การจำลองสถานการณ์การโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ทางความมั่นคงปลอดภัยไซเบอร์

Cyber Attack with Social Engineering Simulation Using Deep Leaning to Enhance Cybersecurity Awareness

สิรินทรา พิฤทธิ์บูรณะ (Sirintra Piritaburana)[1] สุรศักดิ์ มังสิงห์ (Surasak Mungsing)[2]
และประสงค์ ปราณีตพลกรัง (Prasong Praneetpolgrang)[3]
[1]หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
[1, 2, 3] คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม
Sirintra.ben@gmail.com[1], Surasak.mu@spu.ac.th[2], Prasongspu@gmail.com[3]

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อจำลองสถานการณ์ และประเมินการโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ทางความมั่นคงปลอดภัยไซเบอร์ เครื่องมือที่ใช้ คือ โปรแกรม RapidMiner Studio โดยใช้อัลกอริทึมเพื่อจำลองการโจมตีทั้งหมด 4 อัลกอริทึม ได้แก่ การเรียนรู้เชิงลึก (Deep Learning) นิวรอนเน็ตเวิร์ก (Neural Networks) Automated MLP (AutoMLP) และเพอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptrons) ร่วมกับชุดข้อมูล URL Dataset (ISCX-URL2016) แล้วทำการประเมินเพื่อเปรียบเทียบประสิทธิภาพ การวิจัยครั้งนี้พบว่า อัลกอริทึมการเรียนรู้เชิงลึก (Deep Learning) นั้นมีค่าเฉลี่ยความแม่น (Accuracy) กับค่าเฉลี่ยความเที่ยง (Precision) เป็นเปอร์เซ็นต์สูงที่สุด โดยมีค่าเฉลี่ยความแม่นเท่ากับ 96.92% ค่าเฉลี่ยความเที่ยงเท่ากับ 96.96% รวมถึงอัตราค่าเฉลี่ยความระลึก (Recall) เท่ากับ 96.88% อัตราค่าเฉลี่ยความถ่วงดุล (F-Measure) เท่ากับ 96.92% และอัลกอริทึมนิวรอนเน็ตเวิร์ก (Neural Networks) พบว่าอัตราค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error : MAE) เท่ากับ 3.1% และโดยค่าเฉลี่ยรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error : RMSE) เท่ากับ 17.6%

คำสำคัญ: การจำลองสถานการณ์ วิธีวิศวกรรมสังคม
การเรียนรู้เชิงลึก

Abstract

The objective of this research is to simulate the cyber attack with social engineering using deep learning to enhance cybersecurity awareness. The tool used is the RapidMiner Studio program, using algorithms to simulate attacks. A total of 4 algorithms were used, including deep learning, neural networks, automated MLP, and multilayer perceptrons. These algorithms were used together with the URL Dataset (ISCX-URL2016) and evaluated to compare their performance. This research found that deep learning algorithms achieved the highest average accuracy and precision percentages. The average accuracy was 96.92% while the average precision was 96.96% Additionally, the average recall rate was 96.88% and the average F-measure was 96.92% on the other hand, the neural networks algorithm had a mean absolute error (MAE) of 3.1% and a root mean square error (RMSE) of 17.6%

Keywords: Simulation, Social Engineering,
Deep Learning

1. บทนำ

ปัจจุบันนี้หากมองถึงการใช้งานด้านเทคโนโลยีดิจิทัลมีนัยยะสำคัญต่อการพลิกฟื้นปรับปรุง ร่วมถึงการยกระดับประสิทธิภาพการทำงานขององค์กร ซึ่งมีส่วนสำคัญในการสร้างเศรษฐกิจไทยให้เข้มแข็ง อย่างไรก็ตาม การนำเทคโนโลยีดิจิทัลสมัยใหม่มาประยุกต์ใช้ย่อมมีภัยคุกคามและความเสี่ยงทางไซเบอร์ตามมาด้วย อาทิเช่น การโจรกรรม การเรียกค่าไถ่ และการปลอมแปลงข้อมูลเพื่อหลอกลวง ซึ่งจะส่งผลกระทบต่อการมีสุขภาพไซเบอร์ที่ดี (Cyber Hygiene)

