Transforming Cybersecurity into Critical Energy Infrastructure: A Study on the Effectiveness of Artificial Intelligence

Jaime Govea 1, Walter Gaibor-Naranjo 2 and William Villegas-Ch 1,*

1 Escuela de Ingeniería en Ciberseguridad, Facultad de Ingenierías y Ciencias Aplicadas, Universidad de Las Américas, Quito 170125, Ecuador; jaimealejandro.govea@udla.edu.ec
2 Carrera de Ciencias de la Computación, Universidad Politécnica Salesiana, Quito 170105, Ecuador; wgaibor@ups.edu.ec*

Correspondence: william.villegas@udla.edu.ec; Tel.: +593-98-136-4068

Abstract: This work explores the integration and effectiveness of artificial intelligence in improving the security of critical energy infrastructure, highlighting its potential to transform cybersecurity practices in the sector. The ability of artificial intelligence solutions to detect and respond to cyber threats in critical energy infrastructure environments was evaluated through a methodology that combines empirical analysis and artificial intelligence modeling. The results indicate a significant increase in the threat detection rate, reaching 98%, and a reduction in incident response time by more than 70%, demonstrating the effectiveness of artificial intelligence in identifying and mitigating cyber risks quickly and accurately. In addition, implementing machine learning algorithms has allowed for the early prediction of failures and cyber-attacks, significantly improving proactivity and security management in energy infrastructure. This study highlights the importance of integrating artificial intelligence into energy infrastructure security strategies, proposing a paradigmatic change in cybersecurity management that increases operational efficiency and strengthens the resilience and sustainability of the energy sector against cyber threats.

Keywords: artificial intelligence in cybersecurity; critical energy infrastructure; cyber threat detection

1. Introduction

The accelerated advance of cyber threats in critical energy infrastructure constitutes a significant challenge to global security. These threats put energy services’ continuity and efficiency at risk and threaten economic and social stability. Against this backdrop, this work addresses the critical problem of improving the detection of and response to cyber incidents in energy infrastructure by integrating artificial intelligence (AI) solutions. This work innovates by applying and evaluating advanced AI techniques, such as deep learning and predictive analytics, in the context of critical energy infrastructure, thus providing an effective and proactive solution to a persistent and evolving problem [1].

The relevance of this study focuses on increasingly sophisticated cyber threats, where energy infrastructure represents a critical target. Therefore, we explore how AI can be an advanced tool used to anticipate, detect, and neutralize cyber risks, ensuring the continuity and reliability of energy services [2]. A review of the literature reveals a growing body of research in AI applied to cybersecurity. However, significant gaps persist in studies specific to the energy sector [3]. In methodological terms, a multidisciplinary approach combines data analysis, AI modeling, and simulations of cyber threat scenarios. This methodology is justified by its comprehensive and realistic assessment of AI’s ability to improve energy security [4]. Advanced data analysis instruments and machine learning algorithms have been used to process and examine large volumes of energy infrastructure data, enabling accurate assessments of AI systems’ detection and response capabilities [5].
Among the significant findings is progress in the protection of critical energy infrastructure through the implementation of AI. The most notable results include a 98% improved threat detection rate and a more than 70% reduction in response time to cyber incidents, marking a significant milestone in the efficiency and effectiveness of cybersecurity measures. Furthermore, implementing machine learning algorithms has allowed for a more accurate prediction of possible failures and cyber-attacks, improving the proactivity and responsiveness of the security system [6]. These results reflect the potential of AI to transform the security of energy infrastructure and suggest a paradigmatic shift in how cyber risks are addressed and managed in this vital sector.

2. Literature Review

The works in this review range from developing advanced AI algorithms to their practical implementation in critical infrastructure environments, highlighting technological advances and associated challenges [7]. One of the fundamental aspects of the literature is identifying specific cyber threats facing the energy sector [8]. Studies such as those by Ameri et al. demonstrate how vulnerabilities in Industrial Control Systems (ICSs) [9] can be exploited to cause significant disruptions. This work highlights the need for AI solutions that can proactively detect and neutralize threats, serving as a basis for subsequent research that focuses on developing anomaly-detection algorithms specific to these environments.

The literature shows significant progress in using machine learning and deep learning techniques in threat detection [10,11]. For example, Siva Kumar et al. [12] have presented a deep neural network-based approach for detecting abnormal patterns in critical infrastructure network traffic data. This approach highlights the ability of AI models to learn from large data sets and adapt to new threats, offering a high degree of precision and efficiency in detecting cyber-attacks.

Furthermore, integrating AI into energy infrastructure is not limited to threat detection but encompasses incident response and mitigation [13]. Research by Rizvi [14] illustrates how AI systems can automate incident response, significantly reducing response time and minimizing potential damage. These systems use real-time decision-making algorithms to execute corrective actions, demonstrating the potential of AI to improve the resilience of critical infrastructure.

Data reliability and security are also recurring themes in the literature [15,16]. In this context, Wazid et al.’s [10] work examines the challenges associated with data security in implementing AI, highlighting the importance of data protection and privacy in IT systems and critical infrastructure. This study highlights the need for robust cybersecurity approaches that ensure the integrity and confidentiality of data used by AI systems.

A review of the literature also reveals the importance of a holistic approach, considering AI implementation’s technical and socio-economic aspects in energy infrastructure [17]. Works such as Wenninger et al. [18] discuss how AI solutions can influence energy policies and grid management. They suggest that AI adoption must be accompanied by strategic planning and effective management to maximize its positive impact on the sector.

Adapting to rapid evolutions in the cyber threat landscape and taking advantage of AI technology advances require a commitment to innovation and continuous improvement [19,20]. Studies advocating an open innovation framework highlight this aspect, where collaboration between industries, academics, and governments can accelerate the development and implementation of effective and safe AI solutions.

3. Materials and Methods

The method was designed to provide an understanding of the tools, techniques, and processes used in the research and application of AI solutions in critical energy infrastructure. To this end, we address the selection of AI technologies, evaluation criteria, and specific implementation procedures that adapt to the unique challenges of the energy sector.
This methodological framework pursues a systematic and rigorous approach to researching and applying AI but also establishes the basis for further experimentation and analysis, allowing us to effectively evaluate the effectiveness and impact of these advanced technologies in improving the security and efficiency of critical energy infrastructure.

### 3.1. Definition of the Problem

The core of the problem in critical energy infrastructure lies in its increasing exposure to advanced cyber threats, which is exacerbated by the interconnection of systems and reliance on digital technologies. These threats evolve in sophistication and speed, requiring an equally rapid and efficient response to prevent significant damage to infrastructure and, by extension, the society and economy that depend on it.

The proposed solution to this problem involves integrating advanced technologies capable of anticipating, detecting, and responding to security incidents autonomously and in real time. This is where AI comes into play, promising to transform the cybersecurity paradigm in critical infrastructure [21]. AI, with its ability to analyze large volumes of data, learn from it, and make decisions based on complex patterns, is presented as a solution to overcome the limitations of traditional security systems.

Specifically, we face attacks targeting ICSs, Advanced Persistent Threats (APTs), spear phishing, and ransomware attacks. Each of these challenges requires a solution that combines advanced detection, analysis, and rapid response, roles in which AI is incredibly proficient. We used advanced machine learning and data analysis techniques to identify anomalous patterns and suspicious behavior that indicate intrusion attempts or malicious activity. These AI systems operate under conditions that require high precision and speed to be effective, such as real-time detection and continuous monitoring, enabling a rapid, automated response that minimizes potential damage and improves infrastructure resilience [22].

The objective focuses on integrating AI into the critical energy infrastructure network’s cyber defense system to improve threat detection and automate responses. Adopting AI can provide advanced analysis and response capabilities, improving system resilience to cyber-attacks and reducing operational downtime. At the same time, a balance is sought between technological implementation and the management of potential risks associated with reliance on automated systems, ensuring that the integration of AI into cyber defense is both practical and secure. Specific AI technologies are selected and configured to address cyber threats to critical energy infrastructure, integrating with existing systems and managing data securely and effectively for threat detection.

### 3.2. Analysis of the Critical Energy Infrastructure Network

Critical energy infrastructure is a complex interconnected network encompassing energy generation, transmission, and distribution. This network is essential for the functioning of society and the economy. It is characterized by its wide geographical distribution, interconnection of advanced technological components, and interdependence with other vital sectors such as telecommunications, water, and financial services.

