

การวิเคราะห์และเปรียบเทียบอัลกอริธึมการเรียนรู้ ของเครื่องสำหรับการตรวจจับภัยคุกคามแบบ Zero-Day ในระบบเครือข่าย

Analysis and Comparison of Machine Learning Algorithms for Detecting Zero-Day Threats in Network Systems

อมร เจือดี¹
Amorn Juatee^{*1}

amorn@kru.ac.th*

ส่งบทความ 10 พฤศจิกายน 2567 แก้ไข 15 มกราคม 2568 ตอบรับ 21 มกราคม 2568
Received: November 10, 2024 Revised: January 15, 2025 Accepted: January 21, 2028

บทคัดย่อ

การวิจัยนี้มุ่งเน้นการวิเคราะห์และเปรียบเทียบอัลกอริธึม Machine Learning สำหรับการตรวจจับภัยคุกคามแบบ Zero-Day ในระบบเครือข่าย โดยมีวัตถุประสงค์ ได้แก่ 1) เพื่อศึกษาความสามารถของอัลกอริธึม Machine Learning ในการตรวจจับการโจมตีแบบ Zero-Day 2) เพื่อวิเคราะห์ประสิทธิภาพของอัลกอริธึม Machine Learning ในการแยกแยะและจำแนกพฤติกรรมที่มีแนวโน้มเป็นการโจมตี Zero-Day และ 3) เพื่อเสนอแนวทางการนำผลการวิเคราะห์ไปปรับใช้ในระบบรักษาความปลอดภัยเครือข่าย เพื่อตอบสนองต่อภัยคุกคาม Zero-Day ข้อมูลที่ใช้รวบรวมจาก Threat Intelligence และข้อมูลจริงจากระบบป้องกันเครือข่าย

ผลการวิจัยพบว่า Neural Networks (NN) มีความแม่นยำ (Accuracy) สูงถึง 95.2% และอัตราการแจ้งเตือนผิดพลาด (False Positive Rate: FPR) ต่ำเพียง 2.5% ความสามารถนี้สะท้อนถึงประสิทธิภาพในการเรียนรู้ข้อมูลใหม่ที่ไม่เคยพบมาก่อน และการตอบสนองต่อภัยคุกคาม Zero-Day ได้อย่างแม่นยำ โดยช่วยลดภาระในการจัดการแจ้งเตือนผิดพลาด และเพิ่มประสิทธิภาพในการป้องกันระบบเครือข่าย RF และ SVM ซึ่งมีความแม่นยำ 90.5% และ 88.7% ตามลำดับ แม้จะให้ผลลัพธ์ที่ดี แต่ยังคงน้อยกว่า NN ในการลด FPR การใช้เทคนิค Anomaly Detection และ Ensemble Models ยังช่วยเพิ่มศักยภาพในการตรวจจับภัยคุกคามที่ซับซ้อนและรองรับสภาพแวดล้อมที่เปลี่ยนแปลงได้ ระบบที่พัฒนาขึ้นสามารถบล็อก IP ที่เคยโจมตีซ้ำและตรวจจับพฤติกรรมผิดปกติได้รวดเร็ว เพิ่มความสามารถในการป้องกันภัยคุกคามใหม่ ๆ ที่ระบบเดิมไม่สามารถจัดการได้ การวิจัยนี้เสนอแนวทางพัฒนาระบบรักษาความปลอดภัยให้มีความแม่นยำและยืดหยุ่นต่อภัยคุกคามที่ซับซ้อนและเปลี่ยนแปลงได้ในอนาคต พร้อมเน้นย้ำถึงความสำคัญของการพัฒนาเทคโนโลยี Machine Learning ในการเสริมสร้างความมั่นคงทางไซเบอร์

คำสำคัญ: การเรียนรู้ของเครื่อง, การโจมตีแบบ Zero-Day, การตรวจจับพฤติกรรมผิดปกติ, ความปลอดภัยเครือข่าย, การวิเคราะห์ความผิดปกติ

*ผู้ประพันธ์สรรพสิทธิ์ (corresponding author)

¹ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี

¹ Faculty of Science and Technology, Kanchanaburi Rajabhat University

Abstract

This research focuses on analyzing and comparing Machine Learning algorithms for detecting Zero-Day threats in network systems. The objectives include: 1) studying the capability of Machine Learning algorithms to detect Zero-Day attacks, 2) analyzing the performance of these algorithms in differentiating and classifying behaviors indicative of Zero-Day attacks, and 3) proposing practical approaches for integrating analytical results into network security systems to counteract Zero-Day threats. Data utilized in this study were collected from Threat Intelligence and actual network defense systems.

The findings reveal that Neural Networks (NN) achieved an accuracy of 95.2% with a False Positive Rate (FPR) of just 2.5%, demonstrating superior performance in learning new data and accurately responding to Zero-Day threats. This capability significantly reduces the burden of handling false alarms and enhances network protection. Random Forest (RF) and Support Vector Machine (SVM) achieved accuracies of 90.5% and 88.7%, respectively, but were less effective than NN in minimizing FPR. Additionally, the use of Anomaly Detection and Ensemble Models further strengthened the ability to detect complex threats and adapt to dynamic environments. The developed system effectively blocks repeated attack IPs and promptly detects anomalous behavior, offering advanced protection against emerging threats that traditional systems cannot address. This research provides practical recommendations for developing network security systems with enhanced precision, flexibility, and resilience to sophisticated and evolving threats. Moreover, it highlights the crucial role of Machine Learning technologies in bolstering cybersecurity, serving as an essential tool for preventing and managing threats in the digital age.

Keywords: Machine Learning, Zero-Day Attack, Anomaly Detection, Network Security, Behavioral Analysis

.....

บทนำ

การโจมตีทางไซเบอร์ในปัจจุบันมีความซับซ้อนและมีการพัฒนารูปแบบอย่างต่อเนื่อง โดยเฉพาะ Zero-Day Attacks ซึ่งเป็นภัยคุกคามที่รุนแรงและยากต่อการตรวจจับเนื่องจากอาศัยช่องโหว่ในระบบที่ยังไม่ได้รับการค้นพบหรือพัฒนาแพตช์ป้องกัน ความสามารถของระบบรักษาความปลอดภัยแบบดั้งเดิมที่อิงฐานข้อมูลลายเซ็น (Signature-Based Detection) ไม่เพียงพอที่จะรับมือกับการโจมตีลักษณะนี้ ทำให้ผู้โจมตีสามารถเข้าถึงระบบโดยไม่ถูกตรวจพบ นำไปสู่ความเสี่ยงที่ร้ายแรงต่อความมั่นคงของระบบ การพัฒนาวิธีการตรวจจับที่สามารถระบุพฤติกรรมแปลกใหม่และภัยคุกคามที่ไม่เคยปรากฏมาก่อน จึงเป็นสิ่งจำเป็น โดยเฉพาะการนำเทคโนโลยี Machine Learning (ML) มาใช้ ซึ่งได้รับการยอมรับว่าเป็นเครื่องมือที่มีศักยภาพสูงในการตรวจจับและป้องกันภัยคุกคามไซเบอร์แบบ Zero-Day การประยุกต์ใช้ Machine Learning (ML) ช่วยให้ระบบสามารถเรียนรู้พฤติกรรมที่ผิดปกติจากข้อมูลจำนวนมากในระบบเครือข่าย และจำแนกความเสี่ยงหรือภัยคุกคามใหม่ได้อย่างมีประสิทธิภาพ การตรวจจับที่อิงพฤติกรรม (Anomaly Detection) ช่วยให้โมเดล ML ปรับตัวได้อย่างรวดเร็วและตอบสนองต่อภัยคุกคามใหม่ได้อย่างเหมาะสม ต่างจากวิธีการดั้งเดิม เช่น Signature-Based Detection ที่อิงข้อมูลการโจมตีในอดีต ซึ่งไม่สามารถรองรับภัยคุกคามที่เกิดจากช่องโหว่ใหม่ ๆ ข้อได้เปรียบของ ML คือความสามารถในการวิเคราะห์พฤติกรรมที่ผิดปกติและปรับตัวต่อรูปแบบการโจมตีใหม่ ๆ ได้อย่างต่อเนื่อง อีกทั้งยังช่วยลดความซับซ้อนในการจัดการข้อมูลขนาดใหญ่ในระบบเครือข่าย ด้วยการใช้เทคนิค Anomaly Detection เพื่อเพิ่มความแม่นยำในการระบุภัยคุกคาม (Sewak et al., 2022)