การจำลองสถานการณ์การโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม และใช้การเรียนรู้เชิงลึกจะช่วยให้องค์กรสามารถจำลองสถานการณ์การโจมตีที่หลากหลายและซับซ้อนได้ จากการเรียนรู้เชิงลึก (Deep Learning) จะสามารถเสริมสร้างความตระหนักรู้ทางความมั่นคงปลอดภัยไซเบอร์และเป็นเครื่องมือสำคัญในการจำลองสถานการณ์การโจมตี จากการเรียนรู้รูปแบบและพฤติกรรมของมนุษย์ได้อย่างแม่นยำ ทำให้สามารถจำลองสถานการณ์การโจมตีได้อย่างสมจริงและมีประสิทธิภาพ

จากที่กล่าวมาข้างต้นผู้วิจัยจึงใช้อัลกอริทึมการเรียนรู้เชิงลึก เพื่อจำลองสถานการณ์การโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคมโดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ทางความมั่นคงปลอดภัยไซเบอร์ สำหรับองค์กร ผู้วิจัยจึงได้นำโปรแกรม RapidMiner Studio มาสร้างแบบจำลอง เพื่อเปรียบเทียบประสิทธิภาพและเลือกวิธีที่เหมาะสมมากที่สุด จะสามารถเสริมสร้างความตระหนักรู้ ทางความมั่นคงปลอดภัยไซเบอร์

2. ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ได้ดำเนินการรวบรวมทฤษฎี และงานวิจัยที่เกี่ยวข้องตามรายละเอียดดังต่อไปนี้

2.1 อัลกอริทึมที่ใช้ในการจำลองสถานการณ์

2.1.1 อัลกอริทึมการเรียนรู้เชิงลึก [1] เป็นลำดับชั้นของเครือข่ายประสาทเทียม (Artificial Neural Networks) ที่ดำเนินการใช้การเรียนรู้เครื่อง จากเครือข่ายประสาทเทียมถูกสร้างมาจำลองการทำงานของสมองมนุษย์ โดยมีโหนด (Node) ที่เชื่อมต่อกันเหมือนเว็บไซต์ แต่การทำงานของการเรียนรู้เชิงลึก จะช่วยให้การประมวลผลข้อมูลเกิดขึ้นด้วยวิธีที่ไม่ใช่เชิงเส้นทำให้เครื่องจักรสามารถจัดการกับข้อมูลที่ซับซ้อนและปัญหาที่ซับซ้อน จากการวิเคราะห์แบบเชิงเส้นได้มากขึ้น

2.1.2 อัลกอริทึม Neural Networks [2] คือระบบทางคอมพิวเตอร์ที่จำลองการทำงานของระบบประสาทเทียมในสมองมนุษย์ จึงจะให้คอมพิวเตอร์สามารถเรียนรู้และทำนายข้อมูลได้อัตโนมัติ อัลกอริทึมนิวรอนเน็ตถูกออกแบบให้มีลักษณะการทำงานคล้ายกับโครงสร้างของเซลล์ประสาทในสมองมนุษย์ ประสาทเทียมประกอบด้วยชั้นหนึ่งหรือมากกว่าที่ปรับปรุงค่าน้ำหนัก (Weights) เพื่อให้สามารถเรียนรู้และทำนายผลลัพธ์ของข้อมูลที่นำเข้าได้ จากโครงสร้างพื้นฐานประกอบด้วยชั้นนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layers) และชั้นผลลัพธ์ (Output Layer) จากการคำนวณข้อมูลผ่านทางชั้นต่าง ๆ มีการกำหนดค่าน้ำหนักและค่าปรับ (Biases) ในขั้นตอนการฝึกออกมา จากชุดข้อมูลการฝึก (Train Data) พร้อมปรับปรุงค่าเหล่านี้ให้ทำนายผลลัพธ์ที่ถูกต้อง

