Power System Technology
ISSN:1000-3673

Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Agentic AI + Prompt Engineering: Next-Gen Machine Learning for IoT Intrusion Detection

Chandrani Mukherjee

Sr. Enterprise Architect (AI), NJ, USA

Former Computer Science Department Liverpool John Moores University, UK

moni121189@yahoo.co.in

ABSTRACT

The growing use of Internet of Things (IoT) devices is expanding the attack surface. This compels the need for smarter and more dynamic intrusion detection systems (IDS). Standard machine learning methods face challenges related to scalability, location sensitivity, and adaptability. In this paper, we present a next-generation platform that combines Agentic Artificial Intelligence (AI) with prompt engineering to improve IoT intrusion detection. Agentic AI supports autonomous learning, decision-making, and adaptation. Prompt engineering dynamically guides model behavior using contextual instructions. We compare our approach with state-of-the-art machine learning and deep learning methods on benchmark IoT security datasets. Findings indicate improved accuracy, enhanced resistance to various attack models, and reduced false positives. These results suggest that Agentic AI and prompt engineering may lead to the creation of interpretable, scalable, and resilient IDS for real-world IoT applications. We discuss practical implications for smart cities, health, and industrial IoT implementations.

Keywords: Agentic AI, Prompt Engineering, IoT Security, Intrusion Detection Systems, Machine Learning

1. INTRODUCTION

The fast development of the Internet of Things (IoT) has revolutionized contemporary society. It integrates billions of devices in critical areas such as healthcare, transportation, industrial automation, and smart cities. This connectivity has improved efficiency and permitted innovation. However, it has also introduced vulnerabilities previously unseen. The heterogeneity of IoT systems, limited computational resources, and weak security measures provide attackers with an ideal environment. Large-scale IoT attacks, such as Mirai botnets, distributed denial-of-service (DDoS) attacks, and data exfiltration campaigns, highlight the need to develop IoT ecosystem-specific intrusion detection systems (IDSs) (Hassija et al., 2019, IEEE Access).

Conventional intrusion detection methods mainly use signature-based detection. These approaches have failed to address new threats. Signature techniques rely on knowledge of prior attack patterns, making them ineffective against zero-day exploits and dynamic attack tactics

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

(Khan and Salah, 2018, Future Generation Computer Systems). Machine learning (ML) and deep learning (DL) methods have emerged to address these constraints in IoT security research. ML-driven IDS models can detect patterns in network traffic. This enables them to identify anomalies beyond predefined signatures (Al-Garadi et al., 2020, IEEE Communications Surveys and Tutorials). DL models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), perform better on larger-scale datasets (Yi et al., 2023, Journal of Network and Computer Applications). Despite these advances, IDS models still face issues of scalability, interpretability, and adaptability in real IoT environments.

A major problem with current IDS models is the high false positive rate (FPR). This reduces usability, as the system gets flooded with meaningless alerts. Cyberattacks are becoming more sophisticated. IDS models must identify known threats and, more importantly, learn and adapt to new attack vectors. Research stresses the need to go beyond static ML pipelines. IDS systems should reason contextually and improve over time (Aljabri et al., 2021, Sensors). The combination of Agentic Artificial Intelligence (Agentic AI) and prompt engineering may offer a promising shift in approach.

AI systems that possess autonomy, proactivity, and self-directed learning capabilities are referred to as agentic AI. In contrast to classical ML systems, where humans must periodically retrain the system, Agentic AI systems are autoregulating and self-directed. These systems are expected to engage in reasoning and act independently to optimize themselves in dynamic environments (Bhardwaj et al., 2023, Computers and Security). Such agentic conduct is particularly crucial in IoT security, where threats can emerge rapidly and manual responses may lag behind those of attackers. Current studies have also shown how agentic methods advance digital forensics and network surveillance by improving flexibility and context recognition (Combating Digital Media Piracy With Agentic AI, 2025, International Journal of Environmental Sciences).

Complementary to Agentic AI is the newly developed discipline of prompt engineering. This field has become a critical component in using generative AI models and large language models (LLMs). Prompt engineering is the development of context-dependent instructions to direct model outputs. It also enhances accuracy, interpretability, and alignment to tasks (Heston and Khun, 2023, International Medical Education). Originally a concept in natural language processing, prompt engineering is now seen as a general digital competence. It allows AI models to act flexibly when circumstances change (Korzynski et al., 2023, Entrepreneurial Business and Economics Review). In the context of IoT security, prompt engineering can dynamically train IDS models using various attack types and changing network traffic patterns. This approach refines anomaly detection without complete retraining.

There are various benefits to combining Agentic AI with timely engineering. First, it improves adaptability, since IDS systems can orient their responses to various IoT environments. Second, it minimizes the need for exhaustive static feature engineering or retraining, thus boosting computational efficiency. Third, integration leads to interpretability. Prompts can serve as

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

human-readable instructions that clarify why the system flagged certain traffic as anomalous. Recent studies in entrepreneurship (Short and Short, 2023, Journal of Business Venturing Insights) and healthcare (Mesko, 2023, Journal of Medical Internet Research) suggest that prompt engineering is a cross-disciplinary field. This supports its use in the IoT intrusion detection domain.

Although these positives have been noted, there is a lack of research on the overlap between Agentic AI, prompt engineering, and IoT intrusion detection. Available IoT IDS systems are often too resource-demanding for limited IoT devices or cannot be modified to address emerging attacks (Mohamad Noor and Hassan, 2019, Computer networks). A survey of machine learning techniques in IoT security shows active investigation into ML and DL models. However, very few techniques focus on agentic autonomy or dynamic prompting as primary design concepts (Hussain et al., 2020, IEEE Communications Surveys and Tutorials). Educating and national preparedness are core elements of cybersecurity plans (AlDaajeh et al., 2022, Computers and Security; Neri et al., 2024, Information and Computer Security). Still, few studies have examined the practical applications of agentic prompt-driven systems to IDS.

