การตรวจจับการบุกรุกระบบเครือข่ายด้วยวิธีการเรียนรู้แบบไม่มีผู้สอน

NETWORK INTRUSION DETECTION USING AN UNSUPERVISED LEARNING APPROACH

วิทยา ทศพิทักษ์กุล

WITAYA TOSPITAKKUL

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์
ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง
คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2566

KMITL-2023-SC-M-017-026

NETWORK INTRUSION DETECTION USING AN UNSUPERVISED LEARNING APPROACH

WITAYA TOSPITAKKUL

AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE
IN DATA SCIENCE AND ANALYTICS
KMITL DIGITAL ANALYTICS AND INTELLIGENCE CENTER SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2023

KMITL-2023-SC-M-017-026

หัวข้อการค้นคว้าอิสระ การตรวจจับการบุกรุกระบบเครือข่ายด้วยวิธีการเรียนรู้แบบไม่มีผู้สอน
ชื่อนักศึกษา นายวิทยา ทศพิทักษ์กุล
รหัสประจำตัว 64605102
ปริญญา วิทยาศาสตรมหาบัณฑิต (วิทยาการข้อมูลและการวิเคราะห์) ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง
พ.ศ. 2566
อาจารย์ที่ปรึกษาการค้นคว้าอิสระ รองศาสตราจารย์ ดร.ละออ บุญเกษม

บทคัดย่อ

การค้นคว้าอิสระนี้มีวัตถุประสงค์เพื่อศึกษาวิธีการจำแนกการบุกรุกระบบเครือข่ายโดยการใช้การเรียนแบบไม่มีผู้สอน ซึ่งในการศึกษานี้จะใช้วิธีการรู้เชิงลึกประเภทตัวเข้ารหัสอัตโนมัติมาประยุกต์ใช้ในการจำแนกผ่านการใช้ค่าการสูญเสียการสร้างใหม่ โดยที่ชุดข้อมูลที่ใช้ในการศึกษาเป็นชุดข้อมูลที่ได้จากฐานข้อมูล UNSW-NB15 ซึ่งเป็นฐานข้อมูลที่ได้รับความนิยมจากผู้ที่ศึกษาด้านระบบตรวจจับการบุกรุก และสุดท้ายจะนำตัวแบบที่ได้มาวัดประสิทธิภาพโดยใช้ ความแม่น ความเที่ยง ค่าเรียกคืน และคะแนนเอฟ

จากการทดลองพบว่าตัวเข้ารหัสอัตโนมัติมีความสามารถในการจำแนกข้อมูลการบุกรุกจากฐานข้อมูล UNSW-NB15 ได้ โดยได้ค่าความแม่นสูงถึง 82% และจะได้ค่าความแม่นสูงขึ้นไปถึง 92% สำหรับชุดข้อมูลที่ไม่มีการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode ถึงแม้ว่าในการทดสอบจะพบว่าตัวแบบจะมีความสามารถในการจำแนกลดลงถ้าชุดข้อมูลที่ใช้สำหรับฝึกสอนมีอัตราส่วนของข้อมูลประเภทการบุกรุกมากขึ้น แต่เนื่องจากในระบบจริง อัตราส่วนข้อมูลปกติจะมีมากกว่าข้อมูลบุกรุก ทำให้การเรียนรู้แบบไม่มีผู้สอนสามารถทำให้ต้นทุนในการสร้างตัวแบบสำหรับตรวจจับการบุกรุกระบบเครือข่ายลดลง และมีโอกาสถูกนำไปใช้ในระบบต่าง ๆ มากขึ้น

คำสำคัญ : การตรวจจับการบุกรุกระบบเครือข่าย การเรียนรู้แบบไม่มีผู้สอน ชุดข้อมูล UNSW-NB15 ตัวเข้ารหัสอัตโนมัติ

Independent Study Title	Network Intrusion Detection Using an Unsupervised Learning Approach
Student Name	Mr. Witaya Tospitakkul
Student ID	64605102
Degree	Master of Science (Data Science and Analytics) KMITL-Digital Analytics and Intelligence Center
Year	2023
Independent Study Advisor	Assoc.Prof.Dr. Laor Boongasame

Abstract

The purpose of this independent study is to study the methodology of network intrusion detection using an unsupervised learning approach. In this study, a deep learning approach called autoencoder will be applied for intrusion detection by utilizing reconstruction error values. The dataset used in this study was obtained from the UNSW-NB15 database, which is popular among researchers in the field of intrusion detection systems. Finally, the performance of the model will be evaluated using accuracy, precision, recall, and F1-score metrics.

From the experiments, the findings of the study indicate that the autoencoder model can classify intrusion data from the UNSW-NB15 database with a high accuracy up to 82% and up to 92% for datasets without intrusions of Fuzzers, Reconnaissance, and Shellcode types. However, it also revealed that model's classification performance decreases when the training dataset contains a higher proportion of intrusion data, Nevertheless, in the real-world systems, the proportion of normal data is typically higher than that of intrusion data. This allows unsupervised learning to significantly reduce the cost of building intrusion detection models and increases their potential for deployment in various systems.

Keywords : Network Intrusion Detection, Unsupervised Learning, UNSW-NB15 Dataset, Autoencoder

กิตติกรรมประกาศ

การค้นคว้าอิสระฉบับนี้สำเร็จลงได้ด้วยความกรุณาของรศ.ดร.ละออ บุญเกษม อาจารย์ที่ปรึกษาการค้นคว้าอิสระ ผู้ซึ่งกรุณาให้ความรู้ คำแนะนำ และคำปรึกษา ตลอดจนถึงช่วยตรวจทานแก้ไขปรับปรุงข้อบกพร่องต่าง ๆ จนการศึกษาครั้งนี้สำเร็จเสร็จสิ้นลุล่วงด้วยดี ผู้เขียนขอขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณหัวหน้างานและเพื่อนร่วมงานทุกท่าน ที่คอยให้ความช่วยเหลือ ความสะดวกและให้กำลังใจด้วยดีเสมอมา รวมไปถึงอาจารย์และเพื่อน ๆ สาขาวิทยาการข้อมูลและการวิเคราะห์ทุกท่านที่ให้ความรู้และคอยช่วยเหลืออยู่เสมอ และสุดท้ายขอกราบขอบพระคุณมารดาของผู้เขียนที่เข้าใจและคอยให้กำลังใจในการศึกษาอยู่เสมอ

ท้ายสุดนี้หากมีข้อผิดพลาดประการใด ผู้เขียนต้องขออภัยเป็นอย่างสูงในความผิดพลาดและผู้เขียนหวังว่าการค้นคว้าอิสระฉบับนี้ จะมีประโยชน์สำหรับผู้ที่สนใจเพื่อใช้เป็นแนวทางในการศึกษาการตรวจจับการบุกรุกระบบเครือข่ายในวิธีและแนวทางอื่น ๆ ต่อไป

นายวิทยา ทศพิทักษ์กุล

สารบัญ

สารบัญตาราง

สารบัญรูป

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปีพ.ศ. 2564 ที่ผ่านมาจำนวนคนที่ใช้งานอินเตอร์เน็ตมีจำนวนที่เพิ่มขึ้นมากจากข้อมูลที่เก็บโดย DataReportal ซึ่งได้เก็บข้อมูลการใช้อินเตอร์เน็ตของโลก แสดงให้เห็นว่ามีคนใช้งานอินเตอร์เน็ตเพิ่มขึ้น 1 ล้านคนในทุก ๆ วัน ทำให้ในปีพ.ศ. 2565 มีผู้ใช้งานอินเตอร์เน็ตรวมกันมากถึง 4,950 ล้านคน ยิ่งจำนวนผู้ใช้งานอินเตอร์เน็ตมีมากขึ้น ยิ่งทำให้เกิดการบุกรุกระบบคอมพิวเตอร์เป็นจำนวนที่เพิ่มมากขึ้นและความซับซ้อนมากยิ่งขึ้นตามไปด้วย ซึ่งในปีที่ผ่านมามีมูลค่าความเสียหายที่เกิดจากอาชญากรรมไซเบอร์รวมทั้งหมดมากถึง 6.9 พันล้านดอลลาร์สหรัฐ ทำให้เทคโนโลยีหลาย ๆ อย่างเช่นไฟร์วอลล์ การเข้ารหัส และการยืนยันตัวตนได้ถูกนำมาใช้งานเพื่อป้องกันระบบเครือข่ายคอมพิวเตอร์

