cost-effectiveness. Most correlations are weak to moderate, indicating limited relationships between variables. The strongest correlation (r = 0.484, p < 0.05) is between Compliance Adherence and Cost-effectiveness, suggesting that better compliance adherence is associated with higher cost-effectiveness. Accuracy of Threat Detection shows a moderate positive correlation with Compliance Adherence (r = 0.431) and Reduction in Response Time (r = 0.437), indicating a potential link between detection accuracy and system responsiveness. However, Efficiency of Automation has weak correlations with all variables, and its negative correlation with Cost-effectiveness (r = -0.308) suggests that perceived automation efficiency does not always translate into cost savings**.**

Regression:

TABLE 6. Model Summary

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	Durbin-Watson
Efficiency of Automation	.574[a]	0.329	0.15	1.299207	0.329	1.841	4	15	0.173	1.275
Reduction in Response Time	.520[a]	0.271	0.076	1.265663	0.271	1.392	4	15	0.284	1.999
Accuracy of Threat Detection	.638[a]	0.407	0.248	1.303382	0.407	2.571	4	15	0.081	2.232
Compliance Adherence	.733[a]	0.538	0.415	1.187972	0.538	4.363	4	15	0.015	1.926
Cost-effectiveness	.699[a]	0.488	0.352	1.292272	0.488	3.578	4	15	0.031	1.417

Table 6 presents the model summary statistics for the five dependent variables, highlighting the strength of relationships between predictors and outcomes. The R values indicate the strength of correlation, with Compliance Adherence (R = 0.733) and Cost-effectiveness (R = 0.699) showing the strongest relationships with the predictors. The R Square values indicate the proportion of variance explained by the model, where Compliance Adherence (0.538) and Cost-effectiveness (0.488) have the highest explanatory power. The Adjusted R Square values, which account for model complexity, show that Efficiency of Automation (0.15) and Reduction in Response Time (0.076) have weaker predictive models. Significance (Sig. F Change) indicates that Compliance Adherence (p = 0.015) and Cost-effectiveness (p = 0.031) are statistically significant, while others are not. The Durbin-Watson values (1.275–2.232) suggest no strong autocorrelation in residuals.

TABLE 7. ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Efficiency of Automation	12.431	4	3.108	1.841	.173[b]
Reduction in Response Time	8.921	4	2.23	1.392	.284[b]
Accuracy of Threat Detection	17.468	4	4.367	2.571	.081[b]
Compliance Adherence	24.631	4	6.158	4.363	.015[b]
Cost-effectiveness	23.901	4	5.975	3.578	.031[b]

Table 7 presents the ANOVA (Analysis of Variance) results, evaluating whether the independent variables significantly predict each dependent variable. The F-values and Significance (Sig.) values indicate the strength of relationships. Compliance Adherence (F = 4.363, p = 0.015) and Cost-effectiveness (F = 3.578, p = 0.031) show statistically significant results, meaning the independent variables significantly explain their variance. These models are the strongest predictors in the analysis. Other dependent variables—Efficiency of Automation (p = 0.173), Reduction in Response Time (p = 0.284), and Accuracy of Threat Detection (p = 0.081)—have non-significant results, meaning the predictors do not explain their variance effectively. Higher Sum of Squares values for Compliance Adherence (24.631) and Cost-effectiveness (23.901) indicate greater variability explained by the model.

Factor Analysis

TABLE 8. Communalities

Communalities		
	Initial	Extraction
Efficiency of Automation	1	0.634
Reduction in Response Time	1	0.579
Accuracy of Threat Detection	1	0.549
Compliance Adherence	1	0.708
Cost-effectiveness	1	0.816
Extraction Method: Principal Component Analysis.		

Table 8 presents the communalities of the variables, which indicate how much variance in each variable is explained by the extracted components in Principal Component Analysis (PCA). Higher values suggest stronger representation within the factor structure. Cost-effectiveness (0.816) and Compliance Adherence (0.708) have the highest extraction values, meaning they are well-explained by the principal components. Efficiency of Automation (0.634) and Reduction in Response Time (0.579) also have moderate representation, while Accuracy of Threat Detection (0.549) has the lowest value, suggesting it is less strongly captured by the extracted components. All initial communalities are 1, meaning full variance is initially considered, and the extraction values indicate the retained proportion. Cost-effectiveness and Compliance Adherence contribute most to the principal components, implying these factors play key roles in the dataset's structure. Other variables, particularly Accuracy of Threat Detection, may require additional factors for better representation.