2.1.3 อัลกอริทึม Automated MLP (Auto MLP) [3] เป็นอัลกอริทึมจากการเรียนรู้เชิงคำนวณที่ออกแบบมาเพื่อปรับขนาดของโมดูล ในระบบอัตโนมัติแบบ Sequential จะเป็นระบบที่แนะนำไอเท็มตามลำดับของการใช้งานของผู้ใช้จะได้พิจารณาจากความสนใจของผู้ใช้ในระยะสั้นและระยะยาว ฉะนั้น อัลกอริทึม Automated MLP จึงทำงานในรูปแบบการแบ่งข้อมูลในการใช้งานของผู้ใช้ออกเป็น 2 ส่วนคือ ข้อมูลความสนใจระยะสั้น จะถูกปรับขนาดอัตโนมัติตามความสนใจในไอเท็มใหม่ ๆ จากการเรียนรู้ความสนใจใหม่ ๆ ได้อย่างรวดเร็ว ส่วนข้อมูลความสนใจระยะยาว จะถูกปรับขนาดอัตโนมัติตามความสนใจในไอเท็มแบบเดิม ๆ เพื่อเรียนรู้จากสิ่งเดิม ๆ ได้อย่างแม่นยำมากขึ้น ฉะนั้น อัลกอริทึมนี้สามารถเรียนรู้และปรับความยาวของความสนใจระยะสั้นได้ ซึ่งช่วยให้เข้าใจความต้องการที่เปลี่ยนแปลงของผู้ใช้

2.1.4 อัลกอริทึม Multilayer Perceptrons [4] เป็นเป็นอัลกอริทึมการเรียนรู้เชิงลึกของชนิดหนึ่งที่ประกอบด้วย

ชั้นของโหนด (Node) หลายชั้น แต่ละชั้นจะเชื่อมต่อกันด้วยน้ำหนัก (Weight) และฟังก์ชันกระตุ้น (Activation Function) โดยอัลกอริทึม Multilayer Perceptrons สามารถใช้สำหรับการแก้ปัญหาที่ซับซ้อนได้หลากหลาย เช่น การจำแนกประเภท การถดถอย เพื่อนำข้อมูลเข้ามาประมวลผล ซึ่งจะแปลงข้อมูลให้เป็นรูปแบบที่ต้องการ หรือการเรียนรู้ด้วยการปรับพารามิเตอร์ จะใช้กระบวนการเรียนรู้ของอัลกอริทึมซึ่งจะเกิดขึ้น โดยการปรับค่าน้ำหนักของการเชื่อมต่อระหว่างโหนดแต่ละโหนด เพื่อให้เอาต์พุตของอัลกอริทึมอยู่ใกล้เคียงกับเอาต์พุตเป้าหมายมากที่สุด

2.2 งานวิจัยที่เกี่ยวข้อง

สุรชัย ฉัตรเฉลิมพันธุ์ และเทอดพงษ์ แดงสี [5] ได้นำเสนอเกี่ยวกับการเสริมสร้าง ความตระหนักรู้เท่าทันภัยทางไซเบอร์ของบุคลากรในองค์กร: กรณีการจำลองการโจมตีด้วยฟิชชิง ได้ดำเนินการจำลองการโจมตีด้วยอีเมลฟิชชิงในบริษัทแห่งหนึ่ง และผู้ทดสอบได้ดำเนินการทดสอบในรูปแบบไซเบอร์ดริลล์ (Cyber Drill) จากการปฏิบัติของพนักงานภายในบริษัท ที่ผ่านการจำลองการโจมตีแบบฟิชชิง และการใช้กระบวนการถ่ายโอนความรู้ควรเพิ่มความตระหนักถึงความปลอดภัยทางไซเบอร์ในหมู่พนักงาน ซึ่งจะช่วยลดความเสี่ยงที่จะตกเป็นเหยื่อของภัยคุกคามทางไซเบอร์

Almomani, et al [6] ได้นำเสนอเกี่ยวกับการเรียนรู้เชิงลึก โดยการมองเห็นแบบอัตโนมัติแบบจำลองสำหรับการตรวจจับการโจมตีมัลแวร์ ด้วยระบบแอนดรอยด์อย่างมีประสิทธิภาพ ในการใช้รูปแบบการจำลองด้วยอัลกอริทึม Convolutional Neural Networks (CNN) มาตรวจจับความแม่นยำ จากชุดข้อมูลที่สมดุลและไม่สมดุลมาทดสอบ