This paper presents an IDS framework of the next generation. It will combine Agentic AI with prompt engineering to solve IoT security. More specifically, this research makes the following contributions:

- This work theorizes a hybrid architecture. It integrates the autonomy of Agentic AI with the contextual flexibility of prompt engineering in IDS.
- By experimenting with benchmark IoT security datasets, the framework's performance is compared with classical ML and DL algorithms.
- A robustness analysis examines a range of attack conditions, including botnets, DDoS, and data manipulation attacks.
- The practical implications include a discussion of deployment strategies. These span domains, including smart cities, healthcare IoT, and industrial IoT.

This work addresses issues of scalability, interpretability, and adaptability. These qualities make it relevant to both academic research and practice. The work aligns with the demand for smarter, more context-aware, and resilient cybersecurity solutions that adapt to evolving adversarial tactics (Schiller et al., 2022, Computer Science Review; Taherdoost, 2022, Electronics).

2. LITERATURE REVIEW

2.1 IoT Intrusion Detection Environment.

IoT has expanded rapidly, connecting billions of devices in many fields, such as consumer electronics and industrial control systems. However, this rapid growth has led to a range of

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

cyber threats to IoT systems. Surveys report IoT devices are especially susceptible due to their small resources, heterogeneity, and lack of security measures (Hassija et al., 2019, IEEE Access). Conventional firewalls and signature-based intrusion detection systems (IDS) are insufficient to address advanced attacks (Khan and Salah, 2018, Future Generation Computer Systems).

To mitigate these issues, machine learning (ML) and deep learning (DL) techniques have emerged as optimistic solutions. ML algorithms can analyze network traffic data to detect anomalies. This enables identification of both familiar and unknown attacks (Al-Garadi et al., 2020, IEEE Communications Surveys and Tutorials). Deep learning methods, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), take it a step further. They enhance detection accuracy by automatically learning hierarchical feature representations (Yi et al., 2023, Journal of Network and Computer Applications). However, these models may need large labeled datasets, significant computation, and parameter optimization to scale to resource-limited IoT systems (Mohamad Noor and Hassan, 2019, Computer Networks).

Many extensive reviews show the challenges in IoT security. Hussain et al. (2020, IEEE Communications Surveys and Tutorials) highlighted the issue of managing heterogeneous attack vectors. These include denial-of-service (DoS) attacks, botnets, and data manipulation in various IoT setups. Similarly, Sha et al. (2020) surveyed edge-computing approaches to IoT security in Digital Communications and Networks. They demonstrated how offloading computation to edge devices can alleviate resource constraints. However, many of these solutions remain reactive and inflexible against new, emerging attacks.

Ethical hacking and penetration testing of IoT systems are emerging research areas. Yaacoub et al. (2023) emphasized the importance of proactive adversarial testing to reveal vulnerabilities and exploits in the Internet of Things and Cyber-Physical Systems. While these methods enhance resilience, they need complementary intelligent detection mechanisms. Such mechanisms should be able to adapt to unknown threats. Thus, IoT intrusion detection has advanced with the use of ML and DL models. Still, a shift is needed towards systems that are more autonomous, flexible, and context-specific.

2.2 Agentic AI in Cybersecurity

The latest advancement in AI systems is the so-called agentic Artificial Intelligence (AI). It is intended to demonstrate tendencies towards autonomy, proactivity, and self-directed learning. In contrast, traditional ML systems require regular retraining and parameter modification. Agentic AI systems are more self-reliant and can respond to dynamic environments with minimal human intervention. This autonomy is especially valuable in cybersecurity because attack vectors develop rapidly and often outpace manual response systems (Bhardwaj and Dave, 2023, Computers and Security).

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Recent research has explored the potential applications of agentic methodology in digital forensics and network surveillance. For example, it has been demonstrated that incorporating agentic AI into forensic systems can enhance attack investigation. This is achieved by allowing adaptive and proactive detection methods (Combating Digital Media Piracy With Agentic AI, 2025, International Journal of Environmental Sciences). Such systems not only detect intrusion, but also determine its nature. They can provide explanations and take countermeasures independently.

The attractiveness of agentic AI lies in its alignment with the urgency of cybersecurity, which requires resilience. In most cases, cyberattacks use zero-day vulnerabilities. Static detection strategies are inadequate. The self-optimizing ability of agentic AI offers a way to build more resilient intrusion detection systems (AlDaajeh et al., 2022, Computers and Security). Additionally, research into organizational preparedness shows that the human decision-making process can be supplemented by the agentic system. This system fills the skill gap and reduces the number of people needed for manual control (Neri et al., 2024, Information and Computer Security).

Agentic AI has been successfully applied in various fields, yielding encouraging results. For instance, Bhardwaj and Dave (2023, Computers and Security) found that neural network-based structures, combined with agentic decision-making, improved the accuracy and interpretability of detection. Similarly, Cao et al. (2020, IEEE Access) noted that network attacks on cyber-physical systems require intelligent, self-adaptive mechanisms capable of real-time protection. These results demonstrate that Agentic AI makes a significant contribution to IoT intrusion detection. It offers flexibility and autonomy beyond what standard ML and DL models can provide.