TABLE 9. Total Variance Explained

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.943	38.856	38.856	1.943	38.856	38.856
2	1.344	26.883	65.739	1.344	26.883	65.739
3	0.875	17.51	83.249			
4	0.623	12.459	95.707			
5	0.215	4.293	100			
Extraction method: Principal component analysis						

The Total Variance Explained table presents the results of a Principal Component Analysis (PCA), which identifies the key components contributing to the dataset's variance. From the Initial Eigenvalues, the first two components have eigenvalues greater than 1, explaining 38.86% and 26.88% of the variance, respectively. Together, they account for 65.74% of the total variance, suggesting they capture most of the dataset's information. The remaining components have eigenvalues below 1, indicating they contribute less meaningful variance. For instance, Component 3 explains 17.51%, and Component 4 explains 12.46%, bringing the cumulative variance to 95.71%. Component 5 adds 4.29%, reaching 100% but is not retained due to its low eigenvalue. The Extraction Sums of Squared Loadings confirms that the first two components are retained, as they explain the most variance. This suggests that reducing the dataset to two principal components is a reasonable simplification while retaining most of the information.

4. CONCLUSION

The complexity of decision-making in the power sector necessitates a shift from traditional manual approaches to automated, AI-driven frameworks. This study introduces a machine learning-based automated framework designed to streamline decision-making analysis in power utilities and improve efficiency, security, and scalability. Implementation of this framework with two real-world operational datasets and extensive testing have provided compelling evidence of its effectiveness. One of the key findings of this study is that automation significantly reduces the time and effort required for decision-making, while maintaining high levels of accuracy. By integrating AI-driven decision support systems, power utilities can improve their ability to assess vulnerabilities and implement mitigation strategies in a timely manner. This is particularly important in an industry where delays in decision-making lead to increased costs, security risks, and service disruptions. Furthermore, this study highlights the role of automation in improving cybersecurity within power utilities. With the increasing number of cyber threats targeting critical infrastructure, manually managing security vulnerabilities is no longer a viable solution. The machine learning-based framework presented in this research automates the identification and prioritization of vulnerabilities, ensuring a proactive approach to security risk management. Another key advantage of the proposed solution is its scalability. As power utilities expand their operations and integrate new technologies, traditional decision-making methods become increasingly unsustainable. The automated framework is designed to meet evolving industry needs, providing a flexible and future-proof solution. Furthermore, this study advances the field of AI-driven infrastructure management by showing how machine learning can enhance judgment in areas other than cybersecurity. The study's conclusions can also be extended to other fields like predictive maintenance, grid optimization, and energy distribution. The sector has a revolutionary chance to embrace automation in power utility decision-making based on machine learning. The suggested framework offers a strong response to the difficulties faced by contemporary power utilities by increasing safety, cutting expenses, and increasing efficiency. Future studies should concentrate on enhancing the framework's functionality integrating real-time analytics, and exploring cross-sector applications to further improve decision-making in critical infrastructure sectors.

REFERENCES

[1]. Zhang, Fengli, Philip Huff, Kylie McClanahan, and Qinghua Li. "A machine learning-based approach for automated vulnerability remediation analysis." In 2020 IEEE Conference on Communications and Network Security (CNS), pp. 1-9. IEEE, 2020.

[2]. Mohamed, Abdel Mohsen O., Dina Mohamed, Adham Fayad, and Moza T. Al Nahyan. "Enhancing Decarbonation in Solid Waste Management, Wastewater and Contaminated Soil Treatments: The Role of Data Automation and Decision Support Systems." (2024).

[3]. Compastié, Maxime, Antonio López Martínez, Carolina Fernández, Manuel Gil Pérez, Stylianos Tsarsitalidis, George Xylouris, Izidor Mlakar, Michail Alexandros Kourtis, and Valentino Šafran. "Palantir: An nfv-based security-as-a-service approach for automating threat mitigation." Sensors 23, no. 3 (2023): 1658.

[4]. Harrell, Christopher R., Mark Patton, Hsinchun Chen, and Sagar Samtani. "Vulnerability assessment, remediation, and automated reporting: Case studies of higher education institutions." In 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), pp. 148-153. IEEE, 2018.