พงศ์พนธ์ ภาวศุทธิ์ และคณะ [7] ได้นำเสนอเกี่ยวกับสาเหตุเชิงลึกของการถูกโจมตีด้วยวิธีการทางวิศวกรรมสังคมของกลุ่มเจเนอชั่นวาย ในเขตกรุงเทพมหานครและปริมณฑล จากการศึกษาวิเคราะห์ และสำรวจบทบาทของอารมณ์ เช่น ความอยากรู้อยากเห็น และความโลภในการตัดสินใจที่เกี่ยวข้องกับภัยคุกคามทางวิศวกรรม เพื่อลดการเสี่ยงต่อผู้ประสบภัยจะต้องมีการเรียนรู้และตระหนักรู้ในกลุ่มเป้าหมาย

3. วิธีการดำเนินงาน

งานวิจัยนี้ได้ทำการศึกษาจากทฤษฎีต่าง ๆ และงานวิจัยที่เกี่ยวข้อง นำมาใช้ในจำลองสถานการณ์ และประเมินการโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ทางความมั่นคงปลอดภัยไซเบอร์ มีรายละเอียดดังต่อไปนี้

3.1 เครื่องมือในการวิจัย

3.1.1 โปรแกรม RapidMiner Studio

3.1.2 ชุดคำสั่ง Split Validation

3.1.3 อัลกอริทึมใช้ในการจำลอง 4 อัลกอริทึม ได้แก่ Deep Learning, Neural Networks, Automated MLP และ Multilayer Perceptrons เป็นต้น

3.1.4 ชุดข้อมูล URL Dataset (ISCX-URL2016) เป็นชุดข้อมูลนี้ได้รวบรวมข้อมูลจากการสำรวจจากการตรวจจับและจัดหมวดหมู่ URL ที่เป็นอันตรายของประเภทการโจมตีต่าง ๆ เช่น ฟิชชิง และมัลแวร์ [8]

3.2 ขั้นตอนการดำเนินการวิจัย

3.2.1 ทำความเข้าใจเกี่ยวกับข้อมูลการโดนโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยผู้วิจัยได้นำชุดข้อมูล URL Dataset (ISCX-URL2016) ซึ่งเป็นข้อมูลที่ได้รวบรวมข้อมูลจากการสำรวจจากการตรวจจับประเภทการโจมตีทางไซเบอร์ มีจำนวน 79 แอททริบิวต์ และจำนวน 8,454 เรคคอร์ด แสดงดังตารางที่ 1 และ 2

ตารางที่ 1: แสดงรายชื่อและประเภทของแอททริบิวต์

ลำดับ	ชื่อแอททริบิวต์	ชนิดข้อมูล
1	fileNameLen	Discrete
2	domain_token_count	Discrete
3	tld	Discrete
4	SymbolCount_Domain	Discrete
5	Entropy_Afterpath	Discrete
6	delimeter_path	Discrete
7	argPathRatio	Discrete
8	Entropy_Filename	Discrete
9	Entropy_DirectoryName	Discrete
10	Filename_LetterCount	Discrete
…	…	…

ตารางที่ 1: แสดงรายชื่อและประเภทของแอททริบิวต์ (ต่อ)

ลำดับ	ชื่อแอททริบิวต์	ชนิดข้อมูล
69	Entropy_URL	Discrete
70	isPortEighty	Discrete
71	Directory_LetterCount	Discrete
72	pathDomainRatio	Discrete
73	dld_domain	Discrete
74	NumberRate_URL	Discrete
75	sub-Directory_ LongestWordLength	Discrete
76	Path_LongestWordLength	Discrete
77	charcompvowels	Discrete
78	avgdomaintokenlen	Discrete
79	Type	Discrete

ตารางที่ 2: ประเภทของภัยคุกคามทางไซเบอร์

ประเภทการโจมตี	จำนวน	อัตราส่วน
Phishing	4,440	52.52%
Malware	4,014	47.48%
ALL	8,454	100.00%