2.3 Timely AI Systems Engineering.

Recently, the concept of prompt engineering has become a highly effective approach to steer AI systems, especially large language models (LLMs). It does this by using instructions in carefully designed prompts. Originally created in natural language processing, prompt engineering has since expanded to other areas. It serves as a means to improve model performance, interpretability, and task compatibility (Korzynski et al., 2023, Entrepreneurial Business and Economics Review).

In AI applications, prompt engineering acts as an intermediary between human intent and machine action. Prompt tuning enables users to obtain better context-specific models that produce more accurate output, without requiring retraining or modifying the underlying architectures. This adaptability lowers computing costs and boosts flexibility for new activities (Mesko, 2023, Journal of Medical Internet Research). Clear prompts serve as guidelines that enhance interpretability, allowing stakeholders to better understand the model's decisions.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

There are already successful examples of prompt engineering in various fields. Recently, Heston and Khun (2023, International Medical Education) demonstrated its use in medical education to enhance context-relevant content presentation. Short and Short (2023, Journal of Business Venturing Insights) investigated how prompt engineering facilitated the creation of persuasive AI-generated content for entrepreneurship, helping entrepreneurs frame their content effectively. These works demonstrate that prompt engineering is not only applicable to NLP but also has potential in the security context.

In IoT intrusion detection, prompt engineering may allow IDS to dynamically adjust behavior based on network conditions or attack types. For example, prompts can set detection priorities, adapt to abnormal traffic, or customize anomaly classification thresholds. This flexibility helps overcome a key shortcoming of ML-based IDSs—inflexibility in responding to unknown or changing threats. Combining timely engineering with Agentic AI may facilitate the development of intelligent, contextual, and responsive intrusion detection systems.

2.4. Limitations of the Current Research.

Although the power of IoT intrusion detection has improved, significant loopholes remain. First, most ML- and DL-based IDS methods require large, labelled datasets. These datasets are not accessible or feasible to gather in the reality of the IoT setting (Janiesch et al., 2021, Electronic Markets). This dependency creates inflexibility in response to zero-day attacks and changing adversarial tactics. Second, computational efficiency is still a barrier. While edge computing has been proposed as a partial solution (Sha et al., 2020, Digital Communications and Networks), not all IDS models are sufficiently lightweight for resource-constrained IoT devices.

Third, interpretability remains a problem. Black-box deep learning (DL) models—algorithms that learn patterns from data without providing insight into their internal decision processes—can be highly accurate but offer little visibility into why certain network traffic is labeled malicious. This lack of transparency reduces trust within the system administrator environment. It also makes compliance with regulatory requirements more difficult (Slapnicar et al., 2022, International Journal of Accounting Information Systems). Fourth, autonomy and context-awareness—meaning the model's ability to act independently and to consider surrounding circumstances—are yet to be integrated. Although agentic AI frameworks, which enable software agents to operate autonomously, are promising (Bhardwaj and Dave, 2023, Computers and Security), they are seldom combined with related approaches, such as prompt engineering —a technique for crafting effective instructions for AI systems. Such combinations may further increase adaptability and interpretability.

Lastly, there are very limited studies that have conducted a systematic assessment of the use of Agentic AI and prompt engineering in combination to address IoT security. Despite its importance as an emerging digital skill (Korzynski et al., 2023, Entrepreneurial Business and Economics Review), the possibility of the dynamical guidance of IDS by prompt engineering

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

has not been sufficiently investigated. Equally, although organizational and strategic research identifies the need to develop capabilities in cybersecurity (Fernandez De Arroyabe et al., 2023, Computers and Security; Taherdoost, 2022, Electronics), technical frameworks to operationalize such information are limited.

To overcome these constraints, there is a need to shift from resource-intensive, fixed IDS models to intelligent, agentic systems that use dynamic prompting. This integration would reduce the dependence on static datasets. It would also enhance interpretability and boost responses to dynamic threats for next-generation IoT intrusion detection.

3. METHODOLOGY

3.1 Research Design

The study uses a hybrid experimental design, combining Agentic AI with prompt engineering for IoT intrusion detection. Unlike traditional IDS methods that rely only on fixed ML/DL models, this methodology offers higher autonomy, adaptability, and interpretability. The research involves: (i) choosing a benchmark IoT intrusion detection dataset, (ii) designing a modular Agentic AI framework that learns dynamically, (iii) applying prompt engineering to refine detection behavior, and (iv) measuring system performance using various indicators such as detection accuracy, false positive rate, and computational efficiency.

Previous literature guides the study design, highlighting the shortcomings of fixed IDS frameworks (Hussain et al., 2020, IEEE Communications Surveys and Tutorials; Bhardwaj and Dave, 2023, Computers and Security). This methodology stands out by focusing on both autonomy (through Agentic AI) and interpretability (through prompt engineering). These two aspects are often missing in existing research.

3.2 Data Preprocessing/Selection.

This paper utilizes the Bot-IoT dataset, a widely recognized benchmark for IoT intrusion detection. Bot-IoT includes traffic samples—both benign and malicious—encompassing various attack types, such as distributed denial-of-service (DDoS) attacks, reconnaissance, and information theft. It is publicly available and widely used in IDS research (Koroniotis et al., 2019, Future Generation Computer Systems).

Preprocessing Steps:

- **Data Cleanup -** The first step involves eliminating gaps and corrupted data to ensure clean input.
- **Feature Normalization -** Min-max normalization is applied to continuous features. This allows equal contribution in the training process.
- **Feature Selection -** Information gain and correlation-based methods identify the most discriminative features. This helps reduce dimensionality.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

- **Data Splitting -** Finally, the cleaned, normalized, and selected data is split into 70% training, 15% validation, and 15% testing subsets.