อย่างไรก็ตาม เทคโนโลยีเหล่านั้นมีความสามารถไม่เพียงพอและบางครั้งไม่สามารถป้องกันการโจมตีระบบเครือข่ายคอมพิวเตอร์ได้ ดังนั้นระบบตรวจจับการบุกรุก (Intrusion Detection System) ที่มีประสิทธิภาพได้ถูกนำเสนอโดยนักวิจัยมากมาย ซึ่งวิธีการตรวจจับได้ถูกแบ่งเป็นสองประเภทคือ 1. ใช้พื้นฐานของลักษณะเฉพาะ (Signature Based) และ 2. การตรวจจับความผิดปกติ (Anomaly Detection) สำหรับการตรวจจับวิธีแรกใช้พื้นฐานลักษณะเฉพาะจะมีความสามารถในการตรวจจับรูปแบบการโจมตีที่รู้จักซึ่งข้อมูลจะถูกเก็บไว้ในฐานข้อมูลดังนั้นจึงจำเป็นที่จะต้องปรับปรุงฐานข้อมูลให้มีรูปแบบที่ทันสมัยอยู่ตลอดเวลาเพื่อป้องกันการโจมตีรูปแบบใหม่ ๆ แต่สำหรับวิธีการตรวจจับวิธีที่สอง การตรวจจับความผิดปกตินั้นจะมีความสามารถในการตรวจจับรูปแบบการโจมตีที่ไม่เคยพบมาก่อนโดยใช้พื้นฐานทางสถิติ ทำให้มีการศึกษาวิธีการทางการเรียนรู้ของเครื่อง (Machine learning) สำหรับการตรวจจับการบุกรุกมามากกว่า 20 ปี แต่การศึกษาที่ผ่านมาจำนวนมาก ทำการศึกษาโดยการใช้การเรียนรู้แบบมีผู้สอน (Supervised learning) และใช้ชุดข้อมูลตัวอย่างที่มีคำตอบ (Label) ในการศึกษา อย่างไรก็ตามวิธีการดังกล่าว ก่อให้เกิดข้อจำกัดมากในการนำวิธีการเรียนรู้ของเครื่องไปใช้งานจริง เนื่องจากชุดข้อมูลของจริงนั้นมีจำนวนมหาศาลและไม่มีคุณลักษณะ (Feature) ที่เป็นคำตอบมาให้ ซึ่งเป็นเรื่องยากที่จะทำการระบุคำตอบด้วยตัวเองเพื่อที่จะนำไปใช้ในการเรียนรู้ของวิธีการเรียนรู้แบบมีผู้สอน เพราะฉะนั้นจึงเริ่มมีการศึกษาการใช้การเรียนรู้ของเครื่องแบบไม่มีผู้สอน (Unsupervised learning) มาประยุกต์ใช้ในระบบตรวจจับการบุกรุก

ในการค้นคว้าอิสระนี้ได้ศึกษาและนำเสนอวิธีการตรวจจับการบุกรุกระบบเครือข่ายโดยใช้การเรียนรู้เชิงลึกประเภทตัวเข้ารหัสอัตโนมัติ ซึ่งเป็นการเรียนรู้แบบไม่มีผู้สอนมาใช้เพื่อจำแนกข้อมูล โดยที่ชุดข้อมูลที่ใช้ในการทดลองจะเป็นข้อมูลจากฐาน UNSW-NB15 ซึ่งเป็นชุดข้อมูลที่มีความนิยมในการนำมาศึกษาทางด้านการตรวจจับการบุกรุก

1.2 วัตถุประสงค์ของงานวิจัย

1) ศึกษาวิธีการลดคุณลักษณะข้อมูลเพื่อเตรียมไปใช้จำแนก

2) ศึกษาการจำแนกพฤติกรรมการบุกรุกด้วยวิธีการเรียนรู้แบบไม่มีผู้สอน

3) เพื่อหาประสิทธิภาพของวิธีการตรวจจับการบุกรุกระบบเครือข่ายที่นำเสนอ

1.3 ขอบเขตของงานวิจัย

1) ชุดข้อมูลที่ใช้ในการทดลองมาจากฐานข้อมูล UNSW-NB15 ซึ่งจัดทำโดยองค์กร Cyber Range Lab of UNSW Canberra

2) ศึกษาเฉพาะวิธีการสร้างขั้นตอนวิธี (Algorithm) โดยที่ไม่มีการสร้างโปรแกรมประยุกต์ (Application)

3) สร้างต้นแบบของตัวแบบโดยใช้ภาษา Python โดยจะไม่มีส่วนติดต่อกับผู้ใช้ (User Interface)

4) ใช้เทคนิคการเรียนรู้แบบไม่มีผู้สอนในการสร้างตัวแบบสำหรับการจำแนก (Classification)

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1) ได้วิธีการตรวจจับการบุกรุกระบบเครือข่ายแบบใหม่ที่สามารถเรียนรู้จากชุดข้อมูลที่ไม่มีคำตอบได้

2) เพื่อเป็นแนวทางสำหรับการพัฒนาวิธีการตรวจจับการบุกรุกระบบเครือข่ายได้หลายรูปแบบมากขึ้น

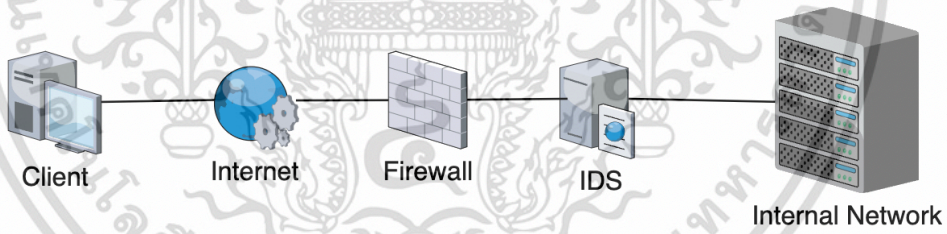
3) ได้ทราบประสิทธิภาพของการนำการเรียนรู้แบบไม่มีผู้สอนมาใช้ในการจำแนกข้อมูล

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ซึ่งจะประกอบไปด้วยพื้นฐานของระบบตรวจจับการบุกรุกระบบเครือข่าย การลดมิติของข้อมูลด้วยวิธีการวิเคราะองค์ประกอบหลัก โครงข่ายประสาทเทียม ตัวเข้ารหัสอัตโนมัติ และส่วนสุดท้ายคืองานวิจัยที่เกี่ยวกับระบบตรวจจับการบุกรุกเครือข่าย

2.1 ระบบตรวจจับการบุกรุกเครือข่าย (Network Intrusion Detection System)

ระบบตรวจจับการบุกรุก คือ ระบบตรวจจับสัญญาณของความผิดปกติต่าง ๆ ที่เกิดขึ้นในระบบที่อยู่ในขอบเขตที่ระบบนี้มีหน้าที่ตรวจสอบ โดยตัวโปรแกรมจะมีความสามารถในการตรวจจับสัญญาณของความผิดปกติที่เกิดขึ้นในระบบไม่ว่าจะเป็นภายในระบบคอมพิวเตอร์ ระบบปฏิบัติการ โปรแกรมที่รันอยู่ในเครื่อง การทำงานกับฐานข้อมูล หรือแม้แต่ข้อมูลที่วิ่งผ่านไปมาในเครือข่าย ซึ่งจะมีสถาปัตยกรรมดังรูปที่ 2.1



รูปที่ 2.1 แสดงสถาปัตยกรรมของระบบที่มีการใช้งานระบบตรวจจับการบุกรุกเครือข่าย

ระบบการตรวจจับการบุกรุกนี้ได้รับการศึกษาและปรับปรุงพัฒนาในหลายด้านอย่างต่อเนื่อง ตัวอย่างเช่น การพัฒนาวิธีการสำหรับใช้ในการตรวจจับการบุกรุกหรือการ ออกแบบสถาปัตยกรรมของระบบการตรวจจับการบุกรุกเพื่อรองรับและป้องกันระบบเครือข่ายคอมพิวเตอร์จากการโจมตีในรูปแบบต่าง ๆ ที่มีอยู่ในปัจจุบัน

2.1.1 ประเภทของการตรวจจับการบุกรุก

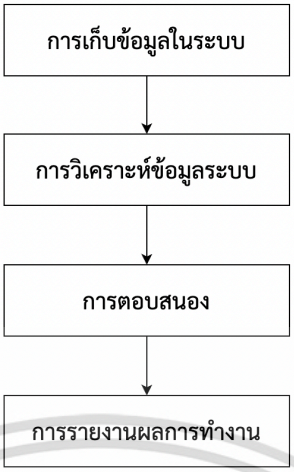
ระบบตรวจจับการบุกรุกแบ่งออกเป็น 2 รูปแบบ [1] ดังนี้