3.2.2 การเตรียมข้อมูล ผู้วิจัยได้ทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วน ในอัตราเปอร์เซ็นต์ 80:20 คือ อัตราส่วน 80 เปอร์เซ็นต์แรก ใช้สำหรับการฝึกสอน (Train Data) และให้อัลกอริทึมเกิดการเรียนรู้ ส่วนอัตรา 20 เปอร์เซ็นต์ที่สองใช้สำหรับการทดสอบอัลกอริทึม (Test Data)

3.3 การสร้างแบบจำลองการประเมินและเปรียบเทียบประสิทธิภาพ

ผู้วิจัยได้สร้างแบบจำลองจากการใช้โปรแกรม RapidMiner Studio ร่วมกับใช้ชุดข้อมูล URL Dataset (ISCX-URL2016) จากการใช้ชุดคำสั่ง Split Validation และเลือกใช้อัลกอริทึมจำนวน 4 อัลกอริทึม เมื่อได้ผลของการสร้างแบบจำลองนี้แล้ว จะต้องมีการประเมินและเปรียบเทียบประสิทธิภาพ ของการใช้อัลกอริทึมที่มีการจำแนกข้อมูลจากการจำลองสถานการณ์การ และประเมินโจมตีทางไซเบอร์ด้วยอัลกอริทึม Deep Learning, Neural Networks, Automated MLP และ Multilayer Perceptrons เพื่อนำมาเปรียบเทียบประสิทธิภาพของแต่ละอัลกอริทึม ได้ตามสมการการวัดค่าความแม่น (Accuracy) ค่าความเที่ยง (Precision) ค่าความระลึก (Recall) ค่าความถ่วงดุล (F-Measure) ที่สูงที่สุด [9] และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error : MAE) และค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error : RMSE) ที่ต่ำที่สุด [10] ได้ตามสมการที่ (1) – (6) ดังนี้

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F\text{-}Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

โดยที่ TP = True Positive คือ ข้อมูลที่ทำนายถูกต้องเมื่อเทียบกับเฉลย

FP = False Positive คือ ข้อมูลที่ทำนายและไม่ถูกต้องเมื่อเทียบกับเฉลย

TN = True Negative คือ ข้อมูลที่อยู่ในเฉลยแต่ไม่มีการทำนาย (ตรงข้ามกับ FN)

Precision คือ ค่าความเที่ยงเกิดจากการนำค่า TP มาเทียบกกับ FP

Recall คือ ค่าความระลึกเกิดจากการนำค่า TP มาเทียบกับ FN

F-Measure คือ ค่าเฉลี่ยของ Precision และ Recall

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2} \quad (6)$$

ทั้งนี้ MAE คือ ค่าความคาดเคลื่อนแบบสัมบูรณ์เฉลี่ย

RMAE คือ ค่าความคาดเคลื่อนแบบกำลังสองเฉลี่ย

n คือ จำนวนข้อมูล

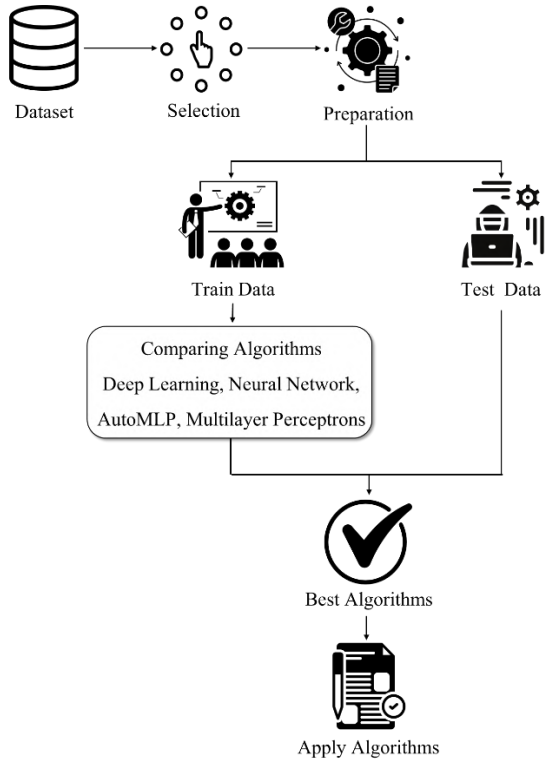
y_i คือ ค่าที่ทำนายของแบบจำลอง

$\hat{y}_i$ คือ ค่าจริง

Σ คือ เครื่องหมายผลรวม (Summation)