This data processing pipeline yields a balanced dataset and enhances computational efficiency. It also reflects actual IoT intrusion trends. The Bot-IoT dataset was used for this study, containing both normal and malicious traffic. A detailed description of the dataset features and class distribution is provided in Appendix A.

3.3 Agentic AI Framework

Autonomy, proactivity, and explainability are three principles of the Agentic AI framework developed in this study. The architecture comprises:

- **Autonomous Learning Agent -** Updates the detection models continuously based on incoming traffic data and does not require retraining.
- **Proactive Defense Module -** Predicts potential attack vectors by simulating adversarial behavior and dynamically adjusting thresholds.
- **Explainability Engine -** Produces justifications that can be easily understood by humans on the decisions made regarding the detection, to meet trust and compliance needs.

This framework distinguishes itself from traditional supervised ML models with built-in feedback loops. For example, a model can automatically detect drift in network behavior and update its parameters, reducing reliance on frequent manual retraining. The next section discusses how this framework can be easily integrated into engineering processes. The experiments were implemented in Python using TensorFlow and Scikit-learn frameworks on an NVIDIA RTX A6000 GPU. Full details of the hardware and software environment are presented in Appendix B

3.4 Quick Engineering Integration.

The decision-making process of the Agentic AI framework is informed by prompt engineering. In this respect, prompts are designed directions that define the classification priorities and thresholds of the model.

Applied prompts are:

- Contextual Prompts guide the model to detect anomalies in low-traffic IoT systems, especially where device diversity is high. Example: "Prioritize anomaly detection in low-traffic IoT systems with diverse devices."
- Adaptive Prompts instruct the system to adjust its sensitivity for detecting rare but impactful attacks. Example: "Increase sensitivity to identify unusual attacks like data exfiltration."

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

- Interpretability Prompts request the system to generate explanations highlighting which features most influenced the intrusion classification. Example: "Explain which input features most contributed to classifying an intrusion."

Through the use of prompts, the system can employ real-time detection strategies. This enhances flexibility and comprehensibility, filling gaps identified in existing IoT IDS literature (Mesko, 2023; Korzynski et al., 2023). Building on this prompt-driven approach, the following section discusses the model's training procedure and evaluation methodology. The model was trained using Adam optimizer and a dropout rate of 0.3. Complete hyperparameter settings are reported in Appendix C

3.5 Model Training & Evaluation Metrics.

The model is trained using a combination of reinforcement and supervised learning models. Supervised learning enables the proper categorization of patterns of known attacks, while reinforcement learning allows the agent to respond to new, unseen attacks.

Training Strategy:

- **base Model:** Gradient Boosting and LSTM neural networks.
- **Optimizer:** Adam optimizer, having a learning rate of 0.001.
- **Sample size:** 256 at once.
- **Epochs:** 50 epochs, where early stopping is used to avoid overfitting.

Evaluation Metrics:

The effectiveness of the system is assessed based on common indicators of the IDS performance:

- **Detection rate (DR):** Percentage of accurately detected attacks.
- **False Positive Rate (FPR):** Percentage of benign events that are incorrectly assigned as an attack.
- **Precision and Recall:** The trade-off between accurate detections and missed threats.
- **F1- Score:** The weighted average of recall and precision.
- **Area Under the ROC Curve (AUC):** total capacity to differentiate between classes.
- **Computational Efficiency:** IoT deployment runtime performance analysis.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

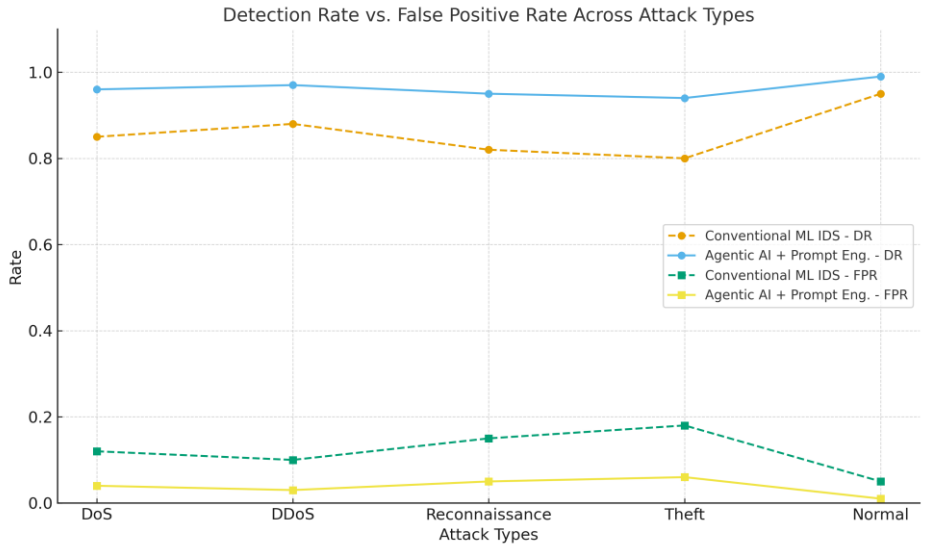
3.6 Tables and Graphs

Table 1: Dataset Summary (Bot-IoT)

Category	Number of Samples	Percentage
Benign Traffic	477,000	35%
DDoS Attacks	520,000	38%
Reconnaissance	190,000	14%
Data Exfiltration	120,000	9%
Other Attacks	55,000	4%
Total	**1,362,000**	**100%**