1. ระบบที่ตรวจหาการทำงานที่ผิดไปจากการทำงานปกติของระบบ (Anomaly Detection) คือ การหาเซตของการทำงานที่เป็นปกติย่อย ๆ ขึ้นมาแล้วนำมารวมกันเพื่อให้ระบบตรวจจับการบุกรุกทราบข้อมูลของเซตการทำงานที่เป็นปกติทั้งหมดในระบบ หลังจากนั้นเมื่อระบบตรวจจับการบุกรุกทำงาน ถ้าเกิดกรณีที่ระบบตรวจจับการบุกรุกตรวจจับการทำงานที่ไม่ได้อยู่ในเซตของการทำงานที่เป็นปกติ ระบบตรวจจับการบุกรุกจะแจ้งเตือนต่อผู้ดูแลระบบทันที สำหรับการสร้างขอบเขตของระบบนั้น อาจจะสร้างได้โดยการหาข้อมูลการทำงานที่เป็นปกติในระบบขึ้นมา โดยเอาข้อมูลการทำงานของผู้ใช้งานแต่ละคน เวลาที่มีการใช้งาน ทรัพยากรที่ผู้ใช้งานคนนั้น ๆ มักจะใช้บ่อย ๆ หรือแม้กระทั่งข้อมูลในระบบ หรือในเครือข่ายก็สามารถนำมาสร้างเป็นเซตของระบบได้เช่นกัน ในการหาเซตของการทำงานที่เป็นปกติทั้งหมดอาจจะเกิดการผิดพลาดขึ้นทำให้เกิดผลลัพธ์ที่เป็นค่าบวกเท็จหรือค่าลบเท็จขึ้นมาได้เช่นกัน

2. ระบบที่ตรวจหาการทำงานที่ไม่ควรเกิดขึ้นในระบบ (Misuse Detection หรือ Signature based) ซึ่งเป็นแนวความคิดที่ตรงข้ามกับระบบที่ตรวจหาการทำงานที่ผิดไปจากการทำงานปกติของระบบ (Anomaly Detection) คือ รูปแบบนี้จะใช้ข้อมูลของการทำงานที่ผิดปกติต่าง ๆ ที่เคยเกิดขึ้นมาแล้ว สร้างเป็นฐานข้อมูลของการทำงานที่ผิดปกติให้ระบบตรวจจับการบุกรุกจดจำไว้ และในการทำงานของตรวจจับการบุกรุกประเภทนี้จะนำข้อมูลที่อยู่ในระบบมาค้นหาในฐานข้อมูลว่ามีอยู่หรือไม่ ถ้าระบบตรวจจับการบุกรุกมีข้อมูลของการทำงานรูปแบบนั้น ๆ อยู่ ก็แสดงว่าเกิดความผิดปกติขึ้น แต่ในการรวบรวมข้อมูลนี้อาจรวมเอาการทำงานที่เป็นปกติเข้าไปด้วย ทำให้เกิดผลลัพธ์ที่เป็นค่าบวกเท็จหรือในบางกรณีที่ไม่ได้เก็บข้อมูลความผิดปกติเอาไว้ก็ทำให้เกิดกรณีผลลัพธ์ที่เป็นค่าลบเท็จได้เช่นกัน ซึ่งในการทำงานประเภทนี้จะมีข้อเสียคือจะไม่สามารถตรวจจับการบุกรุกชนิดใหม่ๆ ได้ เนื่องจากต้องมีข้อมูลของการบุกรุกอยู่ก่อนจึงจะตรวจจับได้

2.1.2 การทำงานของระบบตรวจจับการบุกรุก

ระบบตรวจจับการบุกรุกแต่ละแบบมีหน้าที่การทำงานที่แตกต่างกันออกไป บางตัวจะตรวจจับความผิดปกติในระบบเครือข่าย บางตัวจะตรวจจับความผิดปกติในระบบฐานข้อมูล แต่โดยการทำงานทั้งหมดจะสามารถแบ่งการทำงานของระบบตรวจจับการบุกรุกได้เป็น 4 ขั้นตอนดังรูปที่ 2.2



รูปที่ 2.2 แสดงขั้นตอนการทำงานของระบบตรวจจับการบุกรุก

ขั้นตอนที่ 1. การเก็บข้อมูลในระบบ

การเก็บข้อมูลในระบบที่ต้องการตรวจสอบสามารถแบ่งการเก็บข้อมูลออกเป็นกลุ่มต่าง ๆ ได้ 4 กลุ่มคือ 1. การเก็บข้อมูลในชั้นแอปพลิเคชันเพื่อนำมาตรวจสอบการทำงานของแอปพลิเคชันต่าง ๆ 2. การเก็บข้อมูลของการทำงานของเครื่องเพื่อนำมาตรวจสอบการทำงานของระบบของเครื่องที่ใช้งานอยู่ 3. การเก็บข้อมูลการเปลี่ยนแปลงข้อมูลในระบบเพื่อนำมาตรวจสอบว่าข้อมูลมีการเปลี่ยนแปลงอย่างไร และ 4. การเก็บข้อมูลเครือข่ายเพื่อนำมาตรวจสอบว่ามีการบุกรุกทางระบบเครือข่ายหรือไม่

ขั้นตอนที่ 2. การวิเคราะห์ข้อมูลระบบ

การวิเคราะห์ข้อมูลระบบสามารถแบ่งการทำงานตามรูปแบบการวิเคราะห์ข้อมูลได้ 2 รูปแบบคือ

1. การวิเคราะห์ในขณะเก็บข้อมูลทันที

ในการวิเคราะห์ข้อมูลที่ได้ในขณะที่เก็บข้อมูลนั้น ระบบจะจัดเก็บข้อมูลวิเคราะห์ข้อมูลและรายงานผลการวิเคราะห์ในช่วงเวลาเดียวกัน เมื่อเกิดข้อผิดพลาดขึ้นจะสามารถตอบสนองได้ทันท่วงที แต่ก็ขึ้นอยู่กับความเร็วในการวิเคราะห์ข้อมูลด้วย ถ้าข้อมูลมีความซับซ้อนมากก็จะใช้เวลามากตาม ซึ่งการทำงานที่รวดเร็วดังกล่าวก็ต้องแลกกับการใช้หน่วยความจำปริมาณมาก อีกทั้งการตอบสนองต่อการบุกรุกโดยอัตโนมัติอาจจะทำให้เกิดความเสียหายกับระบบมากกว่าเดิม เพราะว่าในบางครั้งการทำงานที่เร็วเกินไปของระบบ ทำให้เกิดความผิดพลาดในการวิเคราะห์ข้อมูลจนประมวลผลการทำงานปกติกลายเป็นการทำงานที่ผิดปกติ ระบบตรวจจับการบุกรุกที่ทำงานแบบนี้จึงเหมาะกับระบบที่มีข้อมูลที่ต้องพิจารณาน้อย ต้องการการรายงานอย่างรวดเร็วเมื่อผิดปกติ และข้อมูลที่นำมาวิเคราะห์นั้นต้องไม่ซับซ้อนมาก

2. การวิเคราะห์ข้อมูลภายหลังจากที่เก็บข้อมูล

การทำงานในลักษณะนี้เหมาะกับงานที่ไม่จำเป็นต้องตอบสนองทันทีเมื่อเกิดความผิดปกติขึ้น แต่ให้มีการบันทึกและรายงานว่าเกิดความผิดปกติขึ้น การทำงานจะใช้หน่วยความจำและการประมวลผลน้อยกว่าแบบแรก แต่จะใช้เนื้อที่ในการเก็บข้อมูลมากกว่าแบบแรก ข้อเสียของการทำงานแบบนี้คือ มักแก้ปัญหาที่เกิดขึ้นไม่ทัน เพราะกว่าจะทราบว่าเกิดปัญหาขึ้น ปัญหานั้นก็เกิดขึ้นนานมากแล้ว

ไม่ว่าจะเป็นการวิเคราะห์ในขณะเก็บข้อมูลหรือการวิเคราะห์ข้อมูลภายหลังจากที่เก็บข้อมูล ก็จะมีวิธีการวิเคราะห์ระบบที่เหมือนกันคือ ระบบที่ตรวจหาการทำงานที่ผิดไปจากการทำงานปกติของระบบ (Anomaly Detection) ส่วนในอีกรูปแบบหนึ่ง คือ ระบบที่ตรวจหาการทำงานที่ไม่ควรเกิดขึ้นในระบบ (Misuse Detection)

