



Real Time AI-Driven Threat Detection with the Integration of Zero Trust Security Framework

Shreyanshi Srivastava^a, Anshika Prajapati^b and Mr. Mahesh kumar Tiwari^c, Mr. Amit kumar Srivastava^d

^{a,b,c,d} Department of Computer Science, National P.G. College, Lucknow, India

shrivastavashreyanshi111@gmail.com^a, anshika16082005@gmail.com^b, maheshyogi26@gmail.com^c

amit_sri_in@yahoo.com^d

KEYWORD

Artificial Intelligence, ZTA framework, Real time threat detection, cybersecurity, ANN, CNN.

ABSTRACT

This paper elucidate how the Artificial Intelligence with Zero Trust Security determine alert dramatically threat detection and response, this offering a robust security solution against cyber threats. The cooperation of continuous verification, context based authentication, and real-time threat analysis not only reduce risks but also enhances overall security posture. This introduce the use of DL and ML techniques such as ANN and CNN in threat detection that how they mitigate with real time threats. Cybersecurity with zero trust framework and artificial intelligence is an innovative technology in real time threat detection process and adaptive response system. The amalgamation of modern Artificial Intelligence features into zero trust system is forecast to create new opportunities for enhancing adaptive security. This includes automating processes like segmentation, detecting subtle deflections or threats and refining techniques to mitigate risks. Most importantly, it strengthens the principles of Zero Trust and addresses issues associated with minimal and rule based security solutions.

1. Introduction

Real-time threat detection is crucial in modern cybersecurity, enabling organizations to identify threats and malicious activities as they occur. Unlike traditional approaches, where threats are detected after damage occurs, real-time monitoring enables immediate threat identification of threats and other suspicious activities. In this approach an ongoing monitoring occurs as the constant surveillance system of network security that could indicate cyber threat. Whenever a threat is detected automated notifications notify the security teams that they can take a significant action against it. By using the most up to date threat intelligence a business can know the newest attack types, malware identification and vulnerabilities that is being exploited. It also helps organizations quickly address any emerging threats for purpose of

Corresponding Author: Shreyanshi Srivastava^a, Department of Computer Science, National P.G. College, Lucknow, India

Email: shrivastavashreyanshi111@gmail.com

preventing them, which means that the existing security procedures and measures can reduce threats to their barest minimum (Gartner, 2021). Adaptive defense mechanisms enhance real-time threat intelligence with a versatile and reactive security approach [1,20,21,22]. The zero trust security is a framework that assumes no user is trusted within the organization or in network by default. Zero trust reduces the risk of both internal and external threats by requiring continuous verification of every entity whether it is user or any device which is trying to access resources. It is data-centric approach used for better security controls between users, assets and data as they change over time. NIST defines a zero trust security(ZTA) and acknowledging that threat exist both inside and outside the network. It is different from the perimeter-based security architectures; the trust of an object is independent of its physical location and all objects are untrusted by default. The trust of an object can only be obtained by identity authentication and trust evaluation. After the system assigns the relative permissions to the object, the object can perform related operations. In recent years, zero trust architecture has been initially applied, and the most typical example is Google's BeyondCorp model [2,24,25,26]. A survey conducted by Microsoft Security in 2021 according to this mostly cybersecurity professionals believe in zero trust security.

1.1 Review of Existing Research

AI plays a very crucial role in zero trust security framework by introduce an intelligent data analysis and anomaly detection. AI technologies like machine learning(ML) and deep learning are able to process the large amount of data in real-time. Therefore, the AI-driven threat detection in real time can help to detect the threat when they occur.

1.1.1 Zero Trust Security Framework

Author /Year of publication of paper	Core Principle	Description	Key contribution from Research Papers
Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020).	Architectural Framework and Standardization	The recent research focuses on formal models, integration with existing technology and standard approaches to make ZTA actionable.	NIST SP 800-207 (2020) remains the foundational document. Subsequent research builds on it, proposing enhancements for cloud-native environments, IoT, and formal verification of policies.[13,14]
Kindervag, J. (2010).	Never Trust, Always Verify	It focuses on the principle of continuous verification, least privilege access and tight identity management to lessen the risk and prevent unauthorized access.	The paper argued that traditional perimeter defenses are obsolete and defined the three core concepts of Zero Trust: ensuring all resources are accessed securely, adopting a least-privilege strategy, and inspecting and logging all traffic.[15]
R. T. Aboaoja, et al. (2023).	Implementation Challenges &	Critically evaluating the	Identify and categorize the key challenges in ZTA