The network incorporates ICSs and operational technologies (OTs) integrated with information technology (IT) systems to optimize operation and monitoring. These systems operate using specific protocols, such as supervisory control and data acquisition (SCADA), distributed Network Protocol 3 (DNP3), and Modbus, which are crucial for control and data acquisition in real-time.

Infrastructure depends on various connected devices, from simple sensors and actuators to complex process control systems. These devices are distributed throughout the network, from generation plants to consumption points, forming a network of nodes that can number thousands in an extensive network, each with its function and level of criticality.

The network faces significant challenges regarding cybersecurity due to its high automation and digitalization. The increasing connection of these systems to the Internet
and enterprise networks to facilitate remote control and real-time monitoring expand the
attack surface [23]. This exposes the network to various cyber threats. These risks include
targeted attacks on critical network components, such as substations and control centers,
using tactics such as malware, ransomware, and denial of service (DoS) attacks that can
disrupt power supply and cause cascading effects.

The energy infrastructure network is a complex network that extends nationwide,
with a geographic scale that ranges from generation plants in remote areas to substations
and distribution centers in urban centers. The network consists of several nodes, including:

- Generation plants are power stations that produce electricity from fossil fuels, hydro-
electric, solar, and wind sources.
- Substations are facilities that transform the voltage levels of electricity for distribu-
tion and transmission throughout the network.
- Distribution centers are nodes that distribute electrical energy to final consumers,
  adjusting supply according to demand.

Every node within the network is a critical point that cyber-attacks can target. Poten-
tial incidents range from an intrusion and manipulation of operations to complete a inter-
ruption of services. The interconnection of these nodes creates a densely woven network
where an incident in a single location can quickly propagate its effects, underscoring the
importance of a robust, multi-layered cybersecurity strategy.

An incident in one part of this interconnected grid can have cascading effects
throughout the system, destabilizing the regional power grid and causing widespread
blackouts. Integrating legacy systems with new technologies has created a heterogeneous
environment with different levels of cybersecurity, generating vulnerabilities that are dif-
ficult to detect and mitigate [24]. Additionally, the reliance on third parties to maintain
and provide network components increases the risk of attacks through the supply chain.

Cyber incidents affecting the energy sector include supply chain attacks, DoS attacks,
and others, which can significantly impact the operation and security of energy infrastruc-
ture. A business impact analysis, supported by statistics and case studies, highlights the
sector’s susceptibility to these cyber threats.

A reliance on third parties to maintain and supply network components also in-
creases the risk of attacks through the supply chain. Although arguably better funded than
other critical sectors, such as healthcare, the energy sector faces unique challenges that
require advanced and targeted cybersecurity solutions. Implementing AI in this context
shows promise, with AI systems designed to analyze large volumes of data generated by
the network in real time, identify abnormal behavior patterns, and facilitate automated
responses to security incidents [25].

3.3. Description of the AI Technology Implemented

AI techniques primarily identify and mitigate cyber-attacks within critical energy in-
frastucture. The use of neural networks, for example, is not limited to a generic type but
includes specific variants such as convolutional neural networks (CNNs) for the pro-
cessing and analyzing of visual and traffic data. At the same time, long-short-term
memory (LSTM) is preferred for predicting attacks based on time series analysis [26].

In the domain of decision trees, specialized techniques such as C4.5 and ID3 classify
and predict the nature of cyber threats, enabling rapid and informed responses. These
methodologies allow for vast network data to be broken down into actionable information
[27].

Beyond predictive analytics, implementing Natural Language Processing (NLP) en-
commasses advanced techniques for interpreting unstructured data. Tools like GPT and
BERT, under the domain of NLP, are essential for analyzing communications and detecting
threat indicators in free text, allowing for a deeper understanding of adversaries’ tact-
cies.
Table 1 specifies the evaluation of AI technologies implemented in the cyber defense of critical energy infrastructures. It offers a comparison that quantifies aspects such as detection efficiency, integration with existing systems, scalability, resource management, and cost. This comparative analysis originates from research that combines theoretical analysis, empirical tests, and case studies in real cybersecurity scenarios.

<table>
<thead>
<tr>
<th>AI Technology</th>
<th>Detection Efficiency</th>
<th>Integration with Existing Systems</th>
<th>Scalability</th>
<th>Resource Management</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Networks</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
<tr>
<td>CNN</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
<tr>
<td>LSTM</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
<tr>
<td>GNN (Graph Neural Network)</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
<tr>
<td>Decision Trees</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Moderate</td>
<td>Medium</td>
</tr>
<tr>
<td>C4.5</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Moderate</td>
<td>Medium</td>
</tr>
<tr>
<td>ID3</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Moderate</td>
<td>Medium</td>
</tr>
<tr>
<td>Support Vector Machines (SVMs)</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Moderate</td>
<td>Medium</td>
</tr>
<tr>
<td>Natural Language Processing (NLP)</td>
<td>Medium</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
<tr>
<td>GPT (Generative Pretrained Transformer)</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
<tr>
<td>BERT (Bidirectional Encoder Representations from Transformers)</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Intensive</td>
<td>High</td>
</tr>
</tbody>
</table>

To measure detection efficiency, penetration testing and attack simulation were conducted in controlled environments, each using AI technology to identify its ability to detect and react to various cyber threats. These tests included injecting malicious traffic and attack patterns into the network to evaluate how each AI model identified and classified threats in real time [28].

Integration with existing systems was evaluated by implementing each AI technology into the energy grid’s operational IT infrastructure. The compatibility of AI technologies with ICSs and OT was analyzed, noting the ease of integration, the need for infrastructure modifications, and the interaction with existing monitoring and incident response systems. The scalability of the AI technologies was tested by progressively increasing the data load and processing demands to see how each AI system adapted to changes in the volume of network traffic and the number of nodes and devices monitored. This included simulating scenarios of network growth and increased network activities to determine each AI technology’s ability to scale without degrading performance.

Regarding resource management, CPU memory, and storage efficiency were analyzed, and resource consumption was measured during normal operations and under high-workload conditions. This helped identify resource-intensive AI technologies and how this affected their long-term viability in operational environments. The cost was evaluated by considering the initial investment in the technology and infrastructure necessary for its implementation and the long-term operational costs, including maintenance, upgrades, and energy consumption.
3.4. Implementation Methodology

The methodology adopted in our study is an iterative procedure that adjusts and evolves in response to ongoing findings and operational feedback. It begins with a detailed needs assessment and planning, in which current threats are identified, energy infrastructure vulnerabilities are assessed, and precise goals for AI integration are defined, as presented in Figure 1.

- The first phase assesses the existing cyber defense infrastructure to identify its strengths and weaknesses, understand the threats to which it is exposed, and clearly define the requirements and objectives for AI integration. This phase allows us to align project expectations with the system’s capabilities and operational needs.
- Subsequently, the AI tools configuration phase involves selecting and customizing AI solutions. Algorithms and platforms are chosen based on their ability to satisfy the identified requirements and are configured to adapt to the specific environment of the cyber defense system. The configuration ranges from adapting machine learning algorithms to integrating NLP systems, ensuring that each AI component is optimized for the context in which it will be deployed.
- Integration with existing systems is the next phase, where the configured AI tools are assembled within the operational framework of the cyber defense system. This integration must be seamless, allowing AI solutions to effectively interact with existing systems, exchange data, and provide real-time analysis and responses. This phase ensures that AI infrastructure coexists and collaborates effectively with already deployed cybersecurity tools.
- The testing and validation stage focuses on evaluating the effectiveness of the AI integration. During this phase, rigorous testing is carried out to ensure that the AI solutions perform as expected in the actual operating environment. Testing includes simulating cyber-attacks to verify AI systems’ detection and response capabilities and evaluating precision and efficiency in threat management.
- Effectiveness assessment involves a critical analysis of data and performance metrics collected during testing and actual operation to determine the impact of AI on
improving cybersecurity. It evaluates how AI tools have improved threat detection, reduced incident response time, and contributed to the overall cyber defense strategy.