ขั้นตอนที่ 3. การตอบสนอง

เมื่อมีการตรวจพบว่ามีการบุกรุกเกิดขึ้นในระบบ สำหรับระบบตรวจจับที่ทำงานแบบทันทีจะมีการตอบสนองต่อการบุกรุกเพื่อไม่ให้เกิดความเสียหายหรือบรรเทาความเสียหายที่เกิดขึ้น แต่สำหรับระบบที่ทำงานเป็นแบบกลุ่มการตอบสนองอาจจะทำได้ไม่มากนัก เพราะการบุกรุกนั้นเกิดไปแล้วและความเสียหายก็เกิดขึ้นไปแล้ว การตอบสนองอาจจะอยู่ในรูปแบบการบรรเทาไม่ให้เกิดความเสียหายมากขึ้นเท่านั้น ซึ่งการตอบสนองต่อการบุกรุกนั้นจะแบ่งออกได้ 3 รูปแบบดังนี้

1. การเปลี่ยนแปลงสภาพของระบบ

การตอบสนองโดยการเปลี่ยนแปลงสภาพของระบบก็เพื่อแก้ปัญหาหรือลดความเสียหายที่จะเกิดขึ้น เช่น ตัดการเชื่อมต่อระหว่างระบบกับการบุกรุกออกจากกัน การตั้งค่าอุปกรณ์เครือข่ายหรือไฟร์วอลล์ไม่ให้มีการติดต่อกับระบบของการบุกรุก

2. การแก้ไขความผิดพลาดให้ถูก

เป็นการตอบสนองต่อปัญหาที่เกิดขึ้นแล้วในระบบ โดยปกติแล้วการบุกรุกมักเปลี่ยนแปลงค่าต่าง ๆ ในระบบ โดยเฉพาะเข้ามาทำการเปลี่ยนแปลงข้อมูลในระบบตรวจจับการบุกรุกเพื่อไม่ให้สามารถตรวจจับการบุกรุกได้ การแก้ไขระบบก็เพื่อให้ระบบดังกล่าวสามารถทำงานได้อย่างปกติ

3. การแจ้งเตือนผู้ดูแลระบบ

สุดท้ายเป็นการแจ้งเตือนผู้ดูแลระบบ โดยปกติมักแจ้งเตือนผู้ดูแลระบบทันทีเมื่อทำการวิเคราะห์ได้ว่ามีความผิดปกติเกิดขึ้น เพื่อให้ผู้ดูแลระบบรับรู้และสามารถแก้ไขระบบได้ทันท่วงที

ขั้นตอนที่ 4. การรายงานผลการทำงาน

เมื่อระบบตรวจจับการบุกรุกทำการวิเคราะห์ระบบและตรวจพบความผิดปกติในระบบ อาจจะมีการตอบสนองต่อความผิดปกตินั้นถ้าทำได้ จากนั้นระบบจะต้องมีการรายงานผลให้กับผู้ดูแลระบบทราบในรูปแบบต่าง ๆ โดยรายละเอียดของการรายงานผลนั้น จะบอกถึงช่องโหว่ในระบบ การแก้ไขปัญหาคร่าวๆ บางครั้งอาจจะมีรายละเอียดของความรู้พื้นฐานบางอย่างของระบบที่ทำให้เกิดการบุกรุกลักษณะนั้น ๆ ได้ การรายงานผลการทำงานนอกจากจะเป็นการรายงานต่อผู้ดูแลระบบเพื่อให้ทราบการทำงานหรือจุดอ่อนนั้นแล้ว ยังเป็นประโยชน์ต่อการวิเคราะห์สถานะของระบบและการวิเคราะห์ความปลอดภัยในระบบอีกด้วย

2.2 การวิเคราะห์องค์ประกอบหลัก (Principle Component Analysis หรือ PCA)

การวิเคราะห์องค์ประกอบหลัก [3] เป็นวิธีการที่ถูกนำมาใช้มากสำหรับการลดรูปมิติในการเรียนรู้ของเครื่อง ซึ่งจะนิยมใช้ในขั้นตอนการเตรียมข้อมูล ซึ่งการวิเคราะห์องค์ประกอบหลักจะลดรูปข้อมูลจากมิติที่สูงมามิติที่ต่ำกว่าโดยการใช้การฉายภาพเชิงเส้น (Linear Projection) เพื่อที่จะลดมิติของตัวแปรในขณะที่ยังความแปรปรวนของข้อมูลไว้ให้ได้มากที่สุด ซึ่งการทำงานจะแบ่งเป็น 4 ขั้นตอนดังนี้

1. ทำให้ชุดข้อมูลเป็นมาตรฐานดังสมการที่ 2.1

$$Z_{ij} = \frac{x_{ij} - \bar{x_j}}{s_j}, \qquad i = 1,2,\dots,m; j = 1,2,\dots,n \tag{2.1}$$

เมื่อ m คือ จำนวนของชุดข้อมูล
n คือ จำนวนของตัวแปรในชุดข้อมูล
x_{ij} คือ ค่าของข้อมูลของชุดข้อมูลที่ i และตัวแปรที่ j
$\bar{x_j}$ คือ ค่าเฉลี่ยของข้อมูลของข้อมูลตัวแปรที่ j
s_j คือ ค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลตัวแปรที่ j
Z_{ij} คือ ค่าของข้อมูลของชุดข้อมูลที่ i และตัวแปรที่ j ที่ถูกทำให้เป็นมาตรฐาน

2. หาเมทริกซ์สัมประสิทธิ์สหสัมพันธ์ ดังสมการที่ 2.2

$$R = \frac{Z^T Z}{m - 1} \tag{2.2}$$

เมื่อ Z คือ เมทริกซ์ที่ถูกทำให้เป็นค่ามาตรฐาน

Z^T คือ เมทริกซ์สลับเปลี่ยนที่ถูกทำให้เป็นค่ามาตรฐาน

m คือ จำนวนของชุดข้อมูล

3. หาค่าเฉพาะ (eigenvalue) และเวกเตอร์เฉพาะ (eigenvector) ที่สอดคล้อง ดังสมการที่ 2.3

$$a_1 = \begin{Bmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{Bmatrix}, a_2 = \begin{Bmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{Bmatrix}, \ldots, a_n = \begin{Bmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{Bmatrix} \qquad (2.3)$$

เมื่อ a_n คือ เวกเตอร์เฉพาะที่สอดคล้องกับองค์ประกอบหลักที่ n

α_{mn} คือ ค่าของค่าเฉพาะที่ m ในเวกเตอร์ฉพาะขององค์ประกอบหลักที่ n

4. คำนวณหาองค์ประกอบหลัก ดังสมการที่ 2.4

$$t_i = \alpha_{1i}Z_1 + \alpha_{2i}Z_2 + \ldots + \alpha_{ni}Z_n \ , \qquad i = 1, \ldots, n' \qquad (2.4)$$

เมื่อ n' คือ จำนวนขององค์ประกอบหลัก

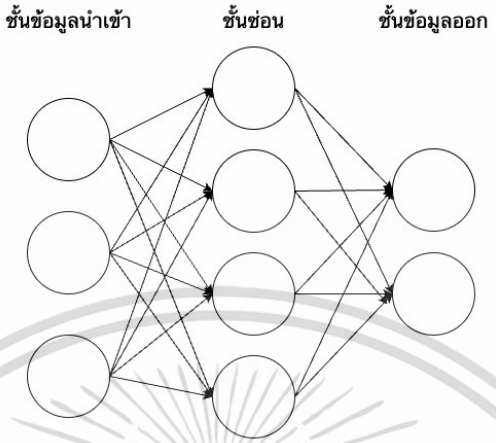
2.3 โครงข่ายประสาทเทียม (Artificial Neural Network)

โครงข่ายประสาทเทียม [4] คือตัวแบบทางคณิตศาสตร์ที่เป็นการจำลองการทำงานของสมองมนุษย์ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบและการอุปมานความรู้แบบเดียวกับสมองมนุษย์ โดยสามารถเรียนรู้ข้อมูลจากชุดข้อมูลที่มีอยู่แล้วในชุดข้อมูลฝึกสอน เพื่อใช้ทำนายข้อมูลในลักษณ์เดียวกันในชุดข้อมูลทดสอบ โดยมีรายละเอียดในส่วนต่าง ๆ [5] ดังนี้

2.3.1 โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า (Feed-Forward Neural Network)

โครงข่ายประสาทเทียมประเภทนี้ จะมีลำดับในการคำนวณและส่งต่อข้อมูลไปในทิศทางเดียว โดยโครงสร้างจะแบ่งออกเป็นลำดับชั้น โดยในแต่ละชั้นจะมีเพอร์เซ็ปตรอนจำนวนหนึ่งซึ่งไม่มีเส้นเชื่อมถึงกันภายในชั้นเดียวกัน แต่จะมีเส้นเชื่อมถึงเพอร์เซ็ปตรอนตัวอื่น ๆ ที่อยู่ในลำดับชั้นที่