Z. T. Al-Saqqa, et al. (2022).	Economic Trade-offs	practical hurdles of ZTA adoption, including performance overhead, complexity, cost, and usability.	implementation, such as interoperability, performance latency, and significant initial investment. Discusses the "human factor," highlighting how complex authentication and continuous validation can lead to user friction and resistance, proposing usability-focused design principles for ZT systems.[16,17]
F. Alawida, et al. (2023). W. Wang, et al. (2021).	Integration with Emerging Paradigms: SASE & Cloud	Exploring how ZTA converges with other modern security frameworks, particularly Secure Access Service Edge (SASE), which combines ZTA with SD-WAN capabilities delivered as a cloud service.	Analyzes the synergy between ZTA and SASE, positioning ZTA as the identity and access control core of the SASE framework. Proposes a specific ZTA model for multi-cloud and hybrid cloud environments.[18,19]

1.1.2 Integration Zero Trust and AI

The National Institute of Standards and Technology SP 800-207 is the most widely used zero trust framework, it emphasizing continuous authentication such as multi factor and behavioural biometrics, micro segmentation like granular network access control, policy enforcement points and policy decision points. BeyondCrown shifts access controls from the network perimeter to individual devices and users, implementing device trust scoring, context aware access policies, user to application encryption. The limitations of this model are scalability challenges in heterogeneous environments and limited AI driven anomaly detection capabilities. Zero trust architecture works on a principle that "Never Trust Always Verify" and the integration of real time threat detection using AI enhances its performance as Zero Trust demands continuous validation and AI predicts anomalies before they strike. The integration of AI in security monitoring and detection has demonstrated remarkable capabilities in cloud-native environments. Analysis by Ramasamy et al. shows that contemporary machine learning models achieve network traffic processing rates of 4.7 terabits per second while maintaining a 99.89% threat detection accuracy rate [9]. Their study of 178 enterprise deployments revealed that AI-powered detection systems reduced false positive rates by 91.3%

compared to traditional rule-based systems, while improving initial threat detection speed by 96.8%[10].

Performance Comparison of Traditional vs AI – driven Threat Detection

1. Traditional Detection Speed: 200 events/sec
2. Traditional Incident Response Time: 18 minutes
3. AI-driven Detection speed: 4000 events/sec
4. Ai- driven Incident Response Time: 2 minutes

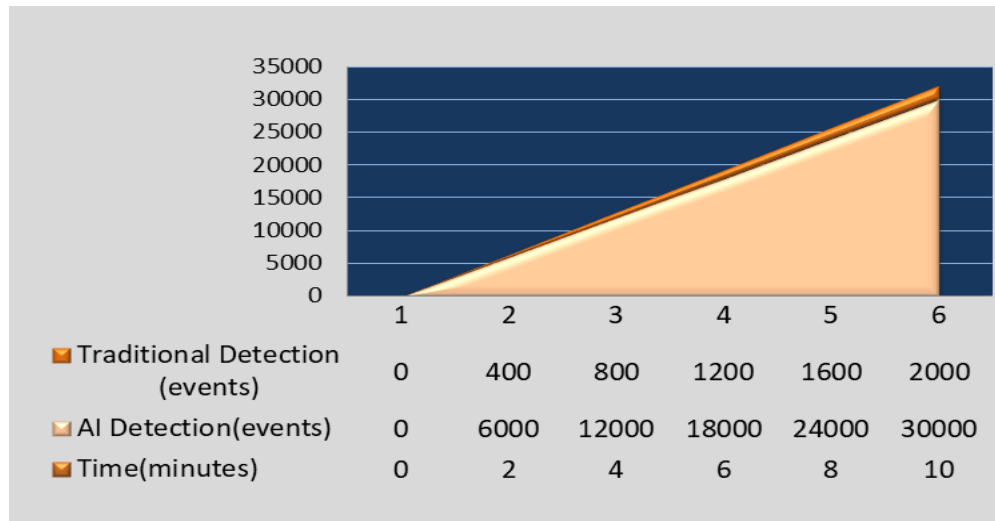


Figure -1: Comparison of Ai-Driven Threat Detection and Traditional Threat Detection

1.2 Challenges in Traditional Threat Detection

Challenge Category	Challenges	Impact
Adaptability and Evolution	1.It is ineffective against Zero day attacks: Means can not detect noval threats or new variants of malware. 2.Unable to detect sophisticated evasion techniques.	Keeps the organization vulnerable to the most harmful and focused attacks until a signature is established and put into action.
Operational overhead	1.Traditionally a constant manual updates were reauired by human analysts.	This create reactive security posture because of large administrative burden,costly and prone to human error. Impact system performance and increases latency in network traffic and endpoint response times.
Detection capabilities	High false positives ,Lack of contextual awareness	High false positives leads to alert fatigue ,causing analysts to miss actual threats buried in noise. Lack of contextual awareness there is a inability to distinguish between malicious action and legitimate action.

Architectural Limitations	Network centric design, Ineffective in distributed environment	This fails in zero trust because of ZT requires continuous validation regardless of origin.
Response capabilities	Slow response time , No augmented investigation	Heavily dependent on human analysts to investigate alerts which cause slow response.