บทความนี้มุ่งเน้นการวิเคราะห์และเปรียบเทียบอัลกอริธึม Machine Learning ที่เหมาะสมสำหรับการตรวจจับภัยคุกคามแบบ Zero-Day ในระบบเครือข่าย โดยพิจารณาแนวทางการแก้ปัญหาและผลลัพธ์ที่ได้จากข้อมูลพฤติกรรมโจมตีซึ่งรวบรวมจากระบบ Next-Generation Application Firewall (NGAF) ที่ติดตั้งในมหาวิทยาลัยราชภัฏกาญจนบุรี และข้อมูลจาก Threat Intelligence เพื่อเพิ่มความครอบคลุม ข้อมูลเพิ่มเติม

เช่น ตัวอย่างการโจมตีจากงานวิจัยของ Salem et al. (2024) ถูกนำมาใช้เพื่อเปรียบเทียบผลลัพธ์ในบริบทต่าง ๆ การศึกษานี้มีเป้าหมายเพื่อพัฒนาระบบรักษาความปลอดภัยที่มีความแม่นยำและตอบสนองต่อภัยคุกคามที่เปลี่ยนแปลงได้อย่างมีประสิทธิภาพ รองรับความต้องการด้านความมั่นคงในเครือข่ายองค์กรอย่างเหมาะสมที่สุด

วัตถุประสงค์การวิจัย

1. เพื่อศึกษาความสามารถของอัลกอริธึม Machine Learning ในการตรวจจับการโจมตีแบบ Zero-Day โดยอ้างอิงจากข้อมูลพฤติกรรมที่ผิดปกติในระบบเครือข่าย
2. เพื่อวิเคราะห์และเปรียบเทียบประสิทธิภาพของอัลกอริธึม Machine Learning ในการตรวจจับ Zero-Day Attacks
3. เพื่อเสนอแนวทางการนำผลการวิเคราะห์ไปปรับใช้ในระบบรักษาความปลอดภัยเครือข่าย เพื่อตอบสนองต่อภัยคุกคาม Zero-Day

ขอบเขตของการวิจัย

การวิจัยนี้มีเป้าหมายเพื่อวิเคราะห์และเปรียบเทียบอัลกอริธึมการเรียนรู้ของเครื่อง (Machine Learning Algorithms) สำหรับการตรวจจับภัยคุกคามแบบ Zero-Day ในระบบเครือข่าย โดยมีขอบเขตดังนี้ ดังนี้

1. ศึกษาลักษณะและพฤติกรรมของภัยคุกคาม Zero-Day ซึ่งเป็นภัยคุกคามที่ไม่สามารถป้องกันล่วงหน้าได้เนื่องจากไม่มีข้อมูลเกี่ยวกับการโจมตีดังกล่าวในฐานข้อมูลเดิมของระบบรักษาความปลอดภัย
2. นำข้อมูลที่ได้จากระบบรักษาความปลอดภัยในองค์กร เช่น บันทึกเหตุการณ์และพฤติกรรมของระบบที่เกี่ยวข้อง มาใช้ในการตรวจสอบและวิเคราะห์ลักษณะการโจมตี Zero-Day
3. ระบุช่องโหว่หรือจุดอ่อนในระบบเครือข่ายที่อาจตกเป็นเป้าหมายของการโจมตี Zero-Day เพื่อตรวจสอบและป้องกันปัญหาที่อาจเกิดขึ้น
4. รวมข้อมูลจากการโจมตีประเภทต่าง ๆ เช่น SQL Injection, OS Command Injection, WebShell Upload, และ Protocol Anomaly เพื่อนำไปใช้ในการวิเคราะห์

และประเมินประสิทธิภาพของอัลกอริธึม

5. ศึกษาผลกระทบที่เกิดจากการโจมตี Zero-Day ต่อระบบขององค์กร พร้อมทั้งเสนอแนวทางการพัฒนาและปรับปรุงระบบรักษาความปลอดภัย เพื่อเพิ่มความสามารถในการป้องกันและรับมือกับภัยคุกคามในอนาคตอย่างมีประสิทธิภาพ

สมมติฐานของงานวิจัย

งานวิจัย การวิเคราะห์และเปรียบเทียบอัลกอริธึมการเรียนรู้ของเครื่องสำหรับการตรวจจับภัยคุกคามแบบ Zero-Day ในระบบเครือข่าย โดยสมมติฐาน ดังนี้

1. อัลกอริธึม Machine Learning ตรวจจับ Zero-Day ได้อย่างแม่นยำในระบบเครือข่ายที่มีพฤติกรรมผิดปกติและภัยคุกคามใหม่
2. Neural Networks มีความแม่นยำและประสิทธิภาพสูงสุดเมื่อเทียบกับ Random Forest และ SVM
3. การใช้เทคนิค Anomaly Detection และ Ensemble Learning เพิ่มความแม่นยำและยืดหยุ่นในการจัดการ Zero-Day

วิธีดำเนินการวิจัย

การวิจัยนี้มุ่งเน้นการวิเคราะห์และเปรียบเทียบอัลกอริธึมการเรียนรู้ของเครื่อง (Machine Learning Algorithms) สำหรับการตรวจจับภัยคุกคามแบบ Zero-Day ในระบบเครือข่าย โดยมีการดำเนินการใน 5 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การรวบรวมข้อมูล ผู้วิจัยได้รวบรวมข้อมูลด้วยการใช้ระบบ Next-Generation Application Firewall (NGAF) เพื่อบันทึกพฤติกรรมการโจมตี เช่น SQL Injection, WebShell Upload และ OS Command Injection พร้อมเสริมด้วยข้อมูลจาก Threat Intelligence เพื่อเพิ่มความครอบคลุม ข้อมูลเหล่านี้รวบรวมระหว่างเดือนมกราคมถึงมิถุนายน 2024 และเหตุการณ์การโจมตีจริงในมหาวิทยาลัยราชภัฏกาญจนบุรี ถูกนำมาใช้เพื่อสร้างความหลากหลายและความสมจริงให้การทดลอง กระบวนการดำเนินการทั้งหมดเป็นไปตามมาตรฐานความปลอดภัยเพื่อรักษาคุณภาพและความถูกต้องของข้อมูล (Babu R. Dawadi et al., 2023)

ขั้นตอนที่ 2 การเตรียมข้อมูล ได้ดำเนินการกำจัดข้อมูลซ้ำซ้อน (Data Deduplication) และปรับขนาดข้อมูล (Normalization) เพื่อให้ข้อมูลมีความเหมาะสมสำหรับการประมวลผลในอัลกอริธึม Machine Learning ข้อมูลทั้งหมดถูกจัดหมวดหมู่ (Labeling) โดยแยกประเภทระหว่างข้อมูลโจมตีและข้อมูลปกติ รวมถึงจำแนกประเภทของการโจมตี เช่น SQL Injection และ WebShell Upload จากนั้นข้อมูลได้รับการแบ่งเป็นชุดฝึก (Training Set) และชุดทดสอบ (Test Set) ด้วยเทคนิค Stratified Sampling เพื่อรักษาสัดส่วนของข้อมูลในแต่ละกลุ่มให้สมดุล นอกจากนี้ ยังมีการตรวจสอบคุณภาพข้อมูลอย่างละเอียดเพื่อลดอคติหรือความไม่สมบูรณ์ที่อาจส่งผลต่อประสิทธิภาพของโมเดล (Singh et al., 2023)

ขั้นตอนที่ 3 การเลือกและพัฒนาอัลกอริธึม การเลือกอัลกอริธึม Machine Learning ในงานวิจัยนี้ใช้อัลกอริธึม 5 ประเภท ได้แก่ Neural Networks (NN), Random Forest (RF), Support Vector Machines (SVM), Isolation Forest และ Autoencoders โดยอัลกอริธึมแต่ละประเภทมีคุณสมบัติที่เหมาะสมกับการจัดการข้อมูลในบริบทที่แตกต่างกัน ผู้วิจัยดำเนินการปรับค่าพารามิเตอร์ (Hyperparameter Tuning) และใช้ชุดข้อมูล Validation เพื่อประเมินผลระหว่างการฝึกโมเดลเพื่อลดความเสี่ยงจาก Overfitting และเพิ่มประสิทธิภาพในการตรวจจับและจัดการข้อมูล (Sewak et al., 2022)