ติดกันทั้งหมด โดยข้อมูลส่งออกของเพอร์เซ็ปตรอนในชั้นก่อนหน้า จะเป็นข้อมูลรับเข้าของเพอร์เซ็ปตรอนในชั้นปัจจุบัน ดังแสดงโครงสร้างได้ในรูปที่ 2.3



รูปที่ 2.3 แสดงโครงสร้างของโครงข่ายประสาทเทียม

และสามารถคำนวณหาค่าของผลลัพธ์ในชั้นถัดไปได้จากสมการดังต่อไปนี้

$$z_j^l = \sum_{k=1}^{n} w_{jk}^l a_k^{l-1} + b_j^l \tag{2.5}$$

$$a_j^l = g(z_j^l) \tag{2.6}$$

เมื่อ a_k^{l-1} คือ ผลลัพธ์ของเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้น l - 1

w_{jk}^l คือ น้ำหนักสำหรับเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l ที่มีเส้นเชื่อมมาจากเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นก่อนหน้า

b_j^l คือ ค่าไบแอสสำหรับเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l

g คือ ฟังก์ชันกระตุ้น

2.3.2 ฟังก์ชันกระตุ้น (Activation Function)

สำหรับข้อมูลส่งออกของแต่ละเพอร์เซ็ปตรอน จะมีการใช้ฟังก์ชันกระตุ้น g(z) ที่มีลักษณะแบบไม่ใช่ฟังก์ชันเชิงเส้น (Non-linear) เพื่อทำให้โครงข่ายประสาทเทียมมีความซับซ้อนและสามารถแก้ปัญหาได้หลากหลายมากยิ่งขึ้น โดยฟังก์ชันกระตุ้นที่นิยมใช้กัน มีดังต่อไปนี้

1. ฟังก์ชันซิกมอยด์ (Sigmoid Function) เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาอยู่ในช่วง 0 ถึง 1 ฟังก์ชันซิกมอยด์สามารถเขียนแทนด้วย σ ซึ่งคำนวณได้จากสมการต่อไปนี้

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{2.7}$$

2. ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ (Hyperbolic Tangent Function) เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาอยู่ในช่วง -1 ถึง 1 ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์สามารถเขียนแทนด้วย tanh ซึ่งคำนวณได้จากสมการต่อไปนี้

$$tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{2.8}$$

3. ฟังก์ชันเรกทิไฟด์เชิงเส้น (Rectified Linear Unit Function) เป็นฟังก์ชันที่เปลี่ยนค่าติดลบที่เข้ามาให้เป็น 0 ส่วนค่าอื่น ๆ คงเดิม ซึ่งคำนวณได้จากสมการต่อไปนี้

$$f(z) = \begin{cases} 0, & if\ z \leq 0 \\ z, & if\ z > 0 \end{cases} \tag{2.9}$$

4. ฟังก์ชันซอฟต์แม็กซ์ (Softmax Function) เป็นฟังก์ชันที่ให้ค่าผลลัพธ์ออกมาอยู่ในช่วง 0 ถึง 1 ซึ่งเป็นค่าที่แสดงความน่าจะเป็นของค่าที่นำเข้ามาแต่ละตัว โดยผลรวมของความน่าจะเป็นที่ได้จะมีค่าเป็น 1 ซึ่งคำนวณได้จากสมการต่อไปนี้

$$f(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{2.10}$$

2.3.3 ฟังก์ชันต้นทุน (Cost Function หรือ Loss Function)

เป็นฟังก์ชันที่แสดงถึงต้นทุนของโครงข่ายประสาทเทียม โดยในกระบวนการเรียนรู้ของโครงข่ายประสาทเทียมนั้น จะทำการปรับค่าน้ำหนักเพื่อที่จะลดค่าของฟังก์ชันต้นทุนเพื่อให้เข้าใกล้กับค่า 0 มากที่สุด ฟังก์ชันต้นทุนที่เป็นที่นิยมมีดังต่อไปนี้

1. ค่าเฉลี่ยความผิดพลาดกำลังสอง (Mean Square Error)

$$J = \frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{2.11}$$

เมื่อ y_i คือ คำตอบจริง ณ ข้อมูลที่ i

$\hat{y}_i$ คือ ผลลัพธ์ที่ทำนายได้จากข้อมูลที่ i

2. ค่าเฉลี่ยครอสเอนโทปีแบบทวิภาค (Binary Cross-entropy)

$$J = -\frac{1}{n}\sum_{i=1}^{n} y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - y_i) \quad (2.12)$$

เมื่อ y_i คือ คำตอบจริงของข้อมูลที่ i

$\hat{y}_i$ คือ ผลลัพธ์ที่ทำนายได้จากข้อมูลที่ i

3. ค่าติดลบลอการิทึมภาวะน่าจะเป็น (Negative Log Likelihood)

$$J = -\frac{1}{n}\sum_{i=1}^{n} y_i \log(\hat{y}_i) \quad (2.13)$$

เมื่อ y_i คือ คำตอบจริง ณ ข้อมูลที่ i

$\hat{y}_i$ คือ ผลลัพธ์ที่ทำนายได้จากข้อมูลที่ i

2.3.4 การหาค่าที่เหมาะสมที่สุด (Optimization)

เป็นวิธีการปรับปรุงอัตราการเรียนรู้ เพื่อให้สามารถลดค่าจากฟังก์ชันต้นทุนได้มากที่สุดในแต่ละรอบ เพื่อเพิ่มโอกาสไปยังจุดต่ำสุดทั้งหมด โดยใช้วิธีการปรับปรุงน้ำหนักของเส้นเชื่อมในโครงข่ายประสาทเทียม วิธีการปรับปรุงน้ำหนักที่ได้รับความนิยมมีดังต่อไปนี้

1. สโตแคสติกเกรเดียนเดสเซนท์ (Stochastic Gradient Descent)

เมื่อกำหนดให้ w แทนค่าพารามิเตอร์ ซึ่งเป็นน้ำหนักที่ต้องการจะปรับค่า α คือ อัตราการเรียนรู้ และ $\frac{\partial J}{\partial w}$ คือ เกรเดียนฟังก์ชันของต้นทุนเทียบกับ w โดยจะทำการปรับค่าด้วยสมการดังต่อไปนี้

$$w_t = w_{t-1} - \alpha\frac{\partial J}{\partial w_{t-1}} \quad (2.14)$$

เมื่อ α คือ อัตราการเรียนรู้

J คือ ฟังก์ชันต้นทุน

w_{t-1} คือ ค่าน้ำหนัก ณ จุด t-1

w_t คือ ค่าน้ำหนักใหม่ที่ได้

แต่ปัญหาที่อาจจะเจอในระหว่างการเรียนรู้คือ การติดอยู่ในโลคอลออพติมา (Local Optima) ดังนั้นจึงมีการนำโมเมนตัม (Momentum) มาใช้โดยมีจุดประสงค์เพื่อทำให้การเรียนรู้มีการลู่เข้าที่ดีขึ้นและหลีกเลี่ยงการติดอยู่ในโลคอลออพติมา โดยกำหนดให้ v แทนค่าความเร็ว ซึ่งมีการปรับค่าพร้อมกับ w และ γ แทนค่าสัมประสิทธิ์ของโมเมนตัม สามารแสดงการเรียนรู้ได้เป็นสมการดังนี้

$$v_t = \gamma v_{t-1} + \alpha \frac{\partial J}{\partial w_{t-1}} \tag{2.15}$$

$$w_t = w_{t-1} - v_t \tag{2.16}$$

เมื่อ α คือ อัตราการเรียนรู้
J คือ ฟังก์ชันต้นทุน
w_{t-1} คือ ค่าน้ำหนัก ณ จุด t-1
v_{t-1} คือ ค่าของความเร็ว ณ จุด t-1
γ คือ สัมประสิทธิ์ของโมเมนตัม
v_t คือ ค่าของความเร็ว ณ จุด t
w_t คือ ค่าน้ำหนักใหม่ที่ได้

2. วิธีเกรเดียนที่ปรับตัวได้ (Adaptive Gradient Method)

เป็นวิธีที่จะมีการปรับอัตราการเรียนรู้ได้ด้วยตัวเองจากค่าเริ่มต้นที่กำหนด โดยการปรับค่าของอัตราการเรียนรู้นั้นจะมีการนำค่าเกรเดียนในอดีตมาใช้ กำหนดให้ g_t แทนเกรเดียนที่เวลา t ซึ่งสามารถแสดงการเรียนรู้ได้เป็นสมการดังนี้

$$g_t = \frac{\partial J}{\partial w_{t-1}} \tag{2.17}$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{\sum_{k=1}^{t} g_i^2}} g_t \tag{2.18}$$