2. AI-Driven Threat Detection

ML(Machine Learning) and DL(Deep Learning) techniques are played a wide role in threat detection. Machine learning technique such as supervised learning for known threats and unsupervised learning for anomaly detection are used to analyze network traffic to unusual patterns, while Deep Learning technique such as Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) process large network data , detecting complex patterns in payloads that signal malware or phishing attempts.

2.1 ANN in Real Time Threat Detection

ANN is foundation of DL models which is inspired by human brain, consisting of interconnected nodes like human brain contains neurons. These nodes organized in hidden layers , it learns patterns from data through training and can generalize to detect anomalies. By continuously monitoring user and device behavior ANNs determinig a standard profiles and flag deviations that may indicate malicious intent. ANNs are different from traditional rule-based systems because they learn from past attack patterns and use that information to dynamically classify network traffic as either benign or malicious. This lets them find both known threats and new ways to attack. Their adaptive learning system also makes sure that they keep getting better. As the model processes new data, it gets better at spotting complex cyber threats in changing Zero Trust environments.

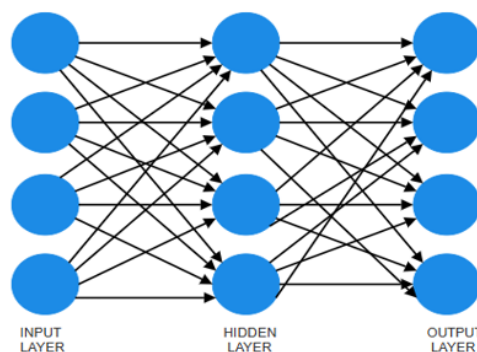


Figure-2 : Working of ANN

2.1.1 How ANN works

ANN consist of many layers such Input layer, Hidden layer, Output layer. The input layer of this ANN is used to convert and prepare the raw data's features into a format that the neural

network understands. It deals with important parameters of packets like source IP, destination IP, packet size and packet timestamps, which gives a more detailed picture of the activities that occur in the network[10]. Hidden layer process through weighted connections and activation functions such as ReLU ,Sigmoid. These layers take the patterns from the input data and clean them to provide the best patterns for the network to distinguish normal from malicious activities[10]. Then comes output layer which produces a prediction, it uses softmax of activation function for detecting threats and classify them in different classes. This allows the model to cluster activities into different types of threats, which may consist of activities such as malware, phishing and benign traffic, with a level of confidence[10].

To explicitly illustrate the integration of Ai-driven threat detection within the Zero Trust framework , here a detailed mathematical example of ANN for real time threat detection. While the use of ANNs for behavioral analytics is well established [11,12], this example explain the specific machanisms by which such a model operates within a Zero Trust control loop. Here , a simulation of data exfiltration scenario, calculating a real time risk score from a set of features and illustrate how this score directly informs a policy decision point to enact the ‘never trust, always verify’ mandate.

1. Input :- Hidden layer (matrix multiply): Each neuron will receive a weighted sum of inputs.

Input vector, $X = [1.0, 0.0, 0.0, 1.0]$.

Formula, $N = (\text{input} * \text{weight}) + \text{bias}$

$W1 = [0.82, 0.25, -1.10, 0.65]$

$W2 = [-0.45, 1.60, 0.50, -0.90]$

$W3 = [1.20, -0.75, 0.95, 0.30]$

- First neuron, $n1: 1.0 \times 0.82 + 0.0 \times 0.25 + 0.0 \times (-1.10) + 1.0 \times 0.65 = 1.47$
- Second neuron, $n2: 1.0 \times (-0.45) + 0.0 \times 1.60 + 0.0 \times 0.50 + 1.0 \times (-0.90) = -1.35$
- Third neuron, $n3: 1.0 \times 1.20 + 0.0 \times (-0.75) + 0.0 \times 0.95 + 1.0 \times 0.30 = 1.50$

So $X \cdot W^h = [1.47, -1.35, 1.50]$.

2. Add hidden biases ($b^h = [-0.30, 0.15, 0.05]$) :

○ $z_1 = 1.47 + (-0.30) = 1.17$

○ $z_2 = -1.35 + 0.15 = -1.20$

○ $z_3 = 1.50 + 0.05 = 1.55$

This is a raw activation , $z^h = [1.17, -1.20, 1.55]$.