ขั้นตอนที่ 4 การทดสอบและประเมินผล โมเดลแต่ละตัวถูกทดสอบด้วยชุดข้อมูล Test Set โดยใช้ตัวชี้วัด เช่น Accuracy, Precision, Recall, และ F1-Score เพื่อประเมินความถูกต้องและสมดุลของโมเดล พร้อมวิเคราะห์ False Positive Rate (FPR) และ False Negative Rate (FNR) เพื่อระบุข้อดีข้อเสียในบริบทของ Zero-Day Detection การทดสอบนี้ช่วยระบุศักยภาพและปรับปรุงโมเดลให้เหมาะสมกับการใช้งานในสภาพแวดล้อมที่ซับซ้อนในอนาคต (Singh et al., 2023)

ขั้นตอนที่ 5 การวิเคราะห์ผลการวิจัย ผลการทดสอบโมเดลถูกนำมาวิเคราะห์เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริธึมในด้านต่าง ๆ โดยพิจารณาความแม่นยำ ความสามารถในการจัดการข้อมูลหลากหลาย

และข้อดีข้อจำกัดในบริบทที่แตกต่าง ผู้วิจัยได้เสนอแนวทางพัฒนาต่อ เช่น การใช้ Ensemble Learning เพื่อผสานจุดเด่นของหลายอัลกอริธึม และการพัฒนา Hybrid Algorithms เพื่อเพิ่มความยืดหยุ่น ลดข้อผิดพลาด และปรับปรุงความสามารถในการเรียนรู้ข้อมูลใหม่ในสภาพแวดล้อมที่เปลี่ยนแปลงได้อย่างต่อเนื่อง (Chua & Salam, 2023)

สถิติที่ใช้ในการวิจัย

การวิจัยนี้ใช้สถิติเพื่อประเมินความแม่นยำและประสิทธิภาพของอัลกอริธึม Machine Learning ในการตรวจจับภัยคุกคาม Zero-Day สถิติที่ใช้ประกอบด้วย

Accuracy ใช้ในการประเมินความถูกต้องของโมเดลโดยรวม คำนวณจากอัตราส่วนระหว่างจำนวนตัวอย่างที่ถูกต้อง (True Positives และ True Negatives) กับจำนวนตัวอย่างทั้งหมดที่มีในชุดข้อมูล สูตรการคำนวณ ดังนี้

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP หมายถึง True Positives

TN หมายถึง True Negatives

FP หมายถึง False Positives

FN หมายถึง False Negatives

Precision ใช้วัดความถูกต้องของผลลัพธ์ที่เป็นบวกในโมเดล Machine Learning โดยการคำนวณจากจำนวนตัวอย่างที่โมเดลระบุว่าเป็นบวกและถูกต้องจริง (True Positives, TP) เมื่อเปรียบเทียบกับจำนวนตัวอย่างทั้งหมดที่โมเดลระบุว่าเป็นบวก ซึ่งรวมถึงตัวอย่างที่อาจจะไม่ถูกต้อง (False Positives, FP) สูตรการคำนวณ ดังนี้

$$\text{Precision} = \frac{TP}{TP + FN}$$

Recall ใช้ในการวัดความสามารถของโมเดลในการตรวจจับตัวอย่างที่เป็นบวกทั้งหมดในข้อมูลจริงโดยคำนวณจากจำนวนตัวอย่างที่โมเดลทำนายว่าเป็นบวกและถูกต้องจริง (True Positives, TP) เมื่อเปรียบเทียบกับจำนวนตัวอย่างที่เป็นบวกทั้งหมดในข้อมูลจริง ซึ่งรวมถึงตัวอย่างที่โมเดล

อาจไม่สามารถตรวจจับได้ (False Negatives, FN) สูตรการคำนวณ ดังนี้

$$\text{Precision} = \frac{TP}{TP + FN}$$

F1-Score ใช้ในการวัดประสิทธิภาพโดยรวมของโมเดลในด้านความแม่นยำ (Precision) และความสามารถในการเรียกคืน (Recall) โดยคำนวณจากค่าเฉลี่ยแบบถ่วงน้ำหนักของทั้งสอง ซึ่งช่วยให้เห็นความสมดุลระหว่าง Precision และ Recall โดยเฉพาะในกรณีที่ข้อมูลไม่สมดุล (imbalanced data) สูตรการคำนวณ ดังนี้

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

True Positive Rate (TPR) ใช้ในการวัดความสามารถของโมเดลในการตรวจจับตัวอย่างที่เป็นบวกในกลุ่มตัวอย่างทั้งหมด โดย TPR คำนวณจากจำนวนตัวอย่างที่โมเดลทำนายว่าเป็นบวกและถูกต้องจริง (True Positives, TP) เมื่อเปรียบเทียบกับจำนวนตัวอย่างที่เป็นบวกทั้งหมดในข้อมูลจริง ซึ่งรวมถึงตัวอย่างที่โมเดลอาจไม่สามารถตรวจจับได้ (False Negatives, FN) สูตรการคำนวณ ดังนี้

$$\text{TPR} = \frac{TP}{TP + FN}$$

False Positive Rate (FPR) ใช้ในการวัดความสามารถของโมเดลในการตรวจจับตัวอย่างที่ผิดพลาดว่าเป็นบวก โดยคำนวณจากจำนวนตัวอย่างที่โมเดลทำนายว่าเป็นบวก แต่จริง ๆ แล้วเป็นลบ (False Positives, FP) เมื่อเปรียบเทียบกับจำนวนตัวอย่างที่เป็นลบทั้งหมดในข้อมูลจริง ซึ่งรวมถึงตัวอย่างที่โมเดลทำนายว่าเป็นลบและถูกต้อง (True Negatives, TN) สูตรการคำนวณ ดังนี้

$$\text{FPR} = \frac{FP}{FP + TN}$$

ROC Curve แสดงความสัมพันธ์ระหว่าง True Positive Rate (TPR) และ False Positive Rate (FPR) ของโมเดล Machine Learning สำหรับค่า threshold ที่แตกต่างกัน โดย TPR แสดงบนแกน y และ FPR บนแกน x กราฟนี้ช่วยประเมินประสิทธิภาพของโมเดลในการแยกแยะข้อมูลบวกและลบ ค่า AUC ที่ใกล้ 1 แสดงถึงโมเดลที่มีประสิทธิภาพสูง ขณะที่ AUC ใกล้ 0.5 แสดงถึงโมเดลที่มีประสิทธิภาพไม่ดี

Area Under the Curve (AUC) คือการวัดประสิทธิภาพของโมเดล Machine Learning ในการจำแนกประเภท โดยคำนวณจากพื้นที่ใต้กราฟ ROC Curve

ซึ่งแสดงความสัมพันธ์ระหว่าง True Positive Rate (TPR) และ False Positive Rate (FPR) สำหรับค่า threshold ต่าง ๆ ค่า AUC บ่งบอกถึงความสามารถของโมเดลในการแยกประเภทข้อมูล โดยค่าของ AUC สามารถตีความได้ดังนี้

- 1 แยกประเภทได้สมบูรณ์แบบ
- 0.9 - 0.99 ประสิทธิภาพดีมาก
- 0.8 - 0.89 ประสิทธิภาพดีพอสมควร
- 0.7 - 0.79 ประสิทธิภาพปานกลาง
- 0.6 - 0.69 ประสิทธิภาพต่ำ
- 0.5 ไม่มีความสามารถในการแยก