เมื่อ J คือ ฟังก์ชันต้นทุน
w_{t-1} คือ ค่าน้ำหนัก ณ จุด t-1
g_t คือ เกรเดียนของฟังก์ชันต้นทุน เทียบกับค่าน้ำหนัก ณ จุด t-1
g_i คือ เกรเดียนของฟังก์ชันต้นทุน เทียบกับค่าน้ำหนัก ณ จุด i-1

α คือ อัตราการเรียนรู้

w_t คือ ค่าน้ำหนักใหม่ที่ได้

3. อาร์เอ็มเอสพรอพ (RMSProp)

เป็นวิธีที่มีการเก็บค่าเกรเดียนของครั้งก่อนหน้าไว้เพื่อที่จะนำมาใช้ในรอบการเรียนรู้ปัจจุบัน โดยนำไปปรับปรุงอัตราส่วนของอัตราการเรียนรู้ โดยนอกเหนือจากการใช้ g_t แล้วยังมีการใช้ $MeanSquare_t$ สำหรับเก็บค่าเฉลี่ยของเกรเดียน และให้ v แทนอัตราการใช้เกรเดียนของอดีตในการเรียนรู้ สามารถแสดงการเรียนรู้ได้เป็นสมการต่อไปนี้

$$g_t = \frac{\partial J}{\partial w_{t-1}} \tag{2.19}$$

$$MeanSquare_t = \gamma MeanSquare_{t-1} + (1-\gamma) g_t^2 \tag{2.20}$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{MeanSquare_t}} g_t \tag{2.21}$$

เมื่อ J คือ ฟังก์ชันต้นทุน

w_{t-1} คือ ค่าน้ำหนัก ณ จุด t-1

g_t คือ เกรเดียนของฟังก์ชันต้นทุน เทียบกับค่าน้ำหนัก ณ จุด t-1

γ คือ สัมประสิทธิ์ของโมเมนตัม

$MeanSquare_{t-1}$ คือ ค่าเฉลี่ยกำลังสองของเกรเดียน ณ จุด t-1

$MeanSquare_t$ คือ ค่าเฉลี่ยกำลังสองของเกรเดียน ณ จุด t

α คือ อัตราการเรียนรู้

w_t คือ ค่าน้ำหนักใหม่ที่ได้

2.3.5 การแพร่กระจายย้อนกลับ (Backpropagation)

เนื่องจากเกรเดียนที่ได้จากการหาค่าความผิดพลาดสุดท้ายจากฟังก์ชันต้นทุนนั้นมีไว้ให้ลำดับชั้นสุดท้ายในโครงข่ายประสาทเทียมเท่านั้น ดังนั้น หากต้องการจะทำการหาค่าเกรเดียนสำหรับปรับค่าของ w ของเพอร์เซ็ปตรอนในลำดับชั้นก่อนหน้า จะต้องใช้วิธีการแพร่กระจายย้อนกลับ โดยจะสามารถเขียนเป็นสมการได้ดังนี้

$$\delta_j^l = \frac{\partial J}{\partial z_j^l} = \frac{\partial J}{\partial a_j^l}\frac{\partial a_j^l}{\partial z_j^l} = \frac{\partial J}{\partial a_j^l} = g'(z_j^l) \tag{2.22}$$

เมื่อ δ_j^l คือ ค่าความผิดพลาดของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l

J คือ ฟังก์ชันต้นทุน

z_j^l คือ ค่าที่คำนวณได้ก่อนผ่านฟังก์ชันกระตุ้น g ของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l

a_j^l คือ ผลลัพธ์ของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้น l

g' คือ ฟังก์ชันกระตุ้น

สำหรับการหาค่า $\frac{\partial J}{\partial a_j^l}$ นั้น ในลำดับชั้นสุดท้ายสามารถคำนวณหาได้โดยตรงจากฟังก์ชันต้นทุนที่เลือกใช้ ส่วนในลำดับชั้นก่อนหน้าจะต้องหาโดยวิธีการแพร่กระจายย้อนกลับ โดยจะคล้ายกับการป้อนไปข้างหน้า เพียงแต่เป็นการกลับทิศทางกันเท่านั้น กำหนด m คือ จำนวนเพอร์เซ็ปตรอนในลำดับชั้นที่ l + 1 สามารถคำนวณได้จากสมการดังนี้

$$\frac{\partial J}{\partial a_j^l} = \sum_{k=1}^{m} \frac{\partial J}{\partial z_j^{l+1}} \frac{\partial z_j^{l+1}}{\partial a_j^l} = \sum_{k=1}^{m} \delta_k^{l+1} w_{kj}^{l+1} \qquad (2.23)$$

เมื่อ J คือ ฟังก์ชันต้นทุน

a_j^l คือ ผลลัพธ์ของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้น l

z_j^{l+1} คือ ค่าที่คำนวณได้ก่อนผ่านฟังก์ชันกระตุ้น g ของเพอร์เซ็ปตรอนตัวที่ j1 ในลำดับชั้นที่ l +1

δ_j^{l+1} คือ ค่าความผิดพลาดของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l + 1

m คือ จำนวนเพอร์เซ็ปตรอนในลำดับชั้นที่ l + 1

w_{kj}^{l+1} คือ ค่าน้ำหนักสำหรับเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นที่ l + 1 ที่มีเส้นเชื่อมมาจากเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นก่อนหน้า

จากนั้น เมื่อคำนวณค่าความผิดพลาดของแต่ละลำดับชั้นได้ ก็สามารถหาค่าความผิดพลาดเทียบกับน้ำหนักและค่าไบแอส ได้จากสมการดังนี้

$$\frac{\partial J}{\partial w_{jk}^l} = \frac{\partial J}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{jk}^l} = \delta_j^l a_k^{l-1} \qquad (2.24)$$

เมื่อ J คือ ฟังก์ชันต้นทุน

w_{jk}^{l} คือ คือ ค่าน้ำหนักสำหรับเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l ที่มีเส้นเชื่อมมาจากเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นก่อนหน้า

z_{j}^{l} คือ ค่าที่คำนวณได้ก่อนผ่านฟังก์ชันกระตุ้น g ของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l

δ_{j}^{l+1} คือ ค่าความผิดพลาดของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l + 1

a_{k}^{l-1} คือ ผลลัพธ์ของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้น l - 1

ในกรณีที่ใช้สโตแคสติกเกรเดียนเดสเซนท์ การปรับปรุงค่าน้ำหนักจะทำโดย

$$w_{jk,t}^{l} = w_{jk,t-1}^{l} - \alpha a_{k,t}^{l-1}\delta_{j,t}^{l} \quad (2.25)$$

เมื่อ $w_{jk,t}^{l}$ คือ ค่าน้ำหนักสำหรับเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l ที่มีเส้นเชื่อมมาจากเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นก่อนหน้า ณ จุด t

$w_{jk,t-1}^{l}$ คือ ค่าน้ำหนักสำหรับเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l ที่มีเส้นเชื่อมมาจากเพอร์เซ็ปตรอนตัวที่ k ในลำดับชั้นก่อนหน้า ณ จุด t – 1

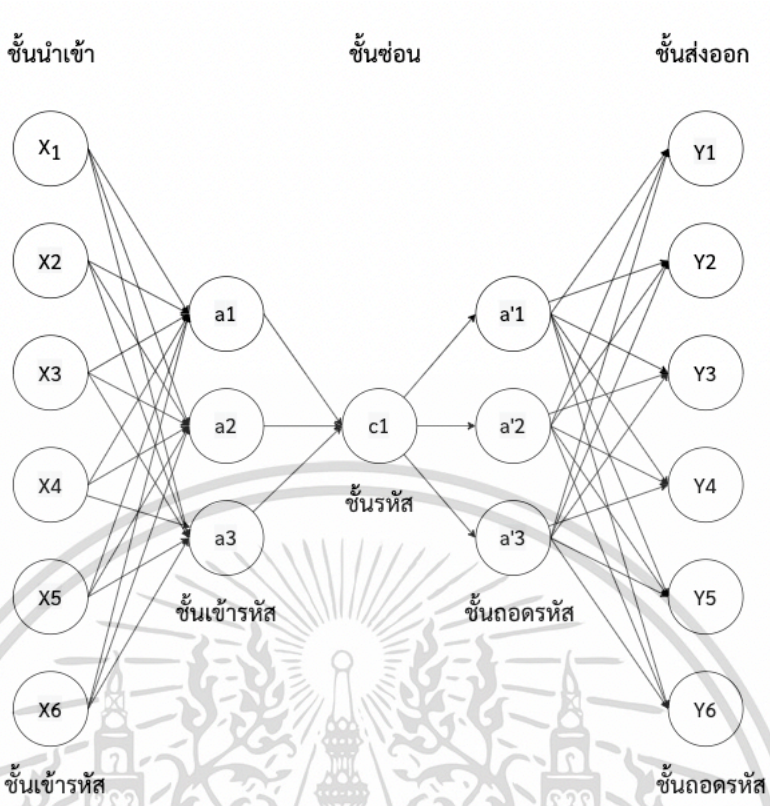
α คือ อัตราการเรียนรู้

$a_{k,t}^{l-1}$ คือ ผลลัพธ์ของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้น l – 1 ณ จุด t