4. ผลการดำเนินงาน

จากผลการสร้างแบบจำลองด้วยอัลกอริทึมจำนวน 4 อัลกอริทึม แสดงผลลัพธ์ดังภาพที่ 1



ภาพที่ 1: แบบจำลองสถานการณ์การโจมตีทางไซเบอร์

จากภาพที่ 1 แสดงถึงการสร้างแบบจำลองด้วยโปรแกรม RapidMiner Studio โดยนำชุดข้อมูลเข้าสู่กระบวนการของโปรแกรม และเริ่มจากทำการคัดเลือกแอททริบิวต์ที่ต้องการ จากนั้นทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วนตามที่ระบุไว้ในขั้นตอนการเตรียมข้อมูลก่อนหน้า เพื่อทำการทดสอบกับ 4 อัลกอริทึม จากนั้นทำการตรวจสอบและคัดเลือกอัลกอริทึมที่มีค่าเฉลี่ยความแม่น และค่าเฉลี่ยความเที่ยงที่สูดสุด เพื่อนำไปเป็นอัลกอริทึมหลักในการพัฒนาแบบจำลอง จากการจำลองสถานการณ์ และประเมินการโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ ทางความมั่นคงปลอดภัยไซเบอร์ เพื่อแสดงผลลัพธ์การเปรียบเทียบประสิทธิภาพของอัลกอริทึม

4.1 ผลทดสอบการเปรียบเทียบประสิทธิภาพของอัลกอริทึม

จากการสร้างแบบจำลองด้วยอัลกอริทึมของการจำลองสถานการณ์ และประเมินการโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ ทางความมั่นคงปลอดภัยไซเบอร์ มีผลการเปรียบเทียบประสิทธิภาพของทั้ง 4 อัลกอริทึม แสดงดังตารางที่ 3 และ 4

ตารางที่ 3: ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
Deep Learning	**96.92**	**96.96**	**96.88**	**96.92**
Neural Networks	95.86	95.83	95.91	95.87
AutoMLP	93.14	93.34	92.99	93.16
Multilayer Perceptrons	47.43	23.73	49.94	32.17

ตารางที่ 4: ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม (ต่อ)

Algorithms	MAE (%)	RMSE (%)
Deep Learning	12.3	35.1
Neural Networks	**3.1**	**17.6**
AutoMLP	5.6	23.8
Multilayer Perceptrons	52.6	72.5

จากตารางที่ 3 และ 4 ผลการเปรียบเทียบค่าทางสถิติการจำลองสถานการณ์ และประเมินการโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคมโดยใช้การเรียนรู้เชิงลึกเพื่อเสริมสร้างความตระหนักรู้ทางความมั่นคงปลอดภัยไซเบอร์ ร่วมกับชุดข้อมูล URL Dataset (ISCX-URL2016) ของอัลกอริทึมที่ใช้คืออัลกอริทึม Deep Learning มีค่าเฉลี่ยความแม่น (96.92%) และค่าเฉลี่ยความเที่ยง (96.96%) เป็นเปอร์เซ็นต์สูงที่สุด และอัลกอริทึม Neural Networks โดยค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (3.1%) และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (17.6%) เป็นค่าต่ำที่สุด

จากผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึมทั้งหมดนั้น จะช่วยเสริมสร้างความตระหนักรู้ในด้านความมั่นคงปลอดภัยไซเบอร์ เช่น การตรวจจับและป้องกันการหลอกลวง การตรวจสอบและปิดกั้นการใช้งานที่ไม่ปกติ ทั้งนี้จะช่วยเพิ่มประสิทธิภาพในการตรวจจับวิเคราะห์ และป้องกันภัยคุกคามจากการหลอกลวงและมัลแวร์ได้เป็นอย่างดี อันส่งผลให้บุคลากรและองค์กรมีความมั่นคงปลอดภัยไซเบอร์มากขึ้น