3. ReLU activation: Activation function is used to learn complex pattern.

If $z > 0$, z remains same

If $z < 0$, $z * 0.01$

- $a_1 = 1.17$
- $a_2 = 0.01 \times (-1.20) = -0.012$
- $a_3 = 1.55$

This is a hidden output, $a^h = [1.17, -0.012, 1.55]$. Here, negative neuron shrunk into -0.012.

$$\begin{aligned}
 4. \text{ Output layer : } W^o &= [2.1, -1.5, 1.8] \text{ and } b^o = -0.9 \\
 z^o &= (1.17 \times 2.1) + (-0.012 \times -1.5) + (1.55 \times 1.8) + (-0.9) \\
 &= 2.457 + 0.018 + 2.79 - 0.9 \\
 &= 4.365
 \end{aligned}$$

$$\begin{aligned}
 5. \text{ Sigmoid:- Sigmoid compressed the value between 0 and 1.} \\
 \hat{y} = 1 / (1 + e^{(-4.365)}) \approx 0.98744498
 \end{aligned}$$

Final result: Risk score $\approx 0.9874 \rightarrow 98.74\%$ (very high)

Interpretation and Zero Trust Enforcement:

The ANN produces a high risk score of 0.9874. This score is influenced by the extreme deviation in data volume (x_1), the suspicious timing (x_2), and the high-risk destination (x_3).

In a Zero Trust framework, this score is sent to the Policy Decision Point (PDP). Based on a dynamic policy the PDP would immediately: Terminate the ongoing data transfer. Revoke the user's session token, requiring re-authentication. Log the incident and alert the Security Operations Center for immediate investigation. Trigger a step-up authentication challenge if the user attempts to re-access the data. This shows the essential Zero Trust principle: trust is never implicit and is continuously validated through real-time, data-driven analysis.

2.2 CNN in Real Time Threat Detection

CNNs are great at finding threats in real time because they look at data in a way that is similar to how the human visual system processes images. CNNs look at network traffic, file binaries, or system logs as visual patterns in cybersecurity. For example, they turn malware code into grayscale images where byte values become pixel intensities. As packets move across the network, slices of the packets are treated as two-dimensional images, and the CNN scans this “digital imagery” using filters. These filters detect various signs of threats, such as repeating byte sequences which may signal the presence of a DDoS attack, or specific patterns of pixels within a binary file which may indicate the presence of ransomware encryption routines.

2.2.1 How CNNworks

Like an ANN, CNN also works on different layers such as convolutional layer, pooling layers, fully connected layers. Convolutional layer applies filters to detect local patterns, pooling layer is used to reduce dimensionality means this is a technique used to decrease number of spatial dimensions in a feature map, while retaining the important information. Fully connected layers is a final classification, it extracts the high level feature from convolutional and pooling layers and use them to differentiate the input as either malicious or benign.

Here an example to describe the working of CNN.

$$X = \begin{bmatrix} 1.0 & 0.95 \\ 0.98 & 0.9 \\ 1.0 & 0.98 \end{bmatrix}$$

Step 1: Convolution Layer

Here we, use filter of size 2×2 .

$$F = \begin{bmatrix} 1.0 & 0.5 \\ 0.8 & 1.2 \end{bmatrix}, \text{ bias } b = -0.1$$

Sliding the filter over the input:

Position 1:

$$Z_1 = (1.0 \times 1.0 + 0.95 \times 0.5 + 0.98 \times 0.8 + 0.9 \times 1.2) + (-0.1)$$

$$=3.339-0.1= 3.239$$

$$\text{Apply ReLU: } A1=\max(0,3.239)=3.239.$$

Position 2:

$$Z2=(0.98*1.0+0.9*0.5+1.0*0.8+0.98*1.2)-0.1$$

$$=3.306$$

$$\text{ReLU: } A2=\max(0,3.306)=3.306.$$

Step 2: Pooling Layer

Max of convolution outputs that specifies the strongest suspicious pattern:

$$P=\max\{A1,A2\}=\max(3.239,3.306)=3.306.$$

Step 3: Fully Connected Layer

weight $W= 1.5$ and bias $B = -0.5$

$$Zout=P*W+B =3.306\times 1.5-0.5=4.459$$

Step 4: Output Layer (Sigmoid)

$$y^{\wedge}=1/(1+e^{-Zout})= 1/(1+e^{-4.459}) \approx 0.988$$

Interpretation: Risk score $\approx 99\%$, very high.

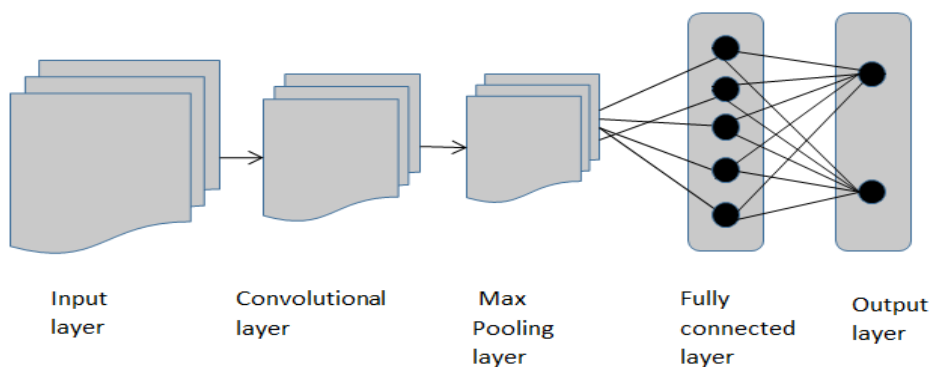


Figure-3 Working of CNN

3.Comparative Analysis of Threat Detection Models

Here is a comparison of different real time threat detection models.