$\delta_{j,t}^{l}$ คือ ค่าความผิดพลาดของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้นที่ l ณ จุด t

2.4 ตัวเข้ารหัสอัตโนมัติ (Autoencoder)

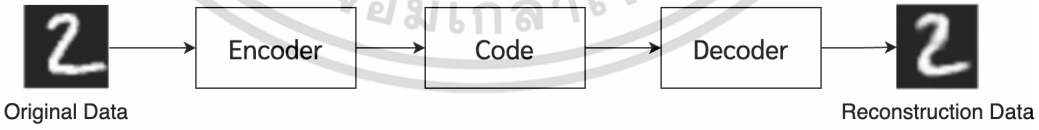
ตัวเข้ารหัสอัตโนมัติ [6] คือ ขั้นตอนวิธีพื้นฐานสำหรับนำไปประกอบเพื่อสร้างโครงข่ายประสาทเทียมชนิดหนึ่ง โดยการใช้งานตัวเข้ารหัสอัตโนมัติจะมีลักษณะคล้ายคลึงกับโครงข่ายประสาทเทียมทั่วไปที่มีชั้นนำเข้า ชั้นซ่อน และชั้นส่งออก โดยที่แตกต่างไปคือโครงข่ายประสาทเทียมจะเป็นการเรียนรู้แบบมีผู้สอน แต่ตัวเข้ารหัสอัตโนมัติจะเป็นการเรียนรู้แบบไม่มีผู้สอน กล่าวคือ เป็นเพียงการเรียนรู้ที่ต้องการให้ข้อมูลส่งออกมีลักษณะใกล้เคียงกับข้อมูลนำเข้า จึงเหมือนกับเป็นการฝึกข้อมูลนำเข้า *x* ให้เข้ารหัสเพื่อให้ได้ชั้นซ่อน $a = f(x)$ และสามารถถอดรหัสออกมาได้เป็น $y = g(a)$ ดังรูปที่ 2.4



รูปที่ 2.4 แสดงโครงสร้างของโครงข่ายประสาทเทียมชนิดตัวเข้ารหัสอัตโนมัติ

ดังนั้นตัวแบบจะพยายามเรียนรู้ความสัมพันธ์ของชุดข้อมูลเพื่อที่จะสามารถบีบอัดหรือเข้ารหัสข้อมูลได้อย่างมีประสิทธิภาพ และยังเรียนรู้ที่จะถอดรหัสข้อมูลกลับออกมาจากชุดข้อมูลที่ถูกบีบอัดแล้วให้มีความใกล้เคียงกับต้นฉบับมากที่สุด ซึ่งการเรียนรู้ของตัวแบบจะใช้การแพร่กลับเพื่อที่จะลดค่าการสูญเสียการสร้างใหม่ให้ได้น้อยที่สุด (Reconstruction Loss)

โครงสร้างของตัวเข้ารหัสอัตโนมัติประกอบไปด้วย 3 ส่วนหลักดังรูปที่ 2.5



รูปที่ 2.5 แสดงตัวอย่างลำดับการทำงานและผลลัพธ์ของตัวเข้ารหัสอัตโนมัติ

โดยในแต่ละส่วนมีความหมายดังนี้

1. ตัวเข้ารหัส (Encoder) เป็นชั้นที่จะเรียนรู้การลดขนาดมิติและบีบอัดข้อมูลต้นฉบับไปยังชั้นรหัส
2. ชั้นรหัส (Code) เป็นชั้นที่จะประกอบไปด้วยข้อมูลที่ถูกบีบอัดแล้ว ซึ่งจะมีมิติน้อยที่สุด

3. ตัวถอดรหัส (Decoder) เป็นชั้นที่จะเรียนรู้การสร้างข้อมูลกลับมาจากข้อมูลที่ถูกบีบอัดแล้วให้ใกล้เคียงต้นฉบับมากที่สุด ซึ่งสามารถวัดความใกล้เคียงได้จากค่าการสูญเสียการสร้างใหม่ระหว่างข้อมูลที่สร้างกลับมาและข้อมูลต้นฉบับ

2.5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการศึกษานี้ทั้ง 6 ชิ้นจะสามารถถูกแบ่งงานวิจัยเป็น 2 ประเภทดังนี้ 1. การนำเอาการเรียนรู้แบบไม่มีผู้สอนมาใช้งานในงานด้านการตรวจจับความผิดปกติ 2. การปรับปรุงประสิทธิภาพของการทำการจำแนก ดังที่แสดงในตารางที่ 2.1 และตารางที่ 2.2 ตามลำดับ

ตารางที่ 2.1 แสดงงานวิจัยที่เกี่ยวข้องกับการนำเอาการเรียนรู้แบบไม่มีผู้สอนมาใช้งานในงานด้านการตรวจจับความผิดปกติ

ชื่องานวิจัย	วิธีที่นำเสนอ
Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction [7]	นำตัวเข้ารหัสอัตโนมัติมาประยุกต์ใช้ในงานตรวจจับความผิดปกติ
Unsupervised Learning Approach for Network Intrusion Detection System Using Autoencoders [8]	นำตัวเข้ารหัสอัตโนมัติมาประยุกต์ใช้ในงานระบบตรวจจับการบุกรุก
Unsupervised Clustering Approach for Network Anomaly Detection [11]	นำการเรียนรู้แบบไม่มีผู้สอนมาประยุกต์ใช้ในงานระบบตรวจจับการบุกรุก

ตารางที่ 2.2 แสดงงานวิจัยที่เกี่ยวข้องกับการปรับปรุงประสิทธิภาพของการทำการจำแนก

ชื่องานวิจัย	วิธีที่นำเสนอ
Dimensionality Reduction and Visualization of Network Intrusion Detection Data [9]	เปรียบเทียบประสิทธิภาพของ PCA และ ตัวเข้ารหัสอัตโนมัติในการลดมิติของข้อมูลก่อนที่จะนำไปฝึกสอนตัวแบบ
A hybrid Intrusion Detection System based on Sparse autoencoder and Deep Neural Network [10]	ประยุกต์ใช้วิธีผสมเทคนิคการทำวิศวกรรมข้อมูลและโครงข่ายประสาทเทียมเชิงลึกมาปรับปรุงประสิทธิภาพของการจำแนก

ตารางที่ 2.2 แสดงงานวิจัยที่เกี่ยวข้องกับการปรับปรุงประสิทธิภาพของการทำการจำแนก (ต่อ)

ชื่องานวิจัย	วิธีที่นำเสนอ
An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks [3]	ประยุกต์ใช้วิธีลดจำนวนมิติข้อมูลเพื่อปรับปรุงประสิทธิภาพของการจำแนก

โดยรายละเอียดของงานวิจัยแต่ละกลุ่มตามตารางข้างต้นมีรายละเอียดอย่างย่อดังนี้

2.5.1 การนำเอาการเรียนรู้แบบไม่มีผู้สอนมาใช้งานในงานด้านการตรวจจับความผิดปกติ

1. งานวิจัยของ M. Sakurada และ T. Yairi [7] ได้นำเสนอการใช้ตัวเข้ารหัสอัตโนมัติเข้ามาประยุกต์ใช้ในการตรวจจับความผิดปกติข้อมูลการทำงานของยานอวกาศซึ่งมีจำนวณมิติของข้อมูลที่สูง โดยการใช้การลดมิติของข้อมูล ซึ่งได้มีการเปรียบเทียบกับวิธีวิเคราะห์องค์ประกอบหลัก (Linear PCA) และ วิธีวิเคราะห์องค์ประกอบหลักแบบใช้เคอร์เนล (Kernel PCA) ผลการทดลองปรากฏว่า ตัวเข้ารหัสอัตโนมัติมีประสิทธิภาพสูงสามารถแยกข้อมูลที่ผิดปกติได้ดีและใช้เวลาในการคำนวณต่ำกว่าเมื่อเปรียบเทียบกับการวิเคราะห์องค์ประกอบหลัก ซึ่งการวัดประสิทธิภาพจะใช้ค่า AUC เป็นตัวเปรียบเทียบโดยใช้ชุดข้อมูล Sat-A, Sat-B ได้ผลลัพธ์ดังตารางที่ 2.3