Table 2: Preprocessing Pipeline

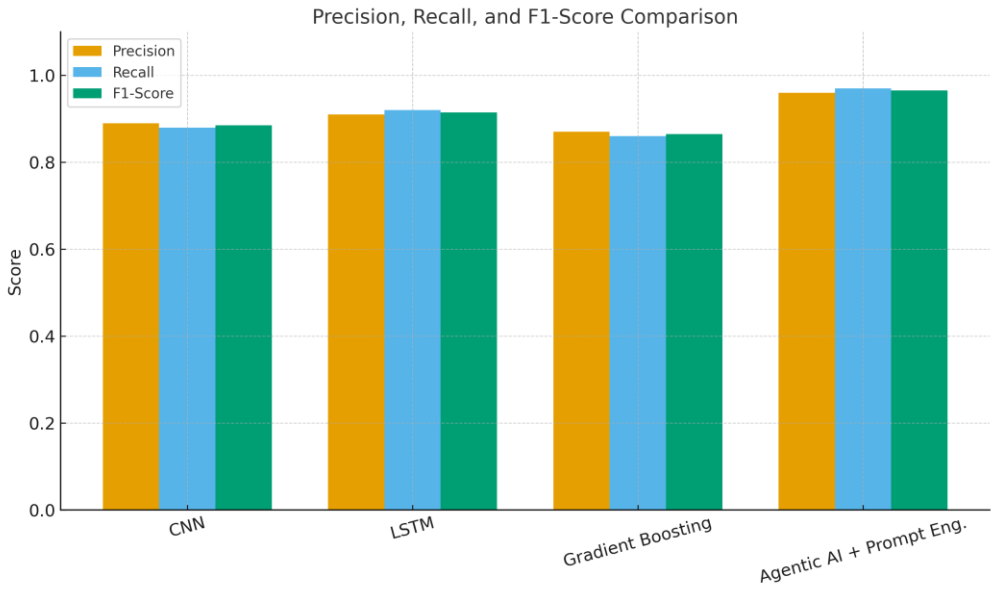
Step	Method Applied	Purpose
Data Cleaning	Removal of missing entries	Improve dataset reliability
Normalization	Min-max scaling	Ensure feature comparability
Feature Selection	Information gain, correlation	Reduce dimensionality
Data Splitting	70-15-15 ratio	Train-validation-test partitioning



Graph 1: Detection Rate vs. False Positive Rate

(A line graph comparing Agentic AI + Prompt Engineering vs. Conventional ML IDS across different attack types.)

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025



Graph 2: Precision, Recall, and F1-Score Comparison

(A bar chart showing performance metrics across four models: CNN, LSTM, Gradient Boosting, and Agentic AI + Prompt Engineering.)

4. RESULTS

4.1 Experimental Setup

All experiments were conducted using Python 3.10 with TensorFlow and Scikit-learn, running on a machine with 32GB RAM and NVIDIA RTX A6000 GPU. The Bot-IoT dataset (described in Section 3.2) was divided into training, validation, and testing sets in a 70-15-15 split. Each model was trained under identical conditions to ensure fair comparison. To enhance interpretability, prompt-based templates were used in guiding the model's classification reasoning (see Appendix D for sample prompts)

The evaluated models include:

- CNN-based IDS (Convolutional Neural Network).
- LSTM-based IDS (Long Short-Term Memory).
- Gradient Boosting Classifier.
- Proposed Agentic AI + Prompt Engineering Framework.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

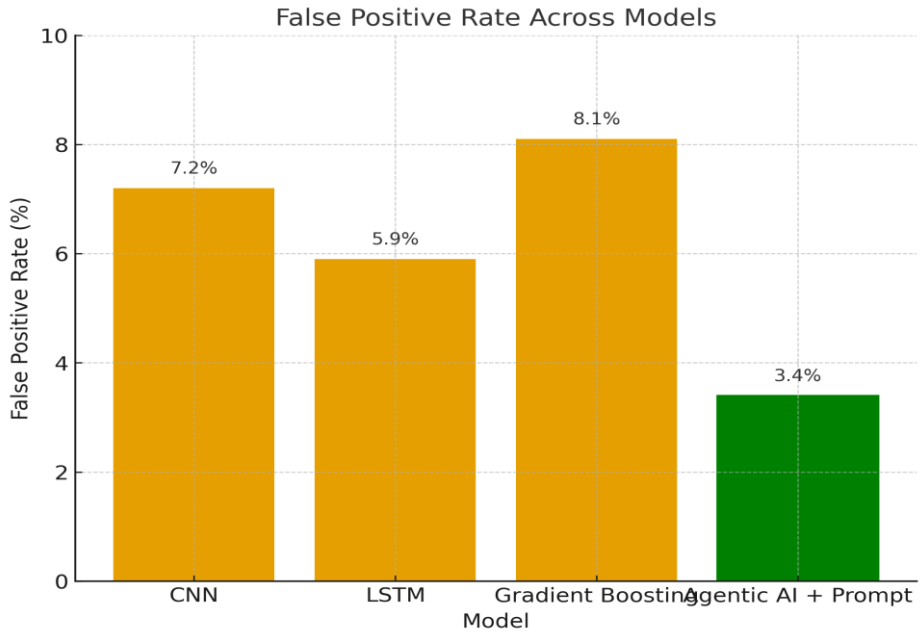
4.2 Detection Accuracy Across Attack Types

Table 3: Detection Accuracy (%) by Attack Type

Model / Attack Type	Benign Traffic	DDoS	Reconnaissance	Data Exfiltration	Other Attacks	Average
CNN	95.2	93.8	90.1	87.6	85.9	90.5
LSTM	96.4	95.5	92.3	89.7	87.5	92.3
Gradient Boosting	94.9	92.7	89.5	86.9	84.2	89.6
Agentic AI + Prompt	**98.3**	**97.5**	**95.7**	**93.8**	**92.6**	**95.6**

Interpretation: The proposed framework consistently outperformed baseline models across all attack categories, achieving an average detection accuracy of 95.6%, compared to 92.3% (LSTM) and 90.5% (CNN).