5. สรุปผล

งานวิจัยนี้ผู้วิจัยได้ทำการจำลองสถานการณ์ และประเมินการโจมตีทางไซเบอร์ด้วยวิธีวิศวกรรมสังคม โดยใช้การเรียนรู้เชิงลึก จากชุดข้อมูล URL Dataset (ISCX-URL2016) ด้วยอัลกอริทึมจำนวน 4 อัลกอริทึม คือ Deep Learning, Neural Networks, Automated MLP (AutoMLP) และ Multilayer Perceptrons จากโปรแกรม RapidMiner Studio ผลการวิจัยพบว่าอัลกอริทึม Deep Learning และ อัลกอริทึม Neural Networks มีประสิทธิภาพในการจำลองสูงที่สุด คือ ค่าเฉลี่ยความแม่น (Accuracy) และ ค่าเฉลี่ยความเที่ยง (Precision) รวมไปถึง ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAE) และค่าเฉลี่ยรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เพื่อตรวจจับและป้องกันมัลแวร์และฟิชชิ่งจากการใช้อัลกอริทึม Deep Learning และ อัลกอริทึม Neural Networks เป็นเครื่องมืออัลกอริทึมที่มีประสิทธิภาพในการป้องกันข้อมูลและระบบคอมพิวเตอร์จากการโจมตีในรูปแบบต่าง ๆ โดยการใช้แบบจำลองเหล่านี้ในการคาดเดาและจำแนกแนวโน้มของการโจมตีที่เป็นอันตราย

เอกสารอ้างอิง

[1] A. Yichiet, G. Ming-Lee, W. Nur Haliza Binti Abdul and G. Goh Hock “Social Engineering Attack Classifications on Social Media Using Deep Learning” *Computers, Materials & Continua, 2023*.

[2] กรสิริณัฐ โรจนวรรณ์ และวิยุดา เพชรจิรโชติกุล, “การเลือกคุณลักษณะที่เหมาะสมสำหรับการแทนค่าข้อมูลสูญหายของข้อมูลอนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล,” *Princess of Naradhiwas University Journal*, 2021.

[3] L. Muyang, et al, “AutoMLP: Automated MLP for Sequential Recommendations,” *In Proceedings of the ACM Web Conference*, pp. 1190-1198, 2023.

[4] อนุวัฒน์ เปพาทย์ และวงกต ศรีอุไร, “การพยากรณ์การออกกลางคันของนักศึกษามหาวิทยาลัยจากการปรับปรุงด้วยการคัดเลือกคุณลักษณะร่วมกับวิธีโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น,” *วารสารวิทยาศาสตร์และวิทยาศาสตร์ศึกษา (JSSE)*, 2022.

[5] สุรชัย ฉัตรเฉลิมพันธุ์ และเทอดพงษ์ แดงสี, “การเสริมความตระหนักรู้เท่าทันภัยทางไซเบอร์ของบุคลากรในองค์กรกรณีการจำลองการโจมตีด้วยฟิชชิง,” *วารสารของคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธนบุรี*, 2563.

[6] I. Almomani, et al., “An automated vision-based deep learning model for efficient detection of android malware attacks” *IEEE Access*, Vol. 10, pp. 2700-2720, 2022.

[7] พงศ์พนธ์ ภาวศุทธิ์, “สาเหตุเชิงลึกของการถูกโจมตีด้วยวิธีการทางวิศวกรรมสังคมของกลุ่มเจเนอเรชั่นวาย ในเขตกรุงเทพมหานครและปริมณฑล,” *วารสารระบบสารสนเทศด้านธุรกิจ*, ปีที่ 5 ฉบับที่ 1, 2561.

[8] J. Clayton, et al, “Towards Detecting and Classifying Malicious URLs Using Deep Learning”. *Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl*, pp. 31-48, 2020.

[9] A. Al-Haija and M. Al-Fayoumi, “An intelligent identification and classification system for malicious uniform resource locators (URLs),” *Neural Computing and Applications*, 2023.

[10] H. Timothy O, “Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not,” *Geoscientific Model Development*, 2022