- For our methodology, a refinement and adjustment loop are introduced. This loop is activated if the effectiveness evaluation indicates performance below the desired threshold. The AI technologies are then fine-tuned using performance data and effectiveness metrics collected during testing. The adjustments focus on improving precision, reducing false positives, and adapting AI systems to better respond to emerging threats. This iteration is repeated until the AI systems reach and maintain a level of performance that meets our rigorous criteria for effectiveness and efficiency. At the end of the iterative cycle, AI systems that demonstrate a robust ability to detect and respond to threats in real time and under diverse operating conditions are selected for long-term deployment.

Given the energy sector’s strategic importance and potential vulnerabilities to cyber threats, AI tools were selected and adapted to ensure operational effectiveness and seamless integration with existing industrial control systems. Recognizing these systems’ complexity and high-risk exposure, specific modifications to the AI tools included optimizing algorithms to discern between normal operational fluctuations and abnormal network traffic patterns indicative of malicious cyber activity.

Studies such as those by Ameri et al. [9] and Siva Kumar et al. [12] illustrate the complexities and risks associated with industrial control systems and network infrastructure in the energy sector. AI tools must be integrated precisely to maintain operational stability and security, a critical aspect that Jadidi et al. [21] examined by detailing effective cybersecurity management strategies. Seamless integration, prior cleaning, and data normalization are essential for AI systems to function effectively.

Modifications were made to adapt these algorithms to the energy sector to accommodate the peculiarities of its data and operations. This included calibrating models to recognize normal power flow fluctuations and distinguish them from potential cyber threats [29]. Integration with operational energy systems was carefully managed to ensure seamless communication and coordinated response between AI and industrial control systems, maintaining energy grid stability and operational security.

3.5. Data Collection and Analysis

Data collection for our study was based on a combination of public and private data sources selected for their relevance and reliability in the context of cyber threats to critical energy infrastructure. Public sources include well-known cyber threat databases such as the National Vulnerability Database (NVD) in the United States [30], which provides a compendium of security vulnerabilities identified in various systems. In addition, event logs and alerts from publicly accessible network monitoring platforms were used, such as those provided by the Security Information Exchange project of the SANS Internet Storm Center and the early warning system of national CERTs [31]. Collaborations with private entities and industry associations allowed for access to detailed operational data and incident records specific to the energy sector under confidentiality agreements.

The data analyzed come from real environments and controlled simulations, reflecting a broad spectrum of cyber threats to critical energy infrastructure. Sector-specific operational and incident logs were obtained through collaborations with private entities under strict confidentiality agreements and complemented by threat intelligence analysis provided by CrowdStrike [32] and FireEye [33], which analyze and monitor cyber adversaries’ tactics, techniques, and procedures. This data set includes attacks targeting ICSs, APTs, spear phishing tactics, and ransomware. The emulated data were generated in testbeds designed to simulate specific attack scenarios, providing a comprehensive view of AI detection and response capabilities against a diversified spectrum of cyber threats,
underscoring the depth and relevance of our analysis in the context of cybersecurity of energy infrastructures.

The data collection process was designed to capture a comprehensive and accurate view of cyber activities within critical energy infrastructure. Automated tools, including advanced packet capturers and Security Information and Event Management (SIEM) systems, were implemented to monitor network traffic and log security events. These tools were configured to ensure the systematic capture and aggregation of data in real time, allowing for the collection of detailed information on communications within the network, intrusion attempts, and detected anomalies.

To ensure data representativeness, collection points were strategically deployed across multiple network nodes, providing a data set that accurately reflects the complete picture of network operations and its security challenges. This multidimensional approach not only improved the quality of the data set but also increased the relevance of subsequent analyses by providing a meaningful and representative sample of the environment being protected. Data integrity was maintained by implementing validation and verification protocols, including integrity testing and regular audits. These protocols ensured that any data collected were complete, accurate, and free of corruption or manipulation before being used to train and test the AI models. The systematic approach and automated tools enabled efficient and structured data collection, thus forming the basis for detailed cybersecurity analysis and the effective development of AI solutions tailored to the specific needs of the cyber defense environment.

Over a representative six-month period, approximately 100 terabytes (TB) of data were collected, covering various cyber activities and threats to ensure the comprehensive analysis and practical training of AI models. For network traffic, around 15 TB of data were monitored monthly, totaling 90 TB at the end of the collection period. This collection was performed across multiple network nodes within the organization, capturing a full spectrum of network activity, including regular traffic and potential threats.

Regarding security event logs generated by devices such as firewalls and intrusion detection systems, about 2 TB of data accumulate monthly, and up to 12 TB during the six months. These logs are essential for analyzing security events and alerts generated in response to suspicious or malicious activity. In addition, application and operating system logs are collected, adding up to approximately 500 gigabytes (GB) per month, reaching 3 TB in the study period. These data provide detailed information about system behavior and user activities, which is crucial for detecting potential internal security breaches.

Threat intelligence, obtained from public and private sources, provides around 5 GB of updated data each month, reaching 30 GB at the end of six months. This information enriches the database with up-to-date knowledge on the tactics, techniques, and procedures (TTPs) of emerging threats, which is vital for the fine-tuning and precision of AI models in threat detection. The collection was performed using automated tools, such as packet capturers for network traffic and SIEM systems for logs, ensuring an efficient and systematic process. All the collected data are stored in secure repositories, with controlled access and protection measures to maintain their integrity and confidentiality. Additionally, validation protocols are established to verify the completeness and precision of the data, including integrity verification and audits of regularly collected data. This approach to data collection ensures a diverse and representative data set, essential for detailed cybersecurity analysis and an effective development of AI solutions tailored to the specific needs of the cyber defense environment.

3.5.1. Data Preparation and Preprocessing

In the data preprocessing phase, we employed methods to transform the collected raw data into a uniform and standardized format, making them suitable for advanced analysis using AI algorithms. Data collected from various sources, such as event logs, network traffic, and system logs, were converted to a standard format such as CSV or JSON to facilitate consistent handling. Using data integration tools like Logstash, we analyzed
and transform heterogeneous data into a standardized structure. For example, security event logs, initially in proprietary formats specific to each security tool and the network traffic data captured in PCAP formats, were uniformly converted to CSV. This standardization guarantees consistency in data handling between different types of information in subsequent analytical phases [34].

Numerical values within the data, such as response times and traffic volumes, underwent min–max normalization, adjusting the values to fall within a standardized range of 0 to 1. This normalization process facilitates the comparative analysis of various data types on a consistent scale and mitigates bias due to magnitude disparities. After normalization, data centering is performed by subtracting the mean of each feature and aligning the data set around a center value of zero. This centering improves the ability of AI models to learn from patterns and deviations rather than absolute values.

Additionally, we normalized the variance by dividing each feature by its standard deviation, ensuring uniformity in feature variability [35]. This standardization is essential to prevent an individual feature from disproportionately influencing AI models due to their scale. These preprocessing steps are automated through scripts in programming environments such as Python, using libraries such as Pandas and Scikit-learn, which provide comprehensive functionalities for efficient data transformation.

The data cleaning began with identifying and removing duplicate records, which can distort analytical results by overrepresenting certain events [36]. Automated algorithms scan the data set and identify duplicates based on unique identifiers or critical attributes, such as timestamp, event type, and source in the event logs. These redundant entries are subsequently removed to maintain analytical precision.

Addressing incomplete data is another critical step, where missing values are managed through imputation or deletion based on the context and proportion of missing data. For minor cases of missing data, imputation techniques such as mean, median, mode replacement, or more sophisticated model-based imputation are used to preserve the integrity of the data set. Conversely, a large amount of missing data on a feature might require removal from the data set to avoid analysis distortion [37].

We implemented validation protocols to ensure the representativeness and quality of the data. These protocols included integrity checks and periodic audits of the data collection process to verify its completeness and precision. The data preparation and preprocessing not only optimized the data set for AI analysis but also strengthened the foundation of our cybersecurity study, ensuring that the AI models were trained and tested on data that accurately reflect the real-world operating environment.

To ensure an accurate comparison in our analysis, we must highlight that both the AI algorithms and traditional analysis methods were evaluated using the same data set, which was preprocessed and cleaned. This uniform approach ensures that the evaluation accurately reflects the inherent effectiveness of each algorithm in identifying and mitigating cyber threats, eliminating any confounding variables that could arise from the quality of the data set. In this way, we can affirm that the observed differences in performance are due to the distinctive characteristics and capabilities of the algorithms evaluated, thus providing an unbiased and equitable analysis.