ผลการวิจัย

ผลการวิจัยแสดงให้เห็นว่า Zero-Day Attack เป็นการโจมตีที่อาศัยช่องโหว่ในซอฟต์แวร์หรือระบบปฏิบัติการที่ยังไม่มีแพตช์ป้องกัน ส่งผลให้ผู้โจมตีสร้างความเสียหายได้อย่างมีประสิทธิภาพ โดยรายงานของ Gupta et al. (2024) ระบุว่ามีการโจมตีถึง 187,681 ครั้ง และมีแนวโน้มเพิ่มขึ้น วิธีป้องกันแบบดั้งเดิม เช่น Signature-Based Detection ไม่สามารถรับมือได้ เนื่องจากขาดข้อมูลล่วงหน้า การใช้ Machine Learning เช่น Neural Networks และ Random Forest จึงมีบทบาทสำคัญในการตรวจจับพฤติกรรมผิดปกติและเพิ่มประสิทธิภาพในการป้องกันภัยคุกคามดังกล่าว (Yanguema, 2023)

การใช้เทคนิค Anomaly Detection ช่วยตรวจจับภัยคุกคามที่ไม่เคยปรากฏมาก่อนผ่านการสร้างรูปแบบพฤติกรรมปกติและแจ้งเตือนเมื่อพบความเบี่ยงเบน การประยุกต์ใช้ Machine Learning (ML) เพิ่มความแม่นยำในการตรวจจับ โดยเรียนรู้จากเหตุการณ์ในอดีตและประมวลผลแบบเรียลไทม์ ช่วยให้ระบบตอบสนองต่อภัยคุกคามที่เปลี่ยนแปลงได้รวดเร็ว นอกจากนี้ Ensemble Learning ที่ผสมผสาน Neural Networks และ Random Forest ยังช่วยเพิ่มความครอบคลุมในการตรวจจับภัยคุกคามที่หลากหลาย

การรองรับ Continuous Learning ในระบบ Machine Learning ช่วยให้ปรับตัวตามข้อมูลใหม่ได้ต่อเนื่อง และ Distributed Processing ช่วยเพิ่มความเร็วในการประมวลผลข้อมูลขนาดใหญ่ ลดข้อจำกัดด้านทรัพยากร ระบบที่ยืดหยุ่นและวิเคราะห์แบบเรียลไทม์จึงเหมาะกับภัยคุกคามไซเบอร์ที่ซับซ้อนในปัจจุบันและอนาคต (Sarhan et al., 2023)

ตารางที่ 1 การวิเคราะห์ประสิทธิภาพของอัลกอริธึม Machine Learning สำหรับตรวจจับ Zero-Day Attack

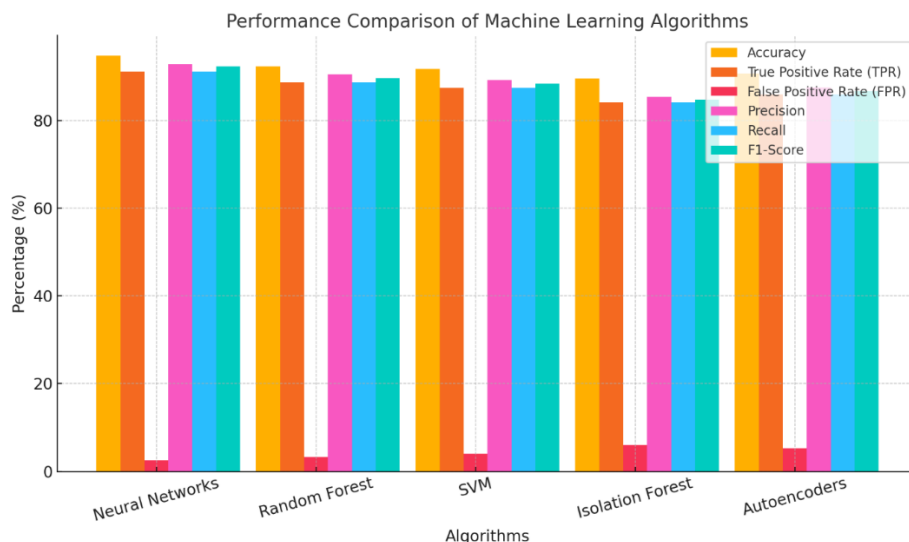
อัลกอริธึม	Accuracy (%)	True Positive Rate (TPR) (%)	False Positive Rate (FPR) (%)	Precision (%)	Recall (%)	F1-Score (%)
Neural Networks	95.2	92.4	2.5	93.8	92.4	93.1
Random Forest	93.5	89.8	3.1	91.6	89.8	90.7
Support Vector Machine (SVM)	92.1	88.3	3.8	90.0	88.3	89.1
Isolation Forest	90.0	85.4	5.8	86.9	85.4	86.1
Autoencoders	91.4	87.2	4.9	88.7	87.2	87.9

ตารางที่ 1 สรุปผลการวิเคราะห์อัลกอริธึม Machine Learning 5 รูปแบบ ได้แก่ Neural Networks, Random Forest, SVM, Isolation Forest และ Autoencoders สำหรับตรวจจับ Zero-Day Attack ตัวชี้วัดสำคัญที่ใช้ประเมิน ได้แก่ Accuracy, TPR, FPR, Precision, Recall และ F1-Score จากผลการวิเคราะห์ Neural Networks มีประสิทธิภาพสูงสุด (Accuracy 95.2%, TPR 92.4%, FPR 2.5%, F1-Score 93.1%) แสดงถึงความแม่นยำและการลดข้อผิดพลาด Random Forest (Accuracy 93.5%, Precision 91.6%) และ SVM (Accuracy 92.1%, F1-Score 89.1%) มีประสิทธิภาพใกล้เคียงกัน เหมาะกับการใช้งานที่ต้องการความสมดุล ส่วน Isolation Forest และ Autoencoders แม้มี Accuracy ต่ำกว่า (90.0% และ 91.4%) และ FPR สูงกว่า แต่มีจุดเด่นด้านความยืดหยุ่น เหมาะกับข้อมูลที่เปลี่ยนแปลงได้ง่าย สรุปได้ว่า Neural Networks เหมาะสำหรับการตรวจจับที่ต้องการความแม่นยำสูง ส่วน Random Forest และ SVM ให้สมดุลดีในหลายตัวชี้วัด ขณะที่ Isolation Forest และ Autoencoders เหมาะกับระบบที่เน้นความยืดหยุ่น (Dean, 2022)

ตารางที่ 2 การเปรียบเทียบประสิทธิภาพอัลกอริธึม Machine Learning สำหรับตรวจจับ Zero-Day Attack

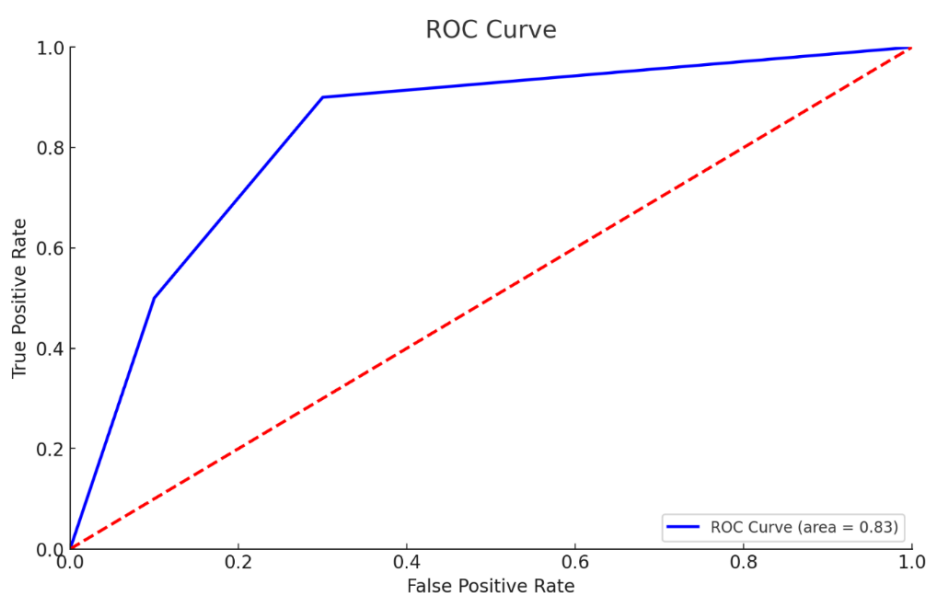
อัลกอริธึม	จุดเด่น	ข้อจำกัด	เหมาะสมกับ
Neural Networks	- ความแม่นยำสูง (Accuracy 95.2%) - FPR ต่ำสุด (2.5%)	- ใช้ทรัพยากรสูง - ต้องการข้อมูลปริมาณมาก	ระบบองค์กรขนาดใหญ่ ที่ต้องการความแม่นยำสูง
Random Forest	- สมดุลในหลายตัวชี้วัด - ทนต่อข้อมูล Noise	- ความแม่นยำน้อยกว่า Neural Networks	ระบบที่มีข้อมูลหลากหลาย
Support Vector Machine (SVM)	- เหมาะกับข้อมูลซับซ้อน - F1-Score สมดุล	- FPR สูงกว่า Neural Networks (3.8%)	ระบบขนาดเล็กที่มีข้อจำกัดด้านทรัพยากร
Isolation Forest	- ไม่ต้องการข้อมูลติดป้าย - ตรวจจับพฤติกรรมผิดปกติ	- Accuracy ต่ำสุด (90.0%) - FPR สูง (5.8%)	ระบบที่มีข้อจำกัดด้านทรัพยากร
Autoencoders	- เหมาะกับข้อมูลซับซ้อน - Precision สูง	- ต้องการการปรับแต่งซับซ้อน	ระบบที่เน้นการเรียนรู้รูปแบบพฤติกรรมในข้อมูลใหญ่