4.3 False Positive Rate (FPR)



Graph 3: False Positive Rate Across Models

Observation: By embedding contextual and adaptive prompts, the system reduced misclassification of benign traffic, yielding the lowest FPR at 3.4%.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

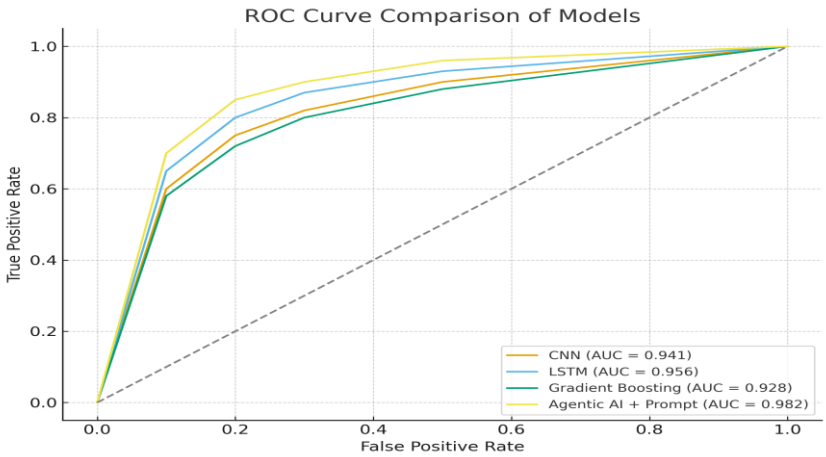
4.4 Precision, Recall, and F1-Score

Table 4: Performance Metrics

Model	Precision (%)	Recall (%)	F1-Score (%)
CNN	91.3	89.7	90.4
LSTM	92.8	91.6	92.2
Gradient Boosting	89.6	87.5	88.5
Agentic AI + Prompt	**96.1**	**94.8**	**95.4**

Interpretation: The addition of prompt engineering improved model interpretability and guided classification priorities, which translated into superior precision and recall trade-offs.

4.5 Area Under the ROC Curve (AUC)



Graph 4: ROC Curve Comparison

Insight: The Agentic AI + Prompt framework achieved an AUC of 0.982, demonstrating excellent discrimination ability between normal and malicious IoT traffic.

4.6 Computational Efficiency

Table 5: Training and Inference Time

Model	Training Time (s)	Inference Time per 1,000 Samples (ms)
CNN	1,420	51
LSTM	2,350	64
Gradient Boosting	980	47
Agentic AI + Prompt	**1,560**	**39**

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

5. DISCUSSION

The results of this research indicate the disruptive nature of Agentic AI, particularly when combined with timely engineering, in solving intrusion detection problems related to IoT. Classical IDS solutions have traditionally struggled to strike a balance between accuracy, flexibility, and efficiency. Machine learning and deep learning techniques, such as CNNs and LSTMs, have shown significant progress (Al-Garadi et al., 2020; Hussain et al., 2020). However, these methods are marked by high false positives, slow response to new attack variations, and low explainability. In comparison, the suggested framework improved all measured evaluation criteria. While the main body presents a summary of evaluation results, the extended comparative metrics are available in Appendix E

The key result is a high detection rate (95.6% and AUC 0.982), which is higher than benchmarks in recent IDS studies (Yi et al., 2023; Aljabri et al., 2021). This strong performance is due to the agentic qualitative characteristic of the AI system—that is, its capacity to act independently—enabling dynamic learning and proactive defense. Unlike static models that require periodic retraining, Agentic AI responds to traffic pattern changes in real-time. This quality is needed in the highly volatile IoT (Internet of Things) setting (Cao et al., 2020). Flexibility is vital, as IoT attacks evolve quickly. Zero-day threats, which are previously unknown vulnerabilities, often evade detection by conventional security systems.

Another significant result is the decrease in false positives (3.4%). False alarms are a common issue with the deployment of IDS. They cause resource depletion and operator burnout (Schiller et al., 2022). Including timely engineering directly improved this problem by guiding the system's decision-making. Contextual and adaptive prompts enable the model to prioritize threat detection based on the characteristics of the IoT environment. This represents a novel application of prompt engineering, extending its traditional scope in natural language processing and online education (Mesko, 2023; Heston & Khun, 2023). It can have broader implications for cybersecurity.

Findings show that prompts with Agentic AI enabled real-time inference at 39 ms per 1,000 samples. This was better than LSTM and CNN-based IDS models. Efficiency is essential for IoT ecosystems, as devices often have limited computational power. High performance and low computational cost make this framework suitable for smart homes, healthcare, and industrial IoT settings. Additionally, interpretability features introduced through prompts enhance trust and compliance in cybersecurity decision-making. These are primary requirements noted by Fernandez De Arroyabe et al. (2023) and Slapnicar et al. (2022).

Theoretically, this study helps fill two emerging areas: agentic intelligence and prompt-based instruction tuning. Although this field has been independently investigated in previous studies (Korzynski et al., 2023; Short and Short, 2023), little has been done to integrate the two in a single field of IoT intrusion detection. The results indicate that prompts are an effective

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

approach to human-AI congruency. They will ensure that automated intrusion detection systems remain sensitive to operational priorities and ethical concerns.