[5]. Namrud, Zakeya, Komal Sarda, Marin Litoiu, Larisa Shwartz, and Ian Watts. "Kubeplaybook: A repository of ansible playbooks for kubernetes auto-remediation with llms." In Companion of the 15th ACM/SPEC International Conference on Performance Engineering, pp. 57-61. 2024.

[6]. Forshaw, Matthew, Gerald Becker, Sanjib Jena, Christian Linke, and Olof Hummes. "Automated hole cleaning monitoring: A modern holistic approach for NPT reduction." In International Petroleum Technology Conference, p. D033S245R001. IPTC, 2020.

[7]. Harzevili, Nima Shiri, Alvine Boaye Belle, Junjie Wang, Song Wang, Zhen Ming, and Nachiappan Nagappan. "A survey on automated software vulnerability detection using machine learning and deep learning." arXiv preprint arXiv:2306.11673 (2023).

[8]. Deb, Supratim, Zihui Ge, Sastry Isukapalli, Sarat Puthenpura, Shobha Venkataraman, He Yan, and Jennifer Yates. "Aesop: Automatic policy learning for predicting and mitigating network service impairments." In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1783-1792. 2017.

[9]. McClanahan, Kylie, and Qinghua Li. "Automatically locating mitigation information for security vulnerabilities." In 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), pp. 1-7. IEEE, 2020.

[10]. Pan, Shengyi, Lingfeng Bao, Jiayuan Zhou, Xing Hu, Xin Xia, and Shanping Li. "Towards More Practical Automation of Vulnerability Assessment." In Proceedings of the IEEE/ACM 46th International Conference on Software Engineering, pp. 1-13. 2024.

[11]. Sarda, Komal, Zakeya Namrud, Marin Litoiu, Larisa Shwartz, and Ian Watts. "Leveraging Large Language Models for the Auto-remediation of Microservice Applications: An Experimental Study." In Companion Proceedings of the 32nd ACM International Conference on the Foundations of Software Engineering, pp. 358-369. 2024.

[12]. Malaiyappan, Jesu Narkarunai Arasu, Sanjeev Prakash, Samir Vinayak Bayani, and Munivel Devan. "Enhancing cloud compliance: A machine learning approach." AIJMR-Advanced International Journal of Multidisciplinary Research 2, no. 2 (2024).

[13]. Kumar, Manoj, and Arun Sharma. "An integrated framework for software vulnerability detection, analysis and mitigation: an autonomic system." Sādhanā 42 (2017): 1481-1493.

[14]. Kram, Mark L., Blayne Hartman, and Cliff Frescura. "Vapor intrusion monitoring method cost comparisons: Automated continuous analytical versus discrete time-integrated passive approaches." Remediation Journal 26, no. 4 (2016): 41-52.

[15]. Lekidis, Alexios, Vasileios Mavroeidis, and Konstantinos Fysarakis. "Towards incident response orchestration and automation for the advanced metering infrastructure." In 2024 IEEE 20th International Conference on Factory Communication Systems (WFCS), pp. 1-8. IEEE, 2024.

[16]. Pargaonkar, Shravan. "Advancements in security testing: A comprehensive review of methodologies and emerging trends in software quality engineering." International Journal of Science and Research (IJSR) 12, no. 9 (2023): 61-66.

[17]. Bhardwaj, Arvind Kumar, P. K. Dutta, and Pradeep Chintale. "Securing Container Images through Automated Vulnerability Detection in Shift-Left CI/CD Pipelines." Babylonian Journal of Networking 2024 (2024): 162-170.

[18]. 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.

[19]. 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.

[20]. 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.

[21]. 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.

[22]. Thoemmes, Felix. "Propensity score matching in SPSS." arXiv preprint arXiv:1201.6385 (2012).

[23]. 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.

[24]. 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.

[25]. 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.

[26]. Von Hippel, Paul T. "Biases in SPSS 12.0 missing value analysis." The American Statistician 58, no. 2 (2004): 160-164.

[27]. 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.

[28]. 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.

[29]. Liang, Guangping, Wenliang Fu, and Kaifa Wang. "Analysis of t-test misuses and SPSS operations in medical research papers." Burns & trauma 7 (2019).