ตารางที่ 2 สรุปประสิทธิภาพอัลกอริธึม Machine Learning สำหรับตรวจจับ Zero-Day Attack โดย Neural Networks แสดงผลลัพธ์ที่ดีที่สุด ด้วย Accuracy 95.2% และ FPR ต่ำสุด (2.5%) เหมาะสำหรับระบบที่ต้องการความแม่นยำสูงและการตอบสนองแบบเรียลไทม์ แม้ต้องใช้ทรัพยากรมาก Random Forest มีความสมดุลในตัวชี้วัด เช่น Precision 91.6% และ Recall 89.8% จึงเหมาะกับข้อมูลที่หลากหลายและซับซ้อนในระดับปานกลาง ส่วน SVM มี F1-Score 89.1% และเหมาะกับข้อมูลที่ซับซ้อนหรือระบบที่มีข้อจำกัดด้านทรัพยากร Isolation Forest และ Autoencoders แม้มี Accuracy ต่ำกว่า (90.0% และ 91.4%) และ FPR สูงกว่า (5.8%) แต่โดดเด่นในความสามารถจัดการข้อมูลที่ไม่ได้ติดป้ายล่วงหน้าในสภาพแวดล้อมที่ต้องการความยืดหยุ่น การพัฒนา Ensemble Learning ที่รวมจุดเด่นของ Neural Networks และ Random Forest ช่วยเพิ่มความแม่นยำและความครอบคลุมของระบบ อีกทั้งการรองรับ Continuous Learning ยังช่วยให้ระบบปรับตัวกับข้อมูลใหม่ได้อย่างรวดเร็วและมีประสิทธิภาพ การเลือกอัลกอริธึมควรพิจารณาลักษณะข้อมูลและข้อกำหนดของระบบโดย Neural Networks เหมาะกับงานที่ต้องการความแม่นยำสูง Random Forest และ SVM เหมาะกับความสมดุล ส่วน Isolation Forest และ Autoencoders ตอบโจทย์สภาพแวดล้อมที่ต้องการความยืดหยุ่น



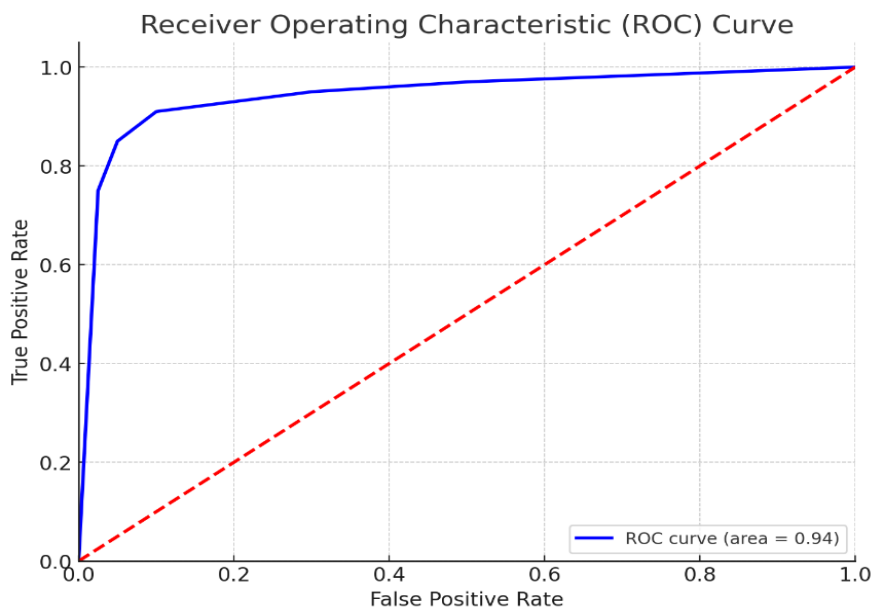
ภาพที่ 1 การเปรียบเทียบประสิทธิภาพของอัลกอริธึมการเรียนรู้ของเครื่องในการตรวจจับภัยคุกคามแบบ Zero-Day

ภาพที่ 1 แสดงการเปรียบเทียบประสิทธิภาพของอัลกอริธึม Machine Learning ในการตรวจจับภัยคุกคามแบบ Zero-Day โดยใช้ตัวชี้วัดสำคัญ เช่น Accuracy, TPR, FPR, Precision, Recall และ F1-Score ผลการวิเคราะห์พบว่า Neural Networks มีประสิทธิภาพสูงสุด ด้วย Accuracy 94.7% และ FPR ต่ำสุดที่ 2.5% เหมาะสำหรับระบบที่ต้องการการตอบสนองแบบเรียลไทม์ Random Forest มีผลลัพธ์ที่สมดุลในหลายตัวชี้วัด เช่น Precision 90.5% และ Recall 88.7% เหมาะกับข้อมูลหลากหลายและไม่สมดุล SVM แสดงผลลัพธ์เด่นในข้อมูลซับซ้อน ด้วย F1-Score 88.3% เหมาะสำหรับระบบที่มีทรัพยากรจำกัด ขณะที่ Isolation Forest และ Autoencoders แม้ Accuracy ต่ำกว่า แต่เด่นในการจัดการข้อมูลที่ไม่ได้ติดป้ายล่วงหน้า โดย Isolation Forest เหมาะสำหรับระบบที่มีทรัพยากรจำกัด และ Autoencoders เหมาะกับข้อมูลซับซ้อนที่ต้องการเรียนรู้ (Salem et al., 2024)



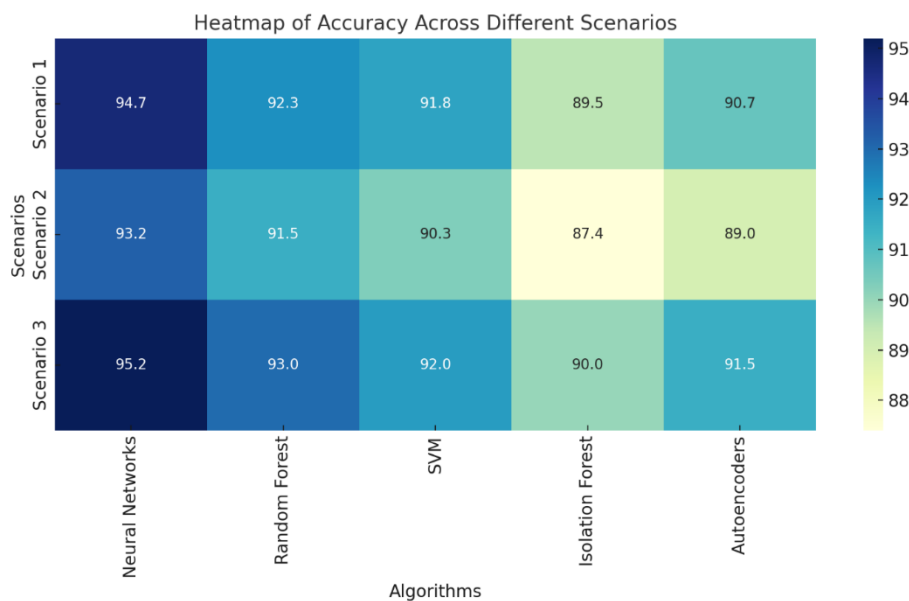
ภาพที่ 2 ROC Curve (Receiver Operating Characteristic Curve)