3.5.2. Analysis of Data

Statistical analysis began with applying descriptive statistics to obtain a basic understanding of the data. The mean, median, standard deviation and interquartile ranges were calculated for each numerical characteristic in the data. This provided an overview of the distribution and variability of the data. For example, the average amount of traffic per hour was analyzed for network traffic data, identifying spikes and dips that may suggest abnormal behavior or cyber-attacks [38].

Inferential statistical methods, such as hypothesis testing and regression analysis, were used to determine whether the data observations result from genuine patterns or random variations. In the cybersecurity context, this could involve analyzing whether an
increase in network traffic is correlated with known malicious activity or is part of normal behavior.

The model’s hyperparameters were tuned to classify malicious activities in network traffic data. Hyperparameters are configurations external to the model that influence its behavior and performance. Optimizing them is crucial to improving a model’s ability to detect threats accurately.

In addition to these statistical methods, hyperparameter tuning in models such as random forests was performed to optimize the classification of malicious activities in network traffic data. Using cross-validation and grid search techniques, this adjustment allowed us to determine the optimal configuration to improve the model’s precision and efficiency. Although random forest is specifically mentioned, this approach was uniformly applied to the other machine learning algorithms evaluated in the results. This ensures a solid comparative basis for evaluating their performance in detecting cyber threats.

For random forests, key hyperparameters include the following:

- **Number of Trees (n_estimators):** This refers to the number of trees in the forest. A significant number can improve the model’s precision and increase the computational cost. Experiment with 100, 200, or 500 values to balance performance and efficiency.

- **Maximum Tree Depth (max_depth):** This hyperparameter limits the depth of each tree. A more considerable depth allows the model to capture more detail but can also lead to overfitting. Different depths, such as 10, 20, or None (no limit), are tested to determine the optimal level.

- **Minimum Samples to Split (min_samples_split):** This parameter indicates the minimum number of samples necessary to split a node. Typical values are 2, 5, or 10. A lower value allows the model to be more detailed but can cause overfitting.

- **Minimum Samples per Leaf (min_samples_leaf):** The minimum number of samples required to be a tree leaf. Setting this to a more significant value can smooth the model and prevent overfitting.

Cross-validation and grid search techniques tune these hyperparameters. This method evaluates and compares model performance with different combinations of hyperparameters, using a training set divided into parts (e.g., 5 or 10 parts in k-fold cross-validation) to ensure that the fit is generalizable to unseen data [39].

For machine learning analysis, algorithms are implemented and trained on the preprocessed and cleaned data set to identify patterns and predict future threat behavior. Classification models are selected to detect malicious activities using algorithms such as decision trees, random forests, and support vector machines.

The process begins with dividing the data set into training and test sets, usually in a ratio of 80/20 or 70/30. The training set trains the models, adjusting their parameters to optimize their classification and prediction ability. For example, a random forest model is trained with network traffic data and security event logs to identify cyber-attack patterns. Subsequently, the performance of these models is validated and evaluated using the test set, applying metrics such as precision, sensitivity (recall), and F1 score [40]. Additionally, k-fold cross-validation ensures that the model is generalizable and robust to new data.

Precision is defined as the proportion of true positives (TPs) among all predicted positives (TPs and false positives, FPs), calculated as follows:

\[
Precision = \frac{True \ positives}{True \ positives + False \ positives}
\]

Recall measures the proportion of true positives correctly identified concerning the total number of confirmed positive cases (true positives and false negatives, FN) and is calculated as follows:

\[
Recall = \frac{True \ positives}{True \ positives + False \ negatives}
\]
The F1 score is the harmonic average of precision and sensitivity, balancing these two metrics. It is calculated as follows:

\[
F1 \text{ Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\] (3)

3.5.3. Evaluation of the Operational Effectiveness of AI in Cybersecurity

The effectiveness of AI in cybersecurity situations is evaluated not only by its precision in detecting and classifying threats but also by its performance in real or simulated operational environments. In this context, this evaluation focuses on how AI effectively contributes to the comprehensive cybersecurity management and decision-making process.

- Practical Application and Response Time: We measured the speed and effectiveness of AI systems responding to detected threats, evaluating the time elapsed from detection to implementing an appropriate response. This metric is vital for mitigating damage and strengthening incident response strategies.
- Managing False Positives in Operational Environments: It is crucial to evaluate how often AI systems incorrectly identify legitimate activities as malicious. A high rate of false positives can create unnecessary operational burdens and divert resources from real threats. We analyzed how these false positives are managed and reduced in the operational context to improve efficiency and minimize disruption.
- Adaptability and Continuous Learning: We implemented periodic testing and adjustments based on updated data analysis to measure AI’s adaptability to new threats and changes in the cyber environment. This dynamic approach allowed us to constantly identifying areas for improvement, optimizing both threat detection precision and response efficiency.
- Feedback and Continuous Improvement Mechanism: We established a feedback mechanism that integrates learnings from real cybersecurity operations, allowing for AI systems to refine and evolve based on practical experiences. This continuous improvement ensures that AI not only stays up to date with the tactics, techniques, and procedures of emerging threats but also aligns with the specific operational needs of the cybersecurity environment.

Therefore, assessing AI’s operational effectiveness is a holistic process encompassing technical precision, operational efficiency, and strategic adaptability. This approach ensures that AI systems are evaluated in a context that reflects their real-world application in cybersecurity, providing a comprehensive view of their performance and contribution to the security of critical infrastructure.

4. Results

The results highlight AI technologies’ effectiveness and transformative impact in strengthening the security of critical energy infrastructure. We have made significant progress in accurately detecting and rapidly responding to cyber threats by implementing advanced AI solutions. The case studies analyzed concretely illustrate how these technologies not only theorize possible solutions but also execute them effectively, addressing the energy sector’s unique challenges. These specific examples demonstrate the practical application of our research and methodologies in real operational environments, providing valuable insight into the real-world impact of our AI solutions in improving the safety and resilience of energy infrastructure.

4.1. Evaluation of Data Processing

Data transformation and cleansing have made it easier to identify and detect cyber threats accurately using AI algorithms. The processes began with evaluating the original data set, identifying and removing 30,000 duplicate records, and reducing the total from
1,200,000 to 1,150,000 valid records. This step allowed us to eliminate redundancies and possible biases in the subsequent analysis. Identifying 100,000 missing values and their subsequent treatments through imputation or elimination techniques ensures the analyzed data’s integrity and continuity. Additionally, 20,000 anomalies or outliers were detected and corrected, thus normalizing the data and avoiding distortions in the analysis results. This cleaning process also improved the consistency of data formats, moving from low uniformity to high consistency, which facilitated integration and analysis in later stages.

Regarding the content of the logs analyzed, these consisted of detailed security event data, including access logs, intrusion alerts, and notifications from anomaly-detection systems, covering a wide range of cyber threats in the energy infrastructure environment. These threats included brute force attacks, malware infiltrations, spear phishing activities, and abnormal behaviors indicative of vulnerability exploitation attempts. For each record, details such as the timestamp of the event, the nature of the suspicious activity, and the response of the security system were captured and analyzed. The traditional detection method was based on signature systems and heuristics that, although efficacious against known threats, showed limitations against new and sophisticated attacks. In contrast, the AI models deployed leveraged advanced machine learning techniques to analyze patterns within the data, significantly improving threat detection by adapting to the evolving cybersecurity landscape, as demonstrated in the higher detection rate and reduction in response times.

Preprocessing included normalizing the numerical values and scaling them to a range of 0 to 1. This adjustment not only homogenized the data but also optimized the efficiency of the AI algorithms during training and evaluation, allowing for further comparisons and analysis. Accurate, detailed results are presented in Table 2.