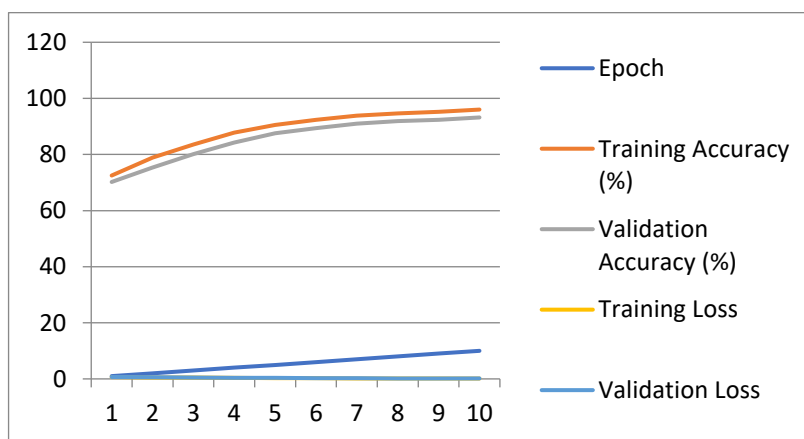


Figure-4 : Signature based model

1. This shows that early performance plateau, model peaks within 2-3 epochs, signature-based learning is quick but has no capacity to improve beyond known rules.
2. The model shows near 100% training accuracy, demonstrating perfect detection of known signatures.

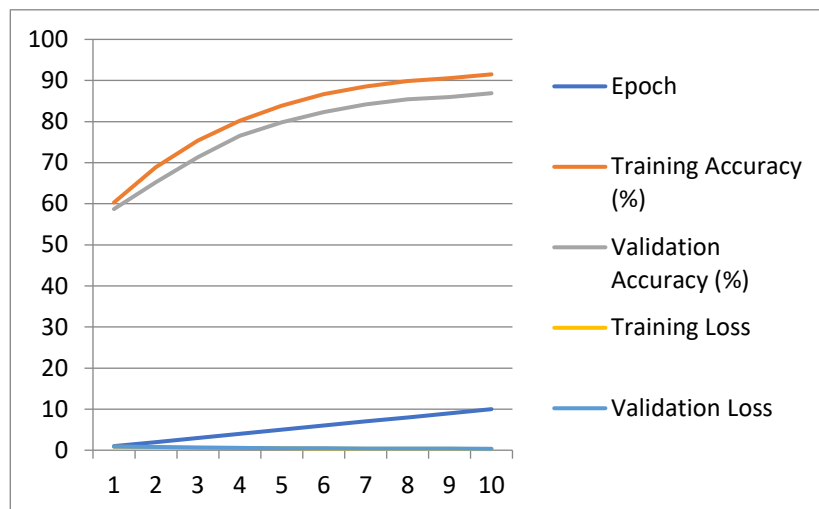


Figure-5 : Anomaly based model

1. This shows slower, progressive learning means accuracy improves gradually over all 10 epochs, showing the model is learning complex patterns, not just memorizing rules.
2. Starts with lower accuracy but ends with higher validation accuracy, showing it learns to detect novel anomalies.

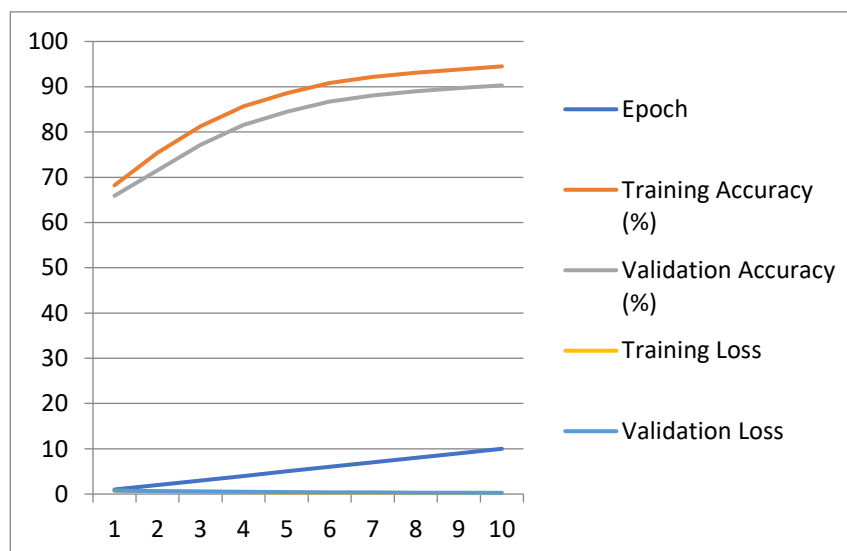


Figure-6 : Machine learning based model

1. This shows steady, continuous improvement means both accuracy and loss show consistent improvement across all 10 epochs, demonstrating true statistical learning.
2. Validation loss reaches a clear minimum point, identifying the optimal stopping epoch and showing well-tuned model complexity.