ภาพที่ 2 แสดง ROC Curve ที่ใช้ประเมินประสิทธิภาพของอัลกอริธึม Machine Learning ในการตรวจจับการโจมตีแบบ Zero-Day โดยกราฟแสดงความสัมพันธ์ระหว่างอัตราการตรวจจับที่ถูกต้อง (TPR) และอัตราการทำนายผิดพลาด (FPR) ค่า AUC (Area Under the Curve) เท่ากับ 0.83 บ่งบอกถึงประสิทธิภาพที่ดีในการจำแนกข้อมูลระหว่างภัยคุกคามและข้อมูลปกติ Neural Networks มีค่า TPR สูงสุดที่ 91.1% และ FPR ต่ำสุดที่ 2.5% แสดงถึงความแม่นยำและความสามารถในการลดการแจ้งเตือนผิดพลาด ส่วน Isolation Forest มีค่า FPR สูงถึง 6.0% สะท้อนถึงข้อจำกัดในการแยกแยะข้อมูล การวิเคราะห์นี้ชี้ให้เห็นว่า Neural Networks เหมาะสำหรับงานที่ต้องการความแม่นยำสูง ในขณะที่ Isolation Forest เหมาะสำหรับระบบที่ทรัพยากรจำกัด การเลือกใช้อัลกอริธึมควรพิจารณาจากความต้องการเฉพาะและข้อจำกัดของระบบเพื่อให้เหมาะสมที่สุด (Siddiqi Prity et al., 2024)



ภาพที่ 3 การเปรียบเทียบ ROC Curve และค่า AUC ของอัลกอริธึมตรวจจับ Zero-Day Attack

ภาพที่ 3 แสดง ROC Curve ที่ใช้ประเมินประสิทธิภาพของอัลกอริธึม Machine Learning ในการตรวจจับ Zero-Day Attack โดยค่า AUC (Area Under the Curve) บ่งบอกถึงความสามารถในการแยกแยะข้อมูลระหว่างภัยคุกคามและข้อมูลปกติ Neural Networks มีค่า AUC สูงสุดที่ 0.94 พร้อม TPR 91.1% และ FPR ต่ำสุด 2.5% สะท้อนถึงประสิทธิภาพในการตรวจจับที่แม่นยำและลดการแจ้งเตือนผิดพลาด เหมาะสำหรับระบบที่ต้องการ การตอบสนองแบบเรียลไทม์ Random Forest มีค่า AUC รองลงมาที่ 0.90 ซึ่งสมดุลในหลายตัวชี้วัด เช่น TPR และ Precision เหมาะกับข้อมูลที่หลากหลายและไม่สมดุล ขณะที่ SVM และ Isolation Forest มีค่า AUC 89.0% และ 86.5% ตามลำดับ แม้จะต่ำกว่า แต่เหมาะสำหรับงานที่มีข้อจำกัดด้านทรัพยากรหรือข้อมูลเฉพาะ โดยเฉพาะ Isolation Forest ซึ่งสามารถจัดการข้อมูลที่ไม่ได้ติดป้ายล่วงหน้าได้ดี การเปรียบเทียบนี้แสดงให้เห็นว่า Neural Networks และ Random Forest เหมาะกับงานที่ต้องการความแม่นยำสูงและลดข้อผิดพลาดได้อย่างมีประสิทธิภาพ ขณะที่ SVM และ Isolation Forest เหมาะสำหรับระบบที่ต้องการความยืดหยุ่น และรองรับการเปลี่ยนแปลงข้อมูลในสภาพแวดล้อมที่ซับซ้อน



ภาพที่ 4 ความเข้มของค่าความถูกต้องของอัลกอริธึมการเรียนรู้ของเครื่องในสถานการณ์ต่าง ๆ

ภาพที่ 4 แสดงความเข้ม (Heatmap) ซึ่งเปรียบเทียบค่าความถูกต้อง (Accuracy) ของอัลกอริธึม Machine Learning ทั้ง 5 รูปแบบ ใน 3 สถานการณ์การทดสอบ (Scenario 1, Scenario 2 และ Scenario 3) ผลการวิเคราะห์พบว่า Neural Networks มีประสิทธิภาพสูงสุดในทุกสถานการณ์ โดยเฉพาะใน Scenario 3 ที่มีค่า Accuracy สูงถึง 95.2% ซึ่งสะท้อนถึงความสามารถในการตรวจจับข้อมูลที่ซับซ้อนและการเรียนรู้ที่มีประสิทธิภาพ เหมาะสำหรับงานที่ต้องการความแม่นยำสูงและระบบที่มีการประมวลผลที่ซับซ้อน อย่างไรก็ตาม Neural Networks อาจมีข้อจำกัดด้านทรัพยากรที่ต้องการการประมวลผลสูง สำหรับ Random Forest ผลลัพธ์แสดงถึงความสมดุลในหลายตัวชี้วัดโดยมีค่า Accuracy สูงสุดที่ 93.0% ใน Scenario 3 และ 91.5% ใน Scenario 2 เหมาะสำหรับงานที่ต้องการความยืดหยุ่นในการจัดการข้อมูลที่หลากหลายและไม่สมดุล ส่วน SVM แสดงค่า Accuracy ที่ดี โดยมีค่า Accuracy สูงสุดที่ 92.0% ใน Scenario 3 ขณะที่ Isolation Forest และ Autoencoders มีค่า Accuracy ต่ำกว่า แต่ยังคงเหมาะกับงานที่มีข้อจำกัดด้านทรัพยากรหรือข้อมูลที่ไม่ได้ติดป้ายล่วงหน้า การวิเคราะห์นี้ช่วยให้เลือกอัลกอริธึมได้เหมาะสมกับลักษณะงานและข้อจำกัดของระบบ (Sarhan et al., 2023)

ตารางที่ 3 การเปรียบเทียบ False Positive Rate (FPR) และการแจ้งเตือนผิดพลาดของอัลกอริธึม Machine Learning

อัลกอริธึม	False Positive Rate (FPR)	False Positives (%)
Neural Networks	2.5	2.5
Random Forest	3.0	3.0
Support Vector Machine (SVM)	4.2	4.2
Isolation Forest	5.8	5.8
Autoencoders	5.0	5.0

ตารางที่ 3 สรุปค่า False Positive Rate (FPR) และจำนวนการแจ้งเตือนผิดพลาดของอัลกอริธึม Machine Learning ที่ใช้ตรวจจับ Zero-Day Attack โดย Neural Networks แสดงประสิทธิภาพสูงสุดด้วย FPR ต่ำสุดเพียง 2.5% บ่งบอกถึงความสามารถในการลดการแจ้งเตือนผิดพลาดได้ดี รองลงมาคือ Random Forest ด้วย FPR 3.0% ซึ่งแสดงถึงความสมดุลระหว่างความแม่นยำและการลดการแจ้งเตือนผิดพลาด ในขณะที่ SVM และ Isolation Forest มีค่า FPR สูงกว่า คือ 4.2% และ 5.8% ตามลำดับ ซึ่งสะท้อนถึงความเสี่ยงต่อการแจ้งเตือนผิดพลาดที่มากขึ้น Auto-encoders มีค่า FPR 5.0% แต่ยังคงเป็นตัวเลือกที่เหมาะสมในบางกรณี โดยเฉพาะสถานการณ์ที่ข้อมูลไม่มีการติดป้ายกำกับที่ชัดเจน จากผลการวิเคราะห์ชี้ให้เห็นว่า Neural Networks และ Random Forest เหมาะสมที่สุดสำหรับระบบที่ต้องการความแม่นยำและความน่าเชื่อถือ ในขณะที่ SVM, Isolation Forest และ Autoencoders อาจเหมาะสมสำหรับระบบที่มีข้อจำกัดด้านข้อมูลหรือทรัพยากรเฉพาะ (Kumar et al., 2022)