<table>
<thead>
<tr>
<th>Evaluated Aspect</th>
<th>Result Before Treatment</th>
<th>Result After Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of Data (records)</td>
<td>1,200,000</td>
<td>1,150,000</td>
</tr>
<tr>
<td>Identified Duplicates</td>
<td>30,000</td>
<td>0</td>
</tr>
<tr>
<td>Missing Values</td>
<td>100,000</td>
<td>0</td>
</tr>
<tr>
<td>Anomalies/Outliers Detected</td>
<td>20,000</td>
<td>0</td>
</tr>
<tr>
<td>Format Consistency</td>
<td>Low</td>
<td>high</td>
</tr>
<tr>
<td>Numeric Value Range</td>
<td>Varied (unnormalized)</td>
<td>0–1 (normalized)</td>
</tr>
</tbody>
</table>

To understand the impact of preprocessing on the effectiveness of AI models in detecting cyber threats, it is crucial to analyze how improvements in data quality have contributed to more accurate and efficient detection. Data preprocessing included removing duplicate records, correcting missing values and anomalies, and standardizing formats. These actions significantly improved the quality of data available for training AI models.

Cleaning and normalizing the data resulted in a more coherent and representative data set, which allowed the AI models to learn from a more accurate and less noisy set of information. This improvement in data quality translated into greater precision in threat detection, as evidenced by the improved detection rate of 95% after the implementation of AI technologies. Additionally, the incident response time was dramatically reduced from minutes to seconds, demonstrating the ability of AI systems to react quickly to identified threats. The analysis of the results highlights the importance of data preprocessing in optimizing AI models for cybersecurity. Improved precision and rapid incident response indicate AI models’ ability to operate effectively in dynamic and complex cybersecurity environments.
4.2. Evaluation of AI Performance in Threat Detection

The ability of AI models to identify cyber threats was evaluated, revealing a significant advance over conventional security methods. The AI algorithms applied to the collected and preprocessed data set demonstrated a high threat detection rate, notably outperforming traditional security solutions in efficiency and effectiveness. Through the analysis, the precision and sensitivity of these models were quantified, with the results highlighting a substantial improvement in accurately identifying malicious activities. The comparative evaluation between various AI algorithms highlighted the superiority of specific techniques that managed to successfully detect a wide range of cyber threats, from brute force attacks to sophisticated intrusions, thus evidencing the potential of AI in reinforcing cybersecurity.

4.2.1. Threat Detection

The obtained results present the improved efficiency of the AI models evidenced by a significantly higher threat detection rate. The analysis began with the implementation of the AI algorithms in a controlled environment, using the same cyber threat data set for the two systems: the AI model and the traditional security system. Special attention was paid to training and tuning the AI models, ensuring they were well-calibrated to recognize malicious behavior patterns from the collected and preprocessed network traffic data, event logs, and system logs.

To measure the threat detection rate, the number of actual threats detected by each system was counted against the total threats in the test data set. The results showed that the AI models achieved a detection rate of 94.7%, detecting 947 of the 1000 threats, while traditional systems managed to identify 749 threats, resulting in a detection rate of 74.9%. As seen in Table 3.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>AI Model</th>
<th>Traditional System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threat Detection Rate (%)</td>
<td>94.7</td>
<td>74.9</td>
</tr>
<tr>
<td>Total Number of Threats Detected</td>
<td>947</td>
<td>749</td>
</tr>
<tr>
<td>Total Number of Threats Present</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>False Positives</td>
<td>53</td>
<td>104</td>
</tr>
<tr>
<td>False Negatives</td>
<td>53</td>
<td>251</td>
</tr>
</tbody>
</table>

A direct comparison of these figures demonstrates the advantage of AI in accurately identifying cyber threats. In addition, the number of false positives and false negatives generated by both systems was analyzed, and was lower in the AI models, indicating greater precision and reliability in threat detection. These results highlight the superiority of AI in effectively detecting malicious activity but also underline its potential to optimize cybersecurity operations, reducing the incidence of false positives and improving the ability to respond to real threats. The quantitative evidence from this study provides a solid basis to argue for adopting AI technologies in advanced cyber defense systems.

4.2.2. Precision and Recall Analysis

Through analysis, precision and sensitivity metrics were quantified, highlighting the superiority of AI systems in identifying and reacting to malicious activities. To arrive at these results, benchmark tests were applied using an extensive data set reflecting a variety of cyber threats. The models selected for the study, including random forest, SVM, and Deep Neural Networks (DNNs), were trained and tuned to maximize their ability to recognize malicious patterns within network traffic and other threat indicators. Each model underwent a series of tests to measure its precision, defined as the proportion of correct
threat identifications among all identifications made, and its sensitivity, which measures the model’s ability to detect all real threats present in the set of data.

The results showed that DNNs improve the performance with 93% precision and 90% sensitivity, followed by random forest with 92% precision and 88% sensitivity. SVMs were also effective, although to a lesser extent, with an accuracy of 89% and a sensitivity of 85%. In contrast, traditional security systems had significantly lower rates, with a precision of 75% and a sensitivity of 70%. This analysis was based on applying advanced statistical methodologies and empirical tests, guaranteeing the reliability and validity of the results obtained. The AI algorithms were evaluated in terms of their ability to detect threats, minimize false positives, and respond effectively to security incidents, thus establishing a compelling argument for their adoption within the modern cybersecurity framework.

4.2.3. Comparison of Algorithms

Various AI algorithms were implemented to perform the comparative analysis, including random forests, SVM, DNN, CNN, and decision trees. Each model was trained using a representative cyber threat data set, and its parameters were optimized to maximize precision and sensitivity in threat detection.

In the analysis, we differentiated between neural networks in general and specific types, such as CNNs. Given that CNNs are a subtype of neural networks, these may seem to overlap. However, this distinction highlights CNNs’ specialized capabilities in processing visual and spatial data, differentiating them from other neural network architectures. CNNs are especially effective in analyzing visual and spatial data due to their unique structure, which mimic the mechanism of visual perception in living beings. This specialization makes them ideal for detecting image patterns or network traffic with visual characteristics. On the other hand, DNNs refer to a broader spectrum of neural networks capable of handling various data analysis tasks, including those that do not necessarily have a clear spatial or visual structure.

In Table 4, we differentiate the types of neural networks to show how each performs in detecting cyber threats based on metrics such as precision, recall, and F1 Score. This differentiation is intended to demonstrate that, although CNNs are a form of neural network, their separate evaluation is justified by their specialization and unique ability to deal with certain types of cyber threat data. When evaluating AI algorithms, it is essential to consider the explainability of each method to ensure its applicability in cybersecurity environments. Decision trees offer a transparent structure that allows users to understand decisions based on specific model characteristics. On the other hand, DNNs and CNNs provide high rates of accuracy and sensitivity but often act as “black boxes,” where the internal decision processes are not easily interpretable. This can complicate understandings of the specific nature of the detected threats. Support vector machines (SVMs), especially with non-linear kernels, also present similar challenges in terms of transparency. Recognizing and addressing these aspects of explainability is critical to developing threat detection systems that are effective, reliable, and understandable to security operators.

Table 4. Performance evaluation of AI algorithms in cyber security.

<table>
<thead>
<tr>
<th>AI Algorithm</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1 Score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>92</td>
<td>88</td>
<td>90</td>
</tr>
<tr>
<td>SVM</td>
<td>89</td>
<td>85</td>
<td>87</td>
</tr>
<tr>
<td>Deep Neural Network</td>
<td>93</td>
<td>90</td>
<td>91.5</td>
</tr>
<tr>
<td>CNN</td>
<td>91</td>
<td>87</td>
<td>89</td>
</tr>
<tr>
<td>Decision Trees</td>
<td>88</td>
<td>84</td>
<td>86</td>
</tr>
<tr>
<td>Traditional System</td>
<td>75</td>
<td>70</td>
<td>72</td>
</tr>
</tbody>
</table>

The results indicate that DNNs obtained the best results, with higher precision and sensitivity, compared to random forests and CNNs. Although effective, the decision trees
and SVMs showed slightly inferior performance. Data collection and analysis allowed for an objective and quantitative evaluation of each algorithm’s performance.

This evaluation process involved training, testing the models, and analyzing the test results to understand how each algorithm processes and responds to cyber threats. Precision, sensitivity, and F1 Score were calculated from actual threat detection on the test data set, providing a standardized measure of model effectiveness.

Furthermore, the analysis revealed the superiority of DNNs in threat detection, marking the importance of model selection in implementing AI-based cyber defense systems. These findings provide a solid foundation for future research and development in cybersecurity, guiding the selection of technologies toward those most effective in preventing and mitigating cyber-attacks.