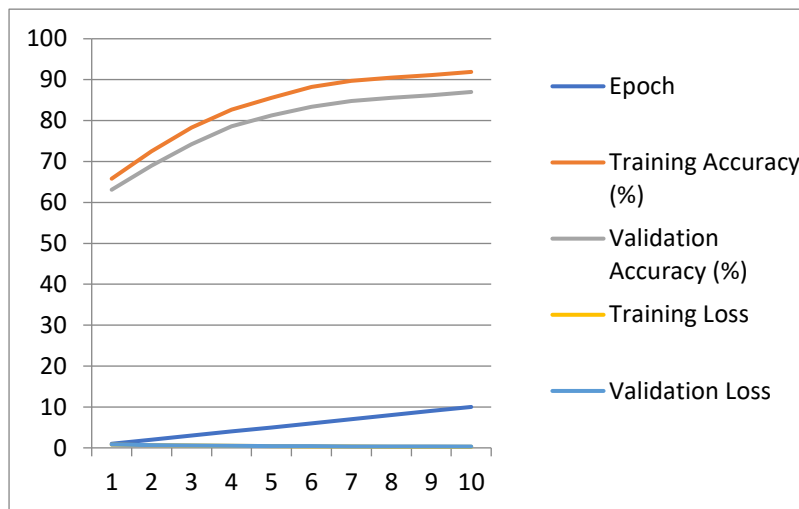


Figure-7 : Heuristic Based model

1. This shows erratic, unreliable learning because accuracy jumps unpredictably between epochs (e.g., 60% to 85%), showing heuristic rules provide inconsistent results.
2. Validation accuracy remains low and volatile (around ~70-75%), proving heuristic-based detection is inadequate for reliable threat identification.

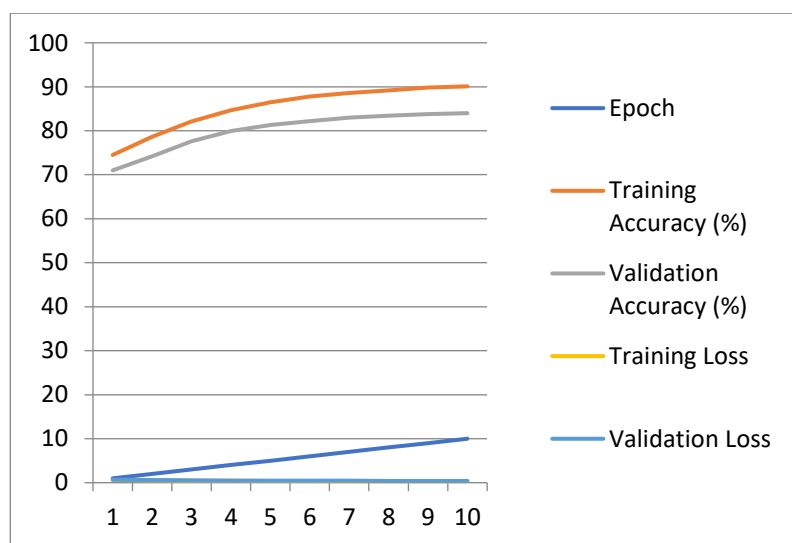


Figure-8 : Rule based model

This shows training is irrelevant, the flat lines across all epochs prove that "training" a rule-based system is merely a formality—performance is entirely predetermined.