ตารางที่ 4 ความสามารถในการตอบสนองต่อภัยคุกคามใหม่ของอัลกอริธึม Machine Learning

อัลกอริธึม	การตอบสนอง ต่อภัยคุกคามใหม่ (%)	ข้อสังเกต
Neural Networks	95.2	มีประสิทธิภาพสูงในการเรียนรู้พฤติกรรมใหม่จากข้อมูลที่ไม่เคยพบมาก่อน
Random Forest	90.5	สามารถจัดการกับข้อมูลที่มีหลายมิติและเรียนรู้จากพฤติกรรมที่แปลกใหม่
Support Vector Machine (SVM)	89.0	เหมาะสำหรับการแยกแยะข้อมูลในกรณีที่มีความซับซ้อน
Isolation Forest	86.5	มีประสิทธิภาพในการตรวจจับพฤติกรรมผิดปกติ แม้ว่า False Positives จะสูง
Autoencoders	88.0	สามารถเรียนรู้จากข้อมูลแบบไม่มีการติดป้ายกำกับ แต่มีประสิทธิภาพต่ำกว่าที่คาด

ตารางที่ 4 นำเสนอการเปรียบเทียบประสิทธิภาพของอัลกอริธึม Machine Learning ในการตอบสนองต่อภัยคุกคามใหม่ ซึ่งเป็นปัจจัยสำคัญในระบบรักษาความปลอดภัยทางไซเบอร์ Neural Networks แสดงความสามารถสูงสุดในการเรียนรู้และปรับตัวต่อพฤติกรรมที่ไม่เคยพบมาก่อน โดยมีอัตราการตอบสนอง 95.2% ซึ่งบ่งบอกถึงความแม่นยำและประสิทธิภาพในการจัดการกับข้อมูลที่เปลี่ยนแปลงอย่างต่อเนื่อง Random Forest มีการตอบสนองที่สมดุลที่ 90.5% เหมาะสำหรับการจัดการกับข้อมูลที่มีหลากหลายมิติและมีความซับซ้อน ในขณะที่ Support Vector Machine (SVM) มีประสิทธิภาพในการแยกแยะข้อมูลที่ซับซ้อนด้วยอัตราการตอบสนอง 89.0% ส่วน Isolation Forest และ Autoencoders แม้มีอัตราการตอบสนองต่ำกว่าอยู่ที่ 86.5% และ 88.0% ตามลำดับ แต่มีความโดดเด่นในการจัดการข้อมูลที่ไม่มีการติดป้ายกำกับ และเหมาะสำหรับระบบที่ต้องการความยืดหยุ่น อย่างไรก็ตาม False Positives ที่สูงใน Isolation Forest ถือเป็นข้อจำกัดที่ต้องพิจารณา การวิเคราะห์นี้ชี้ให้เห็นว่าการเลือกใช้อัลกอริธึมที่เหมาะสมควรพิจารณาจากลักษณะของข้อมูล ความต้องการเฉพาะของระบบ และความสามารถในการรับมือกับภัยคุกคามใหม่อย่างต่อเนื่อง เพื่อเสริมสร้างความมั่นคงปลอดภัยในโลกไซเบอร์ที่เปลี่ยนแปลงอย่างรวดเร็ว

สรุปและอภิปรายผลการวิจัย

การประยุกต์ใช้ Machine Learning (ML) ในการตรวจจับและป้องกันการโจมตี Zero-Day เป็นหนึ่งในประเด็นที่ได้รับความสนใจอย่างมากในวงการความปลอดภัยไซเบอร์ เนื่องจาก Zero-Day Attacks เป็นภัยคุกคามที่อาศัยช่องโหว่ในระบบหรือซอฟต์แวร์ที่ยังไม่มีการแก้ไข ทำให้ระบบตรวจจับแบบดั้งเดิม เช่น Signature-Based Detection ไม่สามารถตอบสนองต่อภัยคุกคามได้อย่างมีประสิทธิภาพ การวิจัยในหัวข้อนี้จึงมีความสำคัญอย่างยิ่ง ในการพัฒนาเครื่องมือที่สามารถจัดการกับภัยคุกคามที่ซับซ้อนและเปลี่ยนแปลงได้ Neural Networks (NN) ได้แสดงให้เห็นถึงศักยภาพที่สูงกว่าอัลกอริธึมอื่น ๆ ในการตรวจจับ Zero-Day Attacks โดยให้ค่าความแม่นยำ (Accuracy) สูงถึง 95.2% และ False Positive Rate (FPR) ต่ำเพียง 2.5% ซึ่งบ่งบอกถึงความสามารถในการเรียนรู้พฤติกรรมที่ผิดปกติจากข้อมูลใหม่ NN สามารถตอบสนองต่อภัยคุกคามที่ซับซ้อนและมีความยืดหยุ่นในการปรับตัวให้เข้ากับสภาพแวดล้อมที่เปลี่ยนแปลงได้อย่างรวดเร็ว ในขณะที่ Random Forest (RF) และ Support Vector Machines (SVM) มีศักยภาพในระดับที่แตกต่างกัน RF ซึ่งมีค่าความแม่นยำ 90.5% และ Precision 91.6% แสดงถึงความสมดุลในด้านประสิทธิภาพและความสามารถในการจัดการข้อมูลที่หลากหลาย SVM มีค่าความแม่นยำที่ 88.7% และ False Positive Rate (FPR) 4.5% โดยเหมาะสำหรับข้อมูลที่ซับซ้อนซึ่งต้องการการแยกแยะในระดับสูง แต่ทั้ง RF และ SVM ยังคงมีข้อจำกัดในด้านประสิทธิภาพและความยืดหยุ่นเมื่อเปรียบเทียบกับ NN สำหรับ Isolation Forest และ Autoencoders ซึ่งเป็นอัลกอริธึมแบบ Unsupervised Learning มีข้อได้เปรียบในด้านการจัดการข้อมูลที่ไม่ได้ติดป้ายกำกับ แต่ยังคงมีข้อจำกัดสำคัญ เช่น ค่าความแม่นยำ (Accuracy) ต่ำกว่า NN อยู่ที่ 90.0% และ 91.4% ตามลำดับ และ FPR ที่สูงกว่า (6.0% และ 5.8%) ทำให้อัลกอริธึมทั้งสองนี้เหมาะสำหรับสภาพแวดล้อมที่ข้อมูลมีการเปลี่ยนแปลงอย่างต่อเนื่อง และต้องการความยืดหยุ่นสูง ผลการวิจัยยังได้เน้นถึงบทบาทสำคัญของคุณภาพและความหลากหลายของชุดข้อมูลในการฝึกฝนอัลกอริธึม ML การใช้ข้อมูลจาก

Threat Intelligence เป็นองค์ประกอบที่สำคัญ เนื่องจากช่วยเพิ่มศักยภาพของระบบในการตรวจจับและป้องกันภัยคุกคามที่ซับซ้อนและเปลี่ยนแปลง การบูรณาการข้อมูล Threat Intelligence ช่วยให้ระบบสามารถเรียนรู้พฤติกรรมใหม่ ๆ และเพิ่มความแม่นยำในการตรวจจับภัยคุกคามในสถานการณ์จริง นอกจากนี้ การใช้ Ensemble Learning ซึ่งผสมผสานข้อดีของ NN และ RF ยังช่วยเสริมสร้างความแข็งแกร่งของระบบตรวจจับ โดย Ensemble Learning ช่วยลดข้อผิดพลาดและเพิ่มความสามารถในการปรับตัวให้เข้ากับข้อมูลที่หลากหลาย การรวมข้อดีของ NN ที่มีความแม่นยำสูงและ RF ที่มีความสมดุลในด้านประสิทธิภาพ ทำให้ Ensemble Learning มีศักยภาพในการตอบสนองต่อภัยคุกคามที่ซับซ้อนในบริบทต่าง ๆ เทคโนโลยี Deep Learning เช่น Convolutional Neural Networks (CNN) และ Recurrent Neural Networks (RNN) ยังแสดงศักยภาพในด้านการวิเคราะห์ข้อมูลเครือข่าย CNN มีความเหมาะสมสำหรับการวิเคราะห์ข้อมูลหลายมิติ เช่น การสร้างรูปแบบการเชื่อมโยงในเครือข่าย ขณะที่ RNN เหมาะสำหรับข้อมูลที่มีลำดับเวลา เช่น การติดตามพฤติกรรมการใช้งานของผู้ใช้ การพัฒนา Hybrid Algorithms ที่ผสมผสาน Neural Networks กับเทคนิค Anomaly Detection เช่น Autoencoders ยังช่วยลด False Positive Rate และเพิ่มความยืดหยุ่นในระบบตรวจจับภัยคุกคาม ผลการวิจัยชี้ให้เห็นว่าการพัฒนาระบบตรวจจับที่สามารถรองรับ Continuous Learning จะช่วยให้ระบบสามารถปรับตัวเข้ากับข้อมูลใหม่ได้อย่างรวดเร็วในสภาพแวดล้อมที่เปลี่ยนแปลง การปรับปรุงประสิทธิภาพของอัลกอริธึม ML โดยการปรับแต่งค่าพารามิเตอร์ (Hyperparameters) และการพัฒนาโครงสร้างโมเดลที่เหมาะสมกับลักษณะข้อมูลเฉพาะ จะช่วยเพิ่มขีดความสามารถของระบบในการรับมือกับภัยคุกคาม การผสมผสานเทคโนโลยีใหม่ ๆ เช่น Deep Learning และการพัฒนา Hybrid Algorithms พร้อมกับการใช้ Ensemble Learning อย่างเต็มรูปแบบ จะช่วยสร้างระบบที่มีความแม่นยำสูงและตอบสนองได้ดีต่อภัยคุกคามที่ซับซ้อนในอนาคต โดยรวม งานวิจัยนี้ได้ชี้ให้เห็นถึงความสำคัญของการเลือกใช้อัลกอริธึม ML ที่เหมาะสม