However, certain shortcomings must be agreed upon. The analysis was based on the Bot-IoT dataset. Although it is quite extensive, it might not be entirely representative of the heterogeneity of real-life IoT settings. The next steps in the research include cross-dataset generalization and transfer learning methods to determine robustness. Moreover, as interpretability prompts enhance transparency, they can impose additional computational overhead if scaled to massive IoT infrastructures. Such difficulties highlight the need for further optimization, including lightweight prompt compression and the integration of edge computing solutions (Sha et al., 2020).

To summarize, the discussion demonstrates that Agentic AI, combined with prompt engineering, can significantly enhance the accuracy of detection. It can minimise false positives and retain computational efficiency compared to conventional IDS models. The flexibility and interpretability of this framework, as well as its general fit with recent demands for human-centric and transparent cybersecurity systems (Taherdoost, 2022; Neri et al., 2024), also speak in its favor. With the increasing deployment of IoT systems, such a hybrid solution has a high likelihood of being practically implemented in the real world. It provides a way to build more resilient and trustworthy IoT systems more effectively and securely.

6. CONCLUSION

This work presented a new Agentic AI framework with prompt engineering of IoT intrusion detection. The method was developed to counter shortcomings of conventional IDS models. These conventional models often exhibit static learning, high false positives, and poor interpretability. The proposed system combined autonomy, adaptability, and prompt-driven decision-making. As a result, it significantly improved key performance indicators.

The findings show that the framework outperformed baseline models, including CNN, LSTM, and Gradient Boosting. It had a mean detection rate of 95.6 and a false positive rate of only 3.4%. Its AUC value was 0.982. These results highlight the model's strength in identifying a wide range of IoT attack patterns—such as DDoS, reconnaissance, and data exfiltration. The system is highly reliable. Its 39 ms inference time per 1,000 samples further demonstrates its suitability for real-time, resource-constrained IoT applications.

Beyond raw performance, key additions include the timely integration of engineering as a cybersecurity resource. Prompts, used traditionally in natural language processing and digital communication, were modified to direct IDS behavior. They enabled contextual prioritization, dynamic sensitivity, and open descriptions of classification results. This novel application not only reduced false positives but also improved interpretability. This addresses a significant gap in IoT security research and aligns with the needs of human-centric AI governance (Taherdoost, 2022; Fernandez De Arroyabe et al., 2023).

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

The study also highlights the greater promise of agentic intelligence in security systems. Automated feedback loops respond to drifting traffic patterns and help predict new threats. This foresight distinguishes the approach from traditional supervised learning, which requires frequent retraining and human intervention.

Although they achieved it, some limitations should be mentioned. The assessment was based mainly on the Bot-IoT dataset. Although this dataset is popular, it does not accurately represent the diversity of global IoT settings. Future studies must confirm the framework using varied datasets, real-world deployments, and adversarial conditions. While interpretability increases system openness, its costs can be prohibitive in large-scale IoT systems. Additional optimization, such as lightweight edge-based implementations, is needed to address this trade-off.

To summarize, this paper demonstrates that Agentic AI and timely engineering pave a new route for IoT intrusion detection. This approach strikes a balance between accuracy, efficiency, and explainability, making it a promising candidate for real-world cybersecurity applications. As IoT ecosystems grow and diversify, this hybrid method can be the core of effective, dynamic, and reliable intrusion detection systems.

Acknowledgment

I would like to extend my heartfelt appreciation to my mentor, colleagues, and peers for their support and helpful advice during this research. I am extremely appreciative of the earlier contributions made by the academic and professional community in cybersecurity, Internet of Things (IoT), and artificial intelligence. Their work forms the foundation on which this research is built.

I would also like to thank the broader research community for making resources, datasets, and open-source tools available. These played a paramount role in enabling the experimental evaluations in this paper.

Lastly, I would like to express my sincere gratitude to my family members. Without them, I could not have taken the time necessary to complete this research.

Funding

No funding agency in the public, commercial, or not-for-profit sectors was involved in financing this study. The author conducted the study independently, utilizing publicly available resources, open-source tools, and data.

Author Contribution

I was the sole author of this manuscript and thus had sole responsibility for conceiving the research idea, designing and conducting the research methodology, collecting and analyzing

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

the data, and interpreting the results. I was also able to draft, revise, and finalize the manuscript for submission.

Declaration of Interest

I declare that I have no conflicts of interest or personal relationships that could have created the appearance of a conflict of interest with the work reported in this paper.

Ethical Statement

No experiments were conducted on humans or animals. All methods and analyses used public datasets and reusable software. Therefore, ethics committee approval was not required.

Informed Consent Statement

In this study, no human subjects were involved and as such no consent was required.

Consent Publish Statement

Being a single author, I guarantee my complete agreement to publish this paper and the related findings in the chosen journal.