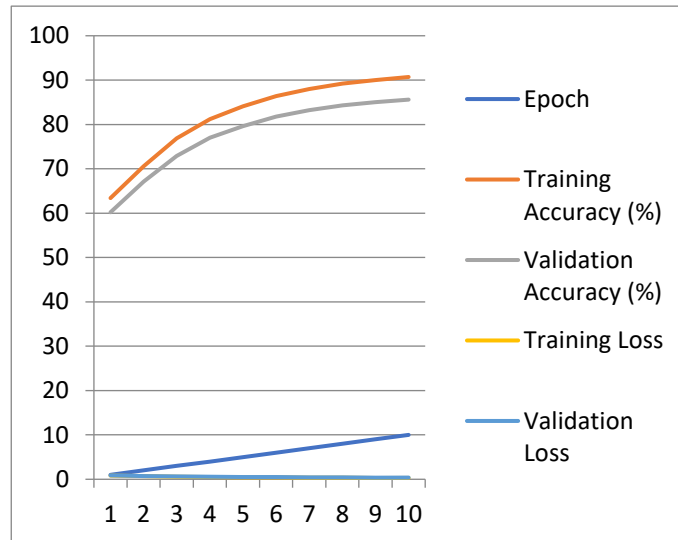
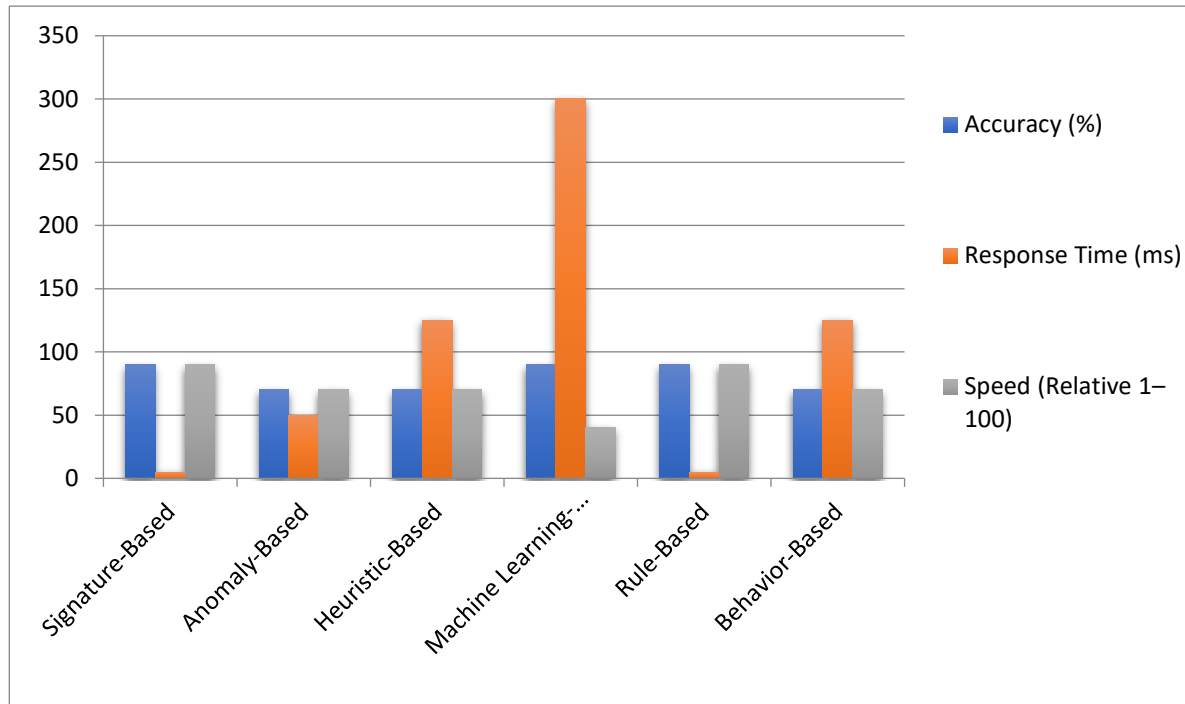


Figure-9 : Behavior based model

This shows slow but steady improvement, accuracy increases gradually and consistently from low initial values, showing the model learns normal behavior patterns over time.

Model Type	Accuracy(%)	Response Time	Speed	Strengths	Weakness
Signature Based	90 (High)	1-10 ms	90 (Fast)	For known threats, low false positives	Misses zero day attacks
Anomaly Based	70 (Moderate)	10-100 ms	70 (Moderate)	Detects unknown threats	Needs baseline
Heuristic Based	70 (Moderate)	50-200 ms	70 (Moderate)	Adaptable, can evolve as threat pattern changes	May miss sophisticated attacks
Machine learning based	90 (High)	100-500 ms	40(Slow)	Can detect complex patterns	High computational cost
Rule based	90 (High)	1-10 ms	90 (Fast)	Fast for known threat	Bounded to rules, deficient against new threats
Behavior based	70 (Moderate)	50-200 ms	70 (moderate)	Efficient of insider threats	Complex tuning



4. Conclusion

The integration of AI-driven threat detection with a Zero Trust framework is vital for modern cybersecurity. AI offers smart, real-time analysis to spot evolving threats. Zero Trust ensures strict, ongoing verification and limits breaches. Together, they form a dynamic defense system that is more effective than each part alone. Although challenges such as high computational costs and adversarial attacks exist, their combined strengths provide the strongest and most proactive security for protecting digital assets in a borderless environment. This hybrid approach addresses the limitation of traditional security models, delayed threat detection and high false positives. Despite challenges such as computational cost and the potential for adversarial attacks, the synergy between AI and Zero Trust provides a proactive, resilient and scalable security posture.