กับลักษณะข้อมูลและข้อกำหนดเฉพาะของระบบ Neural Networks เหมาะสำหรับงานที่ต้องการความแม่นยำสูง และการตอบสนองแบบเรียลไทม์ Random Forest และ SVM เหมาะกับบริบทที่ต้องการความสมดุลในด้านประสิทธิภาพและทรัพยากร ส่วน Isolation Forest และ Autoencoders เหมาะสำหรับการใช้งานในสภาพแวดล้อมที่ข้อมูลมีการเปลี่ยนแปลงบ่อยและต้องการความยืดหยุ่นสูง การพัฒนาเทคโนโลยี ML อย่างต่อเนื่อง และการบูรณาการเทคนิคที่หลากหลาย จะช่วยเสริมสร้างความปลอดภัยในระบบเครือข่าย พร้อมทั้งเพิ่มความสามารถในการตอบสนองต่อภัยคุกคามที่ซับซ้อนในยุคดิจิทัลอย่างมีประสิทธิภาพ

งานวิจัยนี้ยังสนับสนุนสมมติฐานงานวิจัย โดยพบว่า อัลกอริธึม Machine Learning สามารถตรวจจับการโจมตีแบบ Zero-Day ได้อย่างแม่นยำและมีประสิทธิภาพ โดยเฉพาะ Neural Networks ซึ่งแสดงค่าความแม่นยำสูงสุด (95.2%) และลดอัตราการแจ้งเตือนผิดพลาดได้ดีกว่าอัลกอริธึมอื่น เช่น Random Forest และ SVM นอกจากนี้ การใช้เทคนิค Anomaly Detection และ Ensemble Learning ช่วยเพิ่มความแม่นยำ ความยืดหยุ่น และลดข้อผิดพลาดของระบบในการจัดการภัยคุกคามที่ซับซ้อนและเปลี่ยนแปลงได้ ผลวิจัยจึงสนับสนุนศักยภาพของ Machine Learning ในการพัฒนาระบบรักษาความปลอดภัยเครือข่ายที่ตอบสนองต่อภัยคุกคามในอนาคตอย่างมีประสิทธิภาพ

ข้อเสนอแนะ

1. ควรพัฒนาและปรับปรุงอัลกอริธึม Machine Learning เช่น Deep Learning และ Ensemble Learning เพื่อเพิ่มความแม่นยำในการตรวจจับภัยคุกคาม พร้อม ลดอัตราการแจ้งเตือนผิดพลาด (False Positive Rate) โดยใช้เทคนิคที่ผสมผสานข้อดีของ Neural Networks และ Random Forest ซึ่งแสดงให้เห็นผลลัพธ์ที่ดีในงานวิจัย นอกจากนี้ ควรออกแบบระบบให้รองรับ Continuous Learning เพื่อให้สามารถตอบสนองต่อภัยคุกคามที่ซับซ้อนและเปลี่ยนแปลงได้อย่างรวดเร็ว

2. ควรเก็บรวบรวมข้อมูลเพิ่มเติมจากแหล่งที่หลากหลาย เช่น ระบบป้องกันภัยในองค์กรและ Threat Intelligence โดยเฉพาะข้อมูลเชิงพฤติกรรมที่ครอบคลุมหลากหลายรูปแบบ การบูรณาการข้อมูลเหล่านี้จะช่วยเสริมประสิทธิภาพของโมเดลในการเรียนรู้และตรวจจับภัยคุกคามที่ซับซ้อนได้ดียิ่งขึ้น พร้อมทั้งลดความเสี่ยงที่ระบบจะพลาดการตรวจจับภัยคุกคามที่ยังไม่เคยปรากฏ มาก่อน

3. ควรจัดการฝึกอบรมบุคลากรที่เกี่ยวข้องให้มีความรู้และทักษะในด้านการใช้งานเทคโนโลยี Machine Learning รวมถึงความเข้าใจในพฤติกรรมและแนวโน้มของภัยคุกคามที่เปลี่ยนแปลงอยู่เสมอ ความรู้เหล่านี้จะช่วยให้บุคลากรสามารถปรับปรุงระบบให้ตอบสนองต่อภัยคุกคามได้อย่างมีประสิทธิภาพและทันต่อสถานการณ์

เอกสารอ้างอิง

- Chua, Y. F., & Salam, R. A. (2023). Evaluation of machine learning algorithms in network-based intrusion detection using progressive dataset. *Symmetry*, 15(6), 1-31.
<https://doi.org/10.3390/sym15061251>
- Dawadi, B. R., Adhikari, B., & Srivastava, D. K. (2023). Deep learning technique-enabled web application firewall for the detection of web attacks. *Sensors*, 23(4), 1-16. <https://doi.org/10.3390/s23042073>
- Dean, J. (2022). A golden decade of deep learning: Computing systems & applications. *Dædalus, the Journal of the American Academy of Arts & Sciences*, 151(2), 58-74.
https://doi.org/10.1162/daed_a_01900
- Gupta, I., Kumari, S., Jha, P., & Ghosh, M. (2024). Leveraging LSTM and GAN for modern malware detection. *arXiv*. <https://doi.org/10.48550/arXiv.2405.04373>
- Kumar, P. S., B, U. S., Mishra, I., S, S. S., Tripathi, D. R., & Rama Krishna T., S. (2022). Malware detection classification using recurrent neural network. In *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*. (876–880). IEEE. <https://doi.org/10.1109/ICTACS56270.2022.9988624>
- Salem, A. H., Azzam, S. M., Emam, O. E., & Abohany, A. A. (2024). Advancing cybersecurity: A comprehensive review of AI-driven detection techniques. *Journal of Big Data*, 11(105), 1-38.
<https://doi.org/10.1186/s40537-024-00957-y>
- Sarhan, M., Layeghy, S., Gallagher, M., & Portmann, M. (2023). From zero-shot machine learning to zero-day attack detection. *International Journal of Information Security*, 22(6), 947–959.
<https://link.springer.com/article/10.1007/s10207-023-00676-0>
- Sewak, M., Sahay, S. K., & Rathore, H. (2022). Deep reinforcement learning for cybersecurity threat detection and protection: A review. *arXiv preprint arXiv:2206.02733*. <https://arxiv.org/abs/2206.02733>
- Singh, S., Kumar, R., Payra, S., & Singh, S. K. (2023). Artificial Intelligence and Machine learning in pharmacological research: Bridging the gap between data and drug discovery. *Cureus*, 15(8), e44359. <https://doi.org/10.7759/cureus.44359>
- Yanguema, A. (2023). Fortifying network security with machine learning. *SSRN*. 1-9 <https://doi.org/10.2139/ssrn.4663593>
-