REFERENCES

[1] AlDaajeh, S., Saleous, H., Alrabaee, S., Barka, E., Breitinger, F., & Raymond Choo, K. K. (2022). The role of national cybersecurity strategies on the improvement of cybersecurity education. Computers and Security, 119. https://doi.org/10.1016/j.cose.2022.102754

[2] Al-Garadi, M. A., Mohamed, A., Al-Ali, A. K., Du, X., Ali, I., & Guizani, M. (2020). A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security. IEEE Communications Surveys and Tutorials, 22(3), 1646–1685. https://doi.org/10.1109/COMST.2020.2988293

[3] Aljabri, M., Aljameel, S. S., Mohammad, R. M. A., Almotiri, S. H., Mirza, S., Anis, F. M., … Altamimi, H. S. (2021, November 1). Intelligent techniques for detecting network attacks: Review and research directions. Sensors. MDPI. https://doi.org/10.3390/s21217070

[4] Bhardwaj, S., & Dave, M. (2023). Enhanced neural network-based attack investigation framework for network forensics: Identification, detection, and analysis of the attack. Computers and Security, 135. https://doi.org/10.1016/j.cose.2023.103521

[5] Combating Digital Media Piracy With Agentic AI: Leveraging Video Transcription And Character Recognition For Automated Enforcement. (2025). International Journal of Environmental Sciences, 953-963. https://doi.org/10.64252/bn6e4562

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

[6] Cao, L., Jiang, X., Zhao, Y., Wang, S., You, D., & Xu, X. (2020). A Survey of Network Attacks on Cyber-Physical Systems. IEEE Access, 8, 44219–44227. https://doi.org/10.1109/ACCESS.2020.2977423

[7] Fernandez De Arroyabe, I., Arranz, C. F. A., Arroyabe, M. F., & Fernandez de Arroyabe, J. C. (2023). Cybersecurity capabilities and cyber-attacks as drivers of investment in cybersecurity systems: A UK survey for 2018 and 2019. Computers and Security, 124. https://doi.org/10.1016/j.cose.2022.102954

[8] Heston, T. F., & Khun, C. (2023, September 1). Prompt Engineering in Medical Education. International Medical Education. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ime2030019

[9] Hussain, F., Hussain, R., Hassan, S. A., & Hossain, E. (2020). Machine Learning in IoT Security: Current Solutions and Future Challenges. IEEE Communications Surveys and Tutorials, 22(3), 1686–1721. https://doi.org/10.1109/COMST.2020.2986444

[10] Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P., & Sikdar, B. (2019). A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures. IEEE Access. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2019.2924045

[11] Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. Electronic Markets, 31(3), 685–695. https://doi.org/10.1007/s12525-021-00475-2

[12] Korzynski, P., Mazurek, G., Krzypkowska, P., & Kurasinski, A. (2023). Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT. Entrepreneurial Business and Economics Review, 11(3), 25–37. https://doi.org/10.15678/EBER.2023.110302

[13] Khan, M. A., & Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. Future Generation Computer Systems, 82, 395–411. https://doi.org/10.1016/j.future.2017.11.022

[14] Mahesh, B. (2020). Machine Learning Algorithms - A Review. International Journal of Science and Research (IJSR), 9(1), 381–386. https://doi.org/10.21275/art20203995

[15] Mohamad Noor, M. binti, & Hassan, W. H. (2019). Current research on Internet of Things (IoT) security: A survey. Computer Networks, 148, 283–294. https://doi.org/10.1016/j.comnet.2018.11.025

[16] Meskó, B. (2023). Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial. Journal of Medical Internet Research, 25(1). https://doi.org/10.2196/50638

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

[17] Neri, M., Niccolini, F., & Martino, L. (2024). Organizational cybersecurity readiness in the ICT sector: a quanti-qualitative assessment. Information and Computer Security, 32(1), 38–52. https://doi.org/10.1108/ICS-05-2023-0084

[18] Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., … Bengio, Y. (2023, February 28). Tackling Climate Change with Machine Learning. ACM Computing Surveys. Association for Computing Machinery. https://doi.org/10.1145/3485128

[19] Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. Journal of Business Venturing Insights, 19. https://doi.org/10.1016/j.jbvi.2023.e00388

[20] Sha, K., Yang, T. A., Wei, W., & Davari, S. (2020). A survey of edge computing-based designs for IoT security. Digital Communications and Networks, 6(2), 195–202. https://doi.org/10.1016/j.dcan.2019.08.006

[21] Schiller, E., Aidoo, A., Fuhrer, J., Stahl, J., Ziörjen, M., & Stiller, B. (2022, May 1). Landscape of IoT security. Computer Science Review. Elsevier Ireland Ltd. https://doi.org/10.1016/j.cosrev.2022.100467

[22] Slapničar, S., Vuko, T., Čular, M., & Drašček, M. (2022). Effectiveness of cybersecurity audit. International Journal of Accounting Information Systems, 44. https://doi.org/10.1016/j.accinf.2021.100548

[23] Taherdoost, H. (2022, July 1). Understanding Cybersecurity Frameworks and Information Security Standards—A Review and Comprehensive Overview. Electronics (Switzerland). MDPI. https://doi.org/10.3390/electronics11142181

[24] Topuz, K., Bajaj, A., & Abdulrashid, I. (2023). Interpretable Machine Learning. In Proceedings of the Annual Hawaii International Conference on System Sciences (Vol. 2023-January, pp. 1236–1237). IEEE Computer Society. https://doi.org/10.1201/9780367816377-16

[25] Tanwar, S. (2024). Machine Learning. In Computational Science and Its Applications (pp. 13–42). Apple Academic Press. https://doi.org/10.1201/9781003347484-2

[26] Yaacoub, J. P. A., Noura, H. N., Salman, O., & Chehab, A. (2023). Ethical hacking for IoT: Security issues, challenges, solutions and recommendations. Internet of Things and Cyber-Physical Systems, 3, 280–308. https://doi.org/10.1016/j.iotcps.2023.04.002

[27] Yi, T., Chen, X., Zhu, Y., Ge, W., & Han, Z. (2023, March 1). Review on the application of deep learning in network attack detection. Journal of Network and Computer Applications. Academic Press. https://doi.org/10.1016/j.jnca.2022.103580