ตารางที่ 2.3 แสดงค่า AUC ของวิธีการทั้ง 4 แบบกับชุดข้อมูล Lorenz, Sat-A และ Sat-B

	LPCA	AE	dAE	KPCA
Lorenz	0.5104	0.6473	0.7011	0.7045
Sat-A	0.8852	0.8847	0.9354	0.8862
Sat-B	0.9764	0.9763	0.8355	0.7689

2. งานวิจัยของ H. Choi และคณะ [8] ได้นำเสนอระบบตรวจจับการบุกรุกระบบเครือข่ายโดยการใช้ตัวเข้ารหัสอัตโนมัติทั้ง 4 รูปแบบคือ ตัวเข้ารหัสอัตโนมัติแบบพื้นฐาน (Basic Autoencoder), ตัวเข้ารหัสอัตโนมัติประเภทลดสัญญาณรบกวน (Denoising Autoencoder), ตัวเข้ารหัสอัตโนมัติแบบซ้อน (Stacked Autoencoder) และ ตัวเข้ารหัสอัตโนมัติแบบแปรผัน (Variational Autoencoder) โดยผลการทดลองปรากฏว่าประสิทธิภาพที่ได้จากตัวเข้ารหัสอัตโนมัติทุกรูปแบบมีประสิทธิภาพสูงโดยมีความแม่นสูงสุดถึง 91.70% เมื่อเทียบกับวิธีการเรียนรู้แบบไม่มี

ผู้สอนแบบการจัดกลุ่มที่มีความแม่น 80% และงานวิจัยนี้ยังแสดงให้เห็นอีกว่าตัวเข้ารหัสอัตโนมัติสามารถจัดการกับชุดข้อมูลที่มีจำนวนข้อมูลในคลาสที่แตกต่างกันมาก

3. งานวิจัยของ I. Syarif และคณะ [11] ได้อธิบายประโยชน์ของการใช้การเรียนรู้แบบไม่มีผู้สอนในการตรวจจับการบุกรุกระบบเครือข่ายจากการเปรียบเทียบประสิทธิภาพของวิธีการจัดกลุ่มเมื่อถูกนำมาใช้กับการตรวจจับความผิดปกติ จากการทดลองโดยใช้ชุดข้อมูลที่มีการบุกรุกที่ไม่รู้จักปรากฏว่า วิธีการตรวจจับแบบใช้กฎมีค่าความแม่นที่ 63.97% แต่วิธีการจัดกลุ่มที่ได้ถูกนำมาเปรียบเทียบมีความแม่นสูงมากสุดที่ 80.15% ดังตารางที่ 2.4 แต่วิธีการจัดกลุ่มยังมีอัตราผลบวกเท็จที่สูง จึงต้องมีการศึกษาเพิ่มเติมเกี่ยวกับการปรับปรุงประสิทธิภาพของตัวแบบ

ตารางที่ 2.4 แสดงค่าความแม่นของวิธีการจัดกลุ่ม

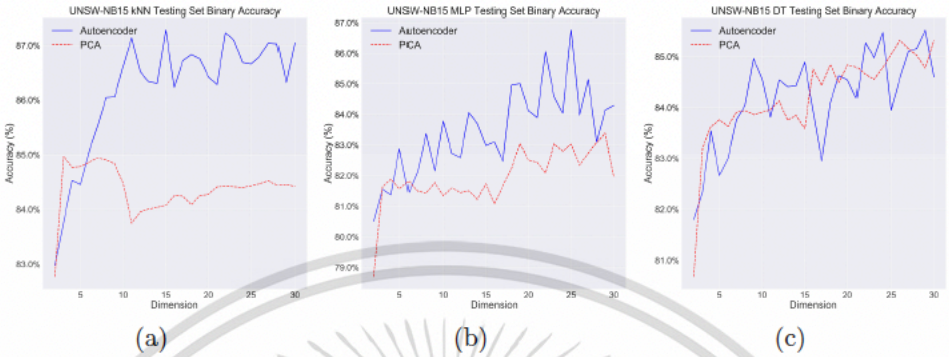
	ความแม่น	**อัตราผลบวกเท็จ**
k-Means	57.81%	22.95%
Improved k-Means	65.40%	21.52%
k-Medoids	76.71%	21.83%
EM Clustering	78.06%	20.74%
Distance-based outlier	80.15%	21.14%

งานวิจัยในกลุ่มนี้แสดงให้เห็นถึงความสามารถในการนำตัวเข้ารหัสอัตโนมัติซึ่งเป็นการเรียนรู้แบบไม่มีผู้สอนเข้ามาใช้งานในการจำแนกข้อมูลที่มีความผิดปกติในขอบเขตระบบตรวจจับการบุกรุกระบบเครือข่ายอีกทั้งยังสามารถจัดการกับชุดข้อมูลที่มีจำนวนข้อมูลในคลาสที่แตกต่างกันมากได้ดี แต่งานวิจัยทั้งสามชิ้นเป็นเพียงแค่การประยุกต์ใช้ตัวเข้ารหัสอัตโนมัติเพียงอย่างเดียวโดยที่ไม่ได้มีการใช้เทคนิคอื่น ๆ เพิ่มเติมจึงอาจจะยังไม่ได้ผลลัพธ์ที่ดีที่สุด

2.5.2 การปรับปรุงประสิทธิภาพของการทำการจำแนก

1. งานวิจัยของ Y. Chow และ W. Susilo [9] ได้นำเสนอถึงการเปรียบเทียบประสิทธิภาพของเทคนิคการลดมิติของข้อมูลเมื่อนำชุดข้อมูลที่ลดมิติแล้วไปใช้ทำการจำแนกต่อ 2 วิธี คือ 1. การวิเคราะห์องค์ประกอบหลัก และ 2. ตัวเข้ารหัสอัตโนมัติ โดยการทดลองจะใช้ชุดข้อมูลของระบบตรวจจับการบุกรุกระบบเครือข่าย 2 ฐานข้อมูลคือ NSL KDD และ UNSW-NB15 โดยขั้นตอนการเปรียบเทียบประสิทธิภาพจะทำการลดมิติของข้อมูลแล้วจึงนำชุดข้อมูลที่ได้ไปจำแนก โดยผลการ

ทดลองปรากฏว่า ชุดข้อมูลที่ถูกลดมิติของข้อมูลด้วยตัวเข้ารหัสอัตโนมัติมีความสามารถในการสร้างตัวแบบที่มีประสิทธิภาพที่ดีกว่าการใช้ชุดข้อมูลที่ลดมิติข้อมูลด้วยการวิเคราะห์องค์ประกอบหลักดังรูปที่ 2.6



รูปที่ 2.6 แสดงการเปรียบเทียบความแม่นยำของวิธีที่งานวิจัยนำเสนอ[9]

2. งานวิจัยของ K. Narayana Rao และคณะ [10] ได้นำเสนอวิธีการแบบผสม 2 ขั้นตอนสำหรับการตรวจจับการบุกรุก โดยในขั้นตอนแรกนั้นจะนำตัวเข้ารหัสอัตโนมัติมาใช้สำหรับการทำวิศวกรรมคุณลักษณะ และขั้นตอนที่สองคือการนำโครงข่ายประสาทเทียมเชิงลึกมาใช้ทำนายและจำแนก โดยผลการทดลองปรากฏว่า SAE-DNN มีความแม่นยำสูงสุดเมื่อเทียบกับวิธีอื่น ๆ บนชุดข้อมูล KDDCup99 NSL-KDD และ UNSW-NB15 ดังที่แสดงในตารางที่ 2.5 – 2.7

ตารางที่ 2.5 แสดงประสิทธิภาพของ SAE-DNN เมื่อเทียบกับวิธีอื่น ๆ บนชุดข้อมูล KDDCup99

วิธีการ	ความแม่นยำ	อัตราตรวจจับ	F1	FPR
LSTM-RNN	96.93	98.88	NA	10.04
RBM-DBN	97.16	NA	NA	0.48
CNN-GRU	98.1	97.6	98.8	NA
Multi-scale CNN	94.11	93.21	NA	2.18
CFA	91.98	91.00	NA	3.917
AE+DBN	92.10	92.20	NA	NA
SAE-DNN	99.03	99.48	99.14	1.55

ตารางที่ 2.6 แสดงประสิทธิภาพของ SAE-DNN เมื่อเทียบกับวิธีอื่น ๆ บนชุดข้อมูล NSL-KDD

วิธีการ	ความแม่นยำ	อัตราตรวจจับ	F1	FPR
SAE-SVM	80.48	NA	NA	NA
S-NDAE	85.42	85.42	87.37	14.58
ID-CVAE