4.3. Effectiveness in Incident Response

AI models have demonstrated a solid ability to react to threats in real time, significantly reducing response times compared to traditional security methods. This improvement in incident response is evidenced by the ability of AI-based systems to automatically execute mitigation actions, minimizing the potential impact of detected threats and improving the overall resilience of the cybersecurity system.

Integrating AI into Security Operations Center (SOC) operations has transformed the dynamics of incident response, where every second counts [41]. Although it seems marginal, the reduction in response time, which can vary between 3 and 6 s compared to traditional methods, is vital in critical security. In situations where threats evolve rapidly, such as ransomware attacks or network intrusions, these seconds can be crucial to preventing an attack from spreading and reducing the resulting damage.

The effectiveness of AI in providing faster responses is based on its ability to analyze large volumes of data efficiently and make automated decisions based on complex and changing patterns. This capability accelerates the detection and response process and enables continuous adaptation to new threats, which is essential to maintaining security against increasingly sophisticated adversaries.

4.3.1. Effectiveness in Incident Response

With the implementation of AI systems, agility in responding to cybersecurity incidents has become innovative. Additionally, integrating AI systems into SOC operations has improved agility in responding to cybersecurity incidents. Figure 2 illustrates a consistent trend: AI-based systems respond faster than traditional methods and maintain improvement over time, optimizing their protocols and reducing latency in response to threats [42,43]. The analysis highlights that the few seconds gained in response time are critical for SOC operations. For example, a faster response can mean the difference between a localized infection and a large-scale spread in a ransomware attack. This time frame allows for decisive actions such as disconnecting compromised devices from the network, mitigating the impact, and preventing potential data breaches.

This analysis focused on measuring response times after detecting a threat. Data were collected bimonthly, reflecting the intervals of ‘Month 1–2’, ‘Month 3–4’, and ‘Month 5–6’, and then subjected to statistical analysis. Data for this evaluation come from timed recorded responses to simulated incidents, measured in seconds. Each two months, data were collected, corresponding to the systems’ reaction to a series of verified threats. The times from initial detection to execution of the appropriate response were recorded.

An improvement was observed in the speed of response of the AI systems, with a trend towards decreasing median response times with each successive measurement, implying the continuous learning and adaptation of the AI system; in contrast, the response times of traditional security systems exhibited more excellent dispersion and a slight tendency to increase in the median over time, which could indicate less adaptability to emerging threats.
Each box on the graph represents the distribution of response times for each system over the two-month intervals. The line inside each box denotes the distribution median, while the boxes’ ends define the first and third quartiles, thus providing a view of the interquartile range that encompasses the middle half of the data. The whiskers extend to the minimum and maximum values within a defined range unless outliers are represented as individual points outside the whiskers.

![Graph showing response speed comparison between AI and traditional systems over time.](image)

**Figure 2.** Comparison of response speed between AI and traditional systems over time.

The graph’s results directly compare the two types of systems and their response capabilities over time. The AI systems’ improved responsiveness is evident and suggests continuous adaptability and optimization in threat management. These findings highlight the importance of incorporating AI into cybersecurity protocols and promise significant improvements in the overall effectiveness of incident response strategies.

### 4.3.2. Response Automation

Organizations can significantly reduce the time between detecting and mitigating threats by automating incident responses. AI, in this process, allows for not only the application of simple rules but also advanced techniques that provide significant value by interpreting and responding to threats in real time. Table 5 presents various types of automatic reactions implemented by AI systems and the frequency of their activation during the study period, illustrating the diversification and agility that AI brings to response capabilities.

**Table 5.** Analysis of automation in response to cyber security incidents.

<table>
<thead>
<tr>
<th>Automated Response Type</th>
<th>Description</th>
<th>Activation Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Isolation</td>
<td>Disconnecting infected devices from the network to contain the spread of a threat</td>
<td>120 times</td>
</tr>
<tr>
<td>Account Deactivation</td>
<td>Temporary suspension of user accounts with abnormal behavior</td>
<td>75 times</td>
</tr>
<tr>
<td>Blocking Network Traffic</td>
<td>Immediate interruption of traffic flows identified as malicious</td>
<td>200 times</td>
</tr>
</tbody>
</table>
The data analysis reveals that the implementation of AI has increased the response capacity, adapting to the type and severity of the perceived threat. The AI systems recorded a series of automatic actions during the study period. These automated responses, from the immediate isolation of compromised devices to the application of security patches without human intervention, demonstrate remarkable efficiency in preventing the infiltration and spread of cyber threats.

The trigger frequency of these responses provides quantitative insight into AI’s contribution to real-time security. For example, system isolation was activated 120 times, reflecting a security policy prioritizing rapid containment of the threat. Equally important, the automatic deployment of security patches was carried out 150 times, highlighting the ability of AI systems to strengthen defenses by fixing known vulnerabilities without delay.

The 300 automated alerts sent to security teams highlight the essential collaboration between AI and security professionals, where technology acts as a force multiplier, allowing human teams to focus on threats that require specialized expertise and judgment.

The automation of responses evidenced by data marks the transformation that AI drives in cybersecurity: an evolution from reactive protocols to proactive and dynamic strategies, maximizing the effectiveness of incident response and strengthening the overall defensive posture.

4.4. Reduction in False Positives and Negatives

Figure 3 evaluates the effectiveness of AI in reducing false positives and negatives and the improvement this represents compared to traditional security systems. The line graph illustrates the variability and performance of the AI system over six months, highlighting the system’s progressive adaptation and learning. The variable line highlights the system’s ability to adjust genuine threat detection over time.

The second bar chart graph compares false positive and negative counts between traditional and AI-based systems, providing a direct visualization of AI’s effectiveness compared to conventional methods. The comparison reveals differences in the incidence of errors, underscoring that AI can improve the speed and accuracy of incident response and decrease the number of incorrect alerts that could divert valuable resources.
Figure 3. Dynamic analysis of error rates in AI-driven cybersecurity systems.

The analysis was based on a set of simulated security incidents, where whether each alert was a true positive, a false positive, a true negative, or a false negative was recorded. These incidents were processed by AI systems and, for comparison purposes, by traditional systems. Performance was evaluated monthly, allowing for the evolution of the AI system and its learning from interactions with new and diverse threats to be tracked. This detailed analysis supports the integration of AI into cyber defense systems. It highlights its critical role in continuously improving security operations. It shows how this technology can reduce errors and build a more reliable and robust system for detecting and managing cyber threats.

4.5. Impact on Global Cyber Security

A detailed analysis of operational metrics reveals an encouraging picture of AI’s role in redefining global cybersecurity. The statistics show a substantial improvement in the ability to detect and respond quickly to threats and notable efficiency in reducing false alerts, fundamental elements for protecting critical infrastructures. Exploring AI’s impact on cybersecurity opens the dialogue on how these advanced solutions are making a difference in the complex fabric of cyber defense globally.

4.5.1. General Security Improvements

Integrating AI into cybersecurity strategies has led to significant improvements in the protection of critical infrastructure globally. Adopting these advanced technologies has redefined the ability to respond and resist complex incidents, evidenced by a notable decrease in the frequency and impact of cyber-attacks. The analyzed data, reflected in Figure 4, show a reduction trend in the number of incidents and their severity since the implementation of AI systems. The first graph reveals a decrease in the frequency of incidents month on month, underscoring the proactive role of AI in preventing attacks before they happen or in quickly neutralizing them. The red line, which represents the frequency of incidents before AI implementation, shows a significantly higher trend than the green line, which illustrates the frequency of incidents after AI implementation. This comparison shows how AI contributes to a notable decrease in security incidents. AI’s predictive capability and continuous surveillance allow one to proactively identify and mitigate vulnerabilities before they are exploited.
Figure 4. Trends in cybersecurity incident frequency and impact with AI integration.

The second graph complements this view by demonstrating a decrease in the impact of incidents that occur, being indicative of more effective risk management and faster recovery. The relationship between the blue and yellow lines exposes the difference in the effects of security incidents before and after integrating AI. The blue line, which indicates a more significant impact of incidents before AI, contrasts significantly with the yellow line, which shows how the effect is reduced after the implementation of AI. This contrast suggests that AI is detecting threats more effectively and implementing more effective responses to contain and neutralize the potential damage from such incidents. The reduction in the impact of incidents can be attributed to AI’s automated and adaptive response, which applies countermeasures in real time and adapts security defenses more dynamically than traditional systems.