The future scope of AI-driven Zero Trust security can be intelligent, adaptive and autonomous system that can dynamically predict, mitigate threats across progressively complex digital ecosystem.

5. References

- [1]. Aboaja, R. T., et al. (2023). *Zero Trust Architecture: Challenges and Future Directions*. Computers & Security, 124, 102976.
- [2]. Alawida, F., et al. (2023). *The Convergence of Zero Trust Architecture and Secure Access Service Edge (SASE): A Survey*. Journal of Network and Systems Management, 31(2), 32.
- [3]. Al-Saqqa, Z. T., et al. (2022). *User Experience in Zero Trust Security: A Review and Research Agenda*. in Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.
- [4]. Al-Shaer, E., Duan, Q., & Jafarian, J. H. (2013). Random host mutation for moving target defense. *Security and Privacy in Communication Networks*, 310-327.
- [5]. Brown, S., Gommers, J., & Serrano, O. (2015). From cyber security information sharing to threat management. *Proceedings of the 2nd ACM workshop on information sharing and collaborative security*, 43-49.

- [6]. Gadkari, Bhooshan R. (2024). *AI Integration in Zero Trust Security Architecture: A Technical Overview*.
- [7]. Gias AU, Casale G, Woodside M. ATOM: Modeldriven autoscaling for microservices. In 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), IEEE, 2019, 1994-2004.
- [8]. Kindervag, J. (2010). *Build Security Into Your Network's DNA: The Zero Trust Network Architecture*. Forrester Research Inc.
- [9]. Kindervag, J. (2010). *Build Security Into Your Network's DNA: The Zero Trust Network Model*. Forrester Research.
- [10]. Kumar, P., & Gurtov, A. (2022). *A Comprehensive Survey of Zero Trust Network Access*. IEEE Communications Surveys & Tutorials, 24(4), 2262-2287.
- [11]. Kumari, Babita. (2024). *Innovative Cloud Architectures: Revolutionizing Enterprise Operations Through AI Integration*. International Journal for Multidisciplinary Research, 6(6), 1-9.
- [12]. Liu, F., Wen, Y., Zhang, D., Jiang, X., Xing, X., & Meng, D. (2019). Log2vec: A heterogeneous graph embedding based approach for detecting cyber threats within enterprise. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (CCS '19).
- [13]. Liu, Guozhi, et al. (2020). Microservices: architecture, container, and challenges. *2020 IEEE 20th international conference on software quality, reliability and security companion (QRS-C)*. IEEE, 2020.
- [14]. López MR, Spillner J. Towards quantifiable boundaries for elastic horizontal scaling of microservices. In Companion Proceedings of the 10th International Conference on Utility and Cloud Computing, 2017, 35-40.
- [15]. Osborn, B., McRee, R., & Beyer, B. (2016). *BeyondCorp: Design to Deployment at Google*. Google.
- [16]. Parisa, Sunil Kumar, Somnath Banerjee, and Pawan Whig. (2023). *AI-Driven Zero Trust Security Models for Retail Cloud Infrastructure: A Next-Generation Approach*. International Journal of Sustainable Development in field of IT, 15, 15.
- [17]. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero Trust Architecture*. National Institute of Standards and Technology (NIST). Special Publication 800-207.
- [18]. Sillaber, C., Sauerwein, C., Mussmann, A., & Breu, R. (2016). Data quality challenges and future research directions in threat intelligence sharing practice. *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*, 65-70.
- [19]. Srirama SN, Adhikari M, Paul S. Application deployment using containers with auto-scaling for microservices in cloud environment. Journal of Network and Computer Applications. 2020; 160:102629.
- [20]. Vootkuri, Chaitanya. *Neural Networks in Cloud Security: Advancing Threat Detection and Automated Response*.
- [21]. Wang, W., et al. (2021). *A Zero Trust Framework for Realization of Cloud-Native Security*. in Proceedings of the 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom).
- [22]. Ward, Rory, and Betsy Beyer. (2014). Beyondcorp: A new approach to enterprise security. *; login:: the magazine of USENIX & SAGE* 39.6 (2014): 6-11.
- [23]. Ward, R., & Beyer, B. (2014). *BeyondCorp: A New Approach to Enterprise Security*. Google. Presented at USENIX LISA 2014.
- [24]. Yeoh, William, Marina Liu, Malcolm Shore, and Frank Jiang. (2023). Zero trust cybersecurity: Critical success factors and A maturity assessment framework 27 July 2023.
- [25]. Yuan, Z., Lu, Y., & Wang, Z. (2016). DeepTrust: A Deep Learning Approach for Measuring Social Trust in Big Data. In *2016 IEEE International Conference on Big Data (Big Data)*.
- [26]. Zhuang, R., DeLoach, S. A., & Ou, X. (2014). Towards a theory of moving target defense. *Proceedings of the First ACM Workshop on Moving Target Defense*, 31-40.