The impact score is defined by a comprehensive approach that considers the severity of the incident, the extent of potential damage, and the speed of response. This comprehensive metric seeks to reflect the seriousness and consequences of cybersecurity incidents before and after implementing AI systems.

The severity of each incident is initially rated on a scale of 1 to 5, with 1 representing minor incidents with marginal impacts and 5 being reserved for critical security breaches with the potential for extensive damage to infrastructure and data integrity. This assessment is based on standardized criteria to ensure consistency over time and between security analysts. The extent of potential damage is estimated by considering the number of compromised systems and their importance to the infrastructure’s operation. A rating is assigned on a scale of 1 to 10, with the highest value indicating a high risk of affecting the organization’s critical systems. The speed of response is measured from the moment the threat is detected until effective corrective action is implemented. More agile and early responses receive higher scores, reflecting the premise that prompt intervention can significantly reduce the magnitude of the impact.

The consolidated impact score is obtained by a weighted average of these factors, creating a quantitative and comparable measure of the impact associated with each incident. This calculation allows for an objective and standardized assessment of the real impact of security incidents. It demonstrates how integrating AI into cybersecurity systems changes the risk landscape over time.
A detailed and systematic monitoring of cybersecurity incidents was carried out over 12 months before and after implementing AI-based solutions to reach these results. Data on the frequency and impact of incidents were meticulously collected through security reports and cyber defense system logs. AI algorithms designed for early detection and rapid response were applied, resulting in a lower frequency of incidents and less impact when they occurred.

According to the ‘Cybersecurity and Infrastructure Security Agency (CISA)’ report published by America’s Cyber Defense Agency, the implementation of AI has been shown to improve the security of critical infrastructure networks and strengthen resilience to complex attacks. This report highlights how AI solutions have enabled organizations to adapt more effectively to emerging threats in the cybersecurity landscape [44]. Additional case studies illustrating the effectiveness of AI in real-world situations will provide deeper insight into its positive impact, showing how the technology has not only detected and mitigated specific incidents but has transformed cybersecurity into a solid and dynamic strength.

4.5.2. Comparison with Other Solutions

The impact of this implementation goes beyond a simple improvement in threat detection rates, encompassing a comprehensive transformation in the way critical infrastructure is defended in the digital realm. When comparing our AI solution to alternative systems, it is evident that there are notable improvements on several fronts. With a threat detection rate of 95%, our solution outperforms the rules-based and hybrid systems, which reach 85% and 90%, respectively. This improvement in the detection rate is not only a reflection of AI’s ability to identify known threats but also its ability to learn and adapt to new tactics employed by cybercriminals.

Average response time is a critical factor in cybersecurity incident management. Our solution responds in an average of 3 s, an impressively faster speed compared to 9 s for the rules-based system and 6 s for the hybrid system. This rapid response is crucial to mitigating the impact of attacks, enabling near-instant defensive action that could be the difference between a minor breach and a devastating security breach.

Regarding the false positive rate, which can overwhelm security teams and divert essential resources, our solution shows a rate of 4%. In contrast, the rule-based system shows a rate of 12%, and the hybrid system is 9%. Although the operational cost of our AI solution is higher, this additional expense is more than justified by the reduction in operational interruptions and downtime, which can have even more significant financial consequences.

Similarly, the advanced AI system significantly improves the false negative rate, representing the threats that evade detection and are potentially the most dangerous. While our solution shows 5%, the rule-based system has a rate of 15%, and the hybrid system 10%, indicating a clear advantage in terms of detection reliability.

This analysis makes the case for investing in advanced AI in cybersecurity. Despite the higher initial operating cost, the return on investment is realized through significant improvements in incident prevention and response and reductions in the disruption and costs associated with false positives and negatives. The complete results are presented in Table 6, which strengthens the argument that AI solutions are not just a complement, but a necessary paradigm shift for cybersecurity in our current digital age.

The effectiveness of our AI solution is based on an analysis of data collected over an observation period, where cyber threats on critical infrastructure were continuously monitored. These data, obtained from security event logs and intrusion detection systems, included a variety of threats, from brute force attacks and ransomware to sophisticated phishing techniques and APTs. During the study period, AI models were refined by incorporating new data, allowing for iterative threat detection and response improvement. This refinement process was based on advanced machine learning techniques, which adapted the models to capture emerging trends and evolving attack tactics.
Lower operating costs are derived from a substantial reduction in response times and the decrease in false positive and negative rates, representing significant savings by avoiding operational interruptions and reducing downtime. These financial benefits and improved detection reliability demonstrate a clear return on investment and justify the higher initial investment in the AI solution.

Table 6. Comparison of efficiency and costs in cyber security systems.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Our AI Solution</th>
<th>Rules Based System</th>
<th>Hybrid System (Rules and Essential ML)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threat Detection Rate (%)</td>
<td>95</td>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td>Average Response Time (sec)</td>
<td>3</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>False Positive Rate (%)</td>
<td>4</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>False Negative Rate (%)</td>
<td>5</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Estimated Operating Cost (USD)</td>
<td>50,000</td>
<td>40,000</td>
<td>30,000</td>
</tr>
</tbody>
</table>

Figure 5 presents the hyperparameter optimization. Key elements such as the number of trees in the random forests and the maximum depth of these trees were adjusted to improve the effectiveness of threat detection. This tuning process was performed using cross-validation techniques, ensuring that the models fit the training data and generalize well to new, unseen data.

- By increasing the number of trees in the random forests, the threat detection rate improved, going from 85% in rule-based systems to 95% in optimized AI systems. This indicates that a more significant number of trees in the model can increase robustness and the ability to capture complex variations in hazard data.
- Adjusting the maximum depth allowed for a balance between capturing detail in the data and preventing overfitting, improving model precision without sacrificing generalization ability.

![Figure 5](image)

In addition, periodic tests were conducted to continually evaluate and improve the effectiveness of AI systems, playing a crucial role in adapting these systems to the changing dynamics of cyber threats. These tests helped identify areas for improvement and make necessary adjustments to the AI models to maintain their relevance and effectiveness, for example:
• Regular system evaluations demonstrated that AI could adapt to new malicious behavior, as reflected in the progressively decreasing incident response time. Initially, the average response time was 6 s. Then, after regular testing and adjustments, this time was reduced to 3 s, highlighting the agility of the AI system in quickly responding to threats.

• The continuous review of AI performance against simulated threats allowed models to be adjusted to improve detection and response, ensuring that the system reacted efficiently to known threats and could anticipate and neutralize new attack tactics.

These results highlight the effectiveness of AI systems in threat detection and response times and show their ability to adapt and continually improve through an iterative testing and optimization process.

4.6. Study Cases

In the first case, we addressed the growing ransomware problem in energy control systems. Ransomware infiltration could cripple a power grid, causing not only significant economic losses but also impacting public safety. We implemented an AI-based system that uses deep learning techniques to monitor network operations in real time, identifying patterns of activity that suggest ransomware attack attempts. This AI system demonstrated an exceptional ability to detect and neutralize ransomware threats in the early stages, achieving a detection rate of 98%. This high level of precision in detection allowed for proactive responses to be implemented, minimizing the operational impact and maintaining the continuity of the energy service.

The second case focused on ICSs, which are essential for energy facilities’ safe and efficient operation. Attacks on these systems can have devastating consequences. We used AI algorithms to analyze behavioral patterns and operational data, distinguishing between normal operations and suspicious or malicious activities. The results were remarkable, with a 90% reduction in undetected security incidents compared to traditional methods. Additionally, the AI system optimized incident response times, moving from a process that could take hours to real-time solutions, significantly improving operational resilience.

The third case examined how AI has been applied for predictive analytics in the power distribution network, focusing on predicting and mitigating disruptions before they occur. Through the collection and analysis of large volumes of operational data, AI models were able to identify patterns that indicated potential failures or cyber-attacks—implementing this predictive analysis system significantly improved outage prevention, with a failure prediction precision of 95%. This proactive approach allows energy companies to take corrective action before problems can escalate, ensuring more stable and secure power delivery.

4.6.1. Quantitative Analysis

Improved Ransomware Detection (Case Study 1): We used deep learning models trained with historical data from known attacks and normal network behaviors. The ransomware detection rate increased to 98% after implementing these solutions, compared to 75% with older methods. This was achieved by optimizing algorithms to recognize specific ransomware patterns and adjusting characteristics such as the learning rate and the number of layers in the neural networks.

Reduction in Safety Incidents in ICSs (Case Study 2): We implemented anomaly-detection systems that constantly analyze the operational data of the ICSs to identify deviations from the norm. The effectiveness of these systems was evaluated in controlled tests, showing a 90% reduction in undetected incidents. This is attributed to the ability of AI systems to learn from past events and adapt to new attack tactics.

Efficiency in Predictive Analysis (Case Study 3): For predictive analysis in the power distribution network, algorithms that process large volumes of operational data were
implemented to identify predictive patterns of failures or attacks. Failure prediction precision reached 95%, enabling preventive actions that reduced outages by 70%. This was achieved by calibrating predictive models that analyze trends and correlations in the data.

4.6.2. Detailed Qualitative Analysis

Confidence in Power System Security: Implementing AI technologies has reinforced confidence in the security of power systems. The ability to proactively detect and respond to threats allows operators to focus more on efficiency and innovation. A qualitative analysis of the attitudes and perceptions of operational staff was carried out, showing a significant improvement in confidence in the system’s security after the implementation of AI.

Improved Operational Stability: The ability to respond quickly to incidents has reduced the time and resources required to recover from attacks or failures. This was evaluated by comparing downtime and associated costs before and after the implementation of AI solutions, showing a notable improvement in operational stability.

Efficient Resource Management: Predictive analysis has allowed for more efficient resource management. By anticipating potential problems, companies can plan and allocate resources more effectively. An analysis of security investment showed that implementing AI led to a more strategic allocation of funds, improving overall profitability.

5. Discussion

The results obtained, which are supported by specific case studies and quantitative and qualitative analyses, illustrate the effectiveness and potential impact of AI solutions in improving energy resilience and security. The methods detail how AI technologies were selected and configured to address specific challenges in energy infrastructure. The implementation methodology was designed considering the peculiarities of the sector, such as precision in detection and response capacity. As demonstrated in our case studies, adapting these systems to the specific needs of critical energy infrastructure reflects a meticulous approach that seeks to apply advanced technology and ensures its relevance and effectiveness in the operational context.

The case studies presented in the results provide a practical view of AI’s application. For example, proactive ransomware detection and improved ICS security demonstrate how AI can identify and mitigate threats before they become significant incidents [45]. These cases validate the technical capability of AI systems and highlight their operational value, delivering tangible improvements in the safety and efficiency of energy operations.

The quantitative analysis revealed significant improvements in security metrics, such as increased threat detection rates and reduced incident response time. These results not only corroborate the effectiveness of AI solutions but also establish a direct link with the applied methodologies, reinforcing the validity of our approach. For example, the improvement in the threat detection rate from 75% to 98% in the context of ransomware attacks evidences an improved ability to protect critical power assets.

From a qualitative perspective, implementing AI has substantially impacted cybersecurity management in energy infrastructure. The ability to anticipate and respond to complex threats has led to greater confidence in the stability and security of the energy system. Additionally, automation and rapid incident response facilitated by AI systems have enabled more efficient resource management, thereby optimizing operations and resource allocation.

Data reliability and security discussions highlight a critical consideration in AI implementation. Data integrity is vital to the effective functioning of AI systems, so protection measures and data security protocols become essential aspects of AI infrastructure [46]. This is reflected in our attention to data quality and security during the preprocessing phase, which ensures that AI systems operate at maximum efficiency and reliability.

Works such as Balaji and Narayanan [30] support the effectiveness of AI solutions in improving energy infrastructure security. They found that using advanced machine
learning techniques significantly improved cyber-attack detection in the energy sector, with detection rates increasing up to 90% in test scenarios. This finding is consistent with our results, where the implementation of AI models increased the ransomware detection rate to 98%, highlighting the effectiveness of these systems in natural operating environments.

In the discussion of the security of ICSs, Ameri et al.’s [9] work provides a valuable analysis of how AI-based defense strategies can mitigate risks in critical infrastructures. It matches our findings, showing a 90% reduction in undetected security incidents at ICSs, underscoring the ability of AI systems to identify and neutralize threats efficiently.

The discussion should also address data management and reliability, as discussed in research by Kumari et al. [47], emphasizing the need to ensure data security and integrity in AI solutions. This aspect is crucial in our study, where preprocessing and data protection were essential to achieving optimal performance of AI systems, demonstrating the importance of data quality in the effectiveness of threat detection and response.

The qualitative impact of AI on the operability and management of energy infrastructure is reflected in our discussion of improving trust in system security, which is an observation supported by the research of Al-Muntaser et al. [48]. In their study, implementing AI solutions led to greater operational efficiency and better strategic decision making, which aligns with our observation that AI facilitated more efficient management and strategic resource allocation in the energy sector.

The work of Alzahrani and Aldhyani [5] discusses the need for continuous adaptation and development in AI technologies. They argue that the changing landscape of cyber threats demands the constant evolution of AI solutions. This point resonates with the conclusion of our study, which highlights the importance of continuous innovation and adaptive learning in AI systems to remain effective in the face of emerging threats in critical energy infrastructure.

The results of this study suggest a vast potential for future explorations and innovative developments regarding the future of AI in energy infrastructure. The continued evolution of cyber threats requires a dynamic and adaptive approach, where AI needs to keep up with current trends and anticipate and prepare for future challenges. This involves a continued commitment to research and development and effective collaboration between the technology, energy, and cybersecurity sectors.

6. Conclusions

This study has shown that integrating AI into the cybersecurity of critical energy infrastructure offers significant improvements in detecting and responding to cyber threats. Advanced AI technologies, such as deep learning and predictive analytics, have achieved a 98% threat detection rate and a reduction in incident response time of more than 70%. These results reflect AI’s ability to process and analyze large volumes of data effectively and underline its potential to act proactively against cyber threats, thereby ensuring the resilience and stability of critical energy infrastructure.

The practical application of AI in the energy sector, illustrated through specific case studies, has enabled a deeper understanding of the operational dynamics and specific threats facing this sector. Customized AI systems have been proven capable of adapting to the complexities of the energy environment, providing more accurate and efficient security solutions. This approach improves cybersecurity and facilitates more effective resource management, improving operational efficiency and strategic decision making in energy infrastructure.

The findings of this study underscore the critical importance of adopting advanced technological approaches in cybersecurity within key energy infrastructures. Integrating AI solutions improves the ability to respond to cyber incidents. It establishes a new paradigm in security management, where prevention, detection, and rapid response become integrated components of daily operations.
It is evident that the field of AI in cybersecurity within energy infrastructure will continue to evolve, driven by technological advances and the changing dynamics of cyber threats. Future research should explore more advanced AI algorithms and adaptive learning systems that can anticipate and evolve in response to cyber attackers’ changing strategies. Additionally, integrating AI with other emerging technologies, such as blockchain for data security and quantum computing for data analytics, could offer new avenues to strengthen energy infrastructure security.

Addressing the challenges associated with ethics and privacy when implementing AI in cybersecurity is crucial. Future research should consider how AI systems can be designed and regulated to protect the privacy and rights of individuals and organizations while maintaining a robust security posture. Interdisciplinary collaborations between AI, cybersecurity, law, and ethics experts will be essential to developing effective and ethically responsible solutions.

**Author Contributions:** Conceptualization, W.V.-C.; methodology, W.G.-N.; software, J.G.; validation, W.V.-C. and J.G.; formal analysis, J.G.; investigation, W.G.-N.; data curation, W.G.-N.; writing—original draft preparation, J.G.; writing—review and editing, W.V.-C.; visualization, W.V.-C.; supervision, W.V.-C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data supporting the results of this study are part of a confidential and protected data set obtained in collaboration with private entities and under confidentiality agreements. Given the privacy restrictions and sensitive nature of the data, they are not publicly available in their raw form. However, we are committed to transparency and scientific collaboration. Therefore, researchers interested in accessing the derived or aggregated data used for this analysis may contact the corresponding author. Access will be granted following a legal process, including signing a non-disclosure agreement and compliance with applicable privacy and ethical guidelines. This measure ensures that data use aligns with ethical principles and confidentiality restrictions while facilitating scientific collaboration and scrutiny of the research results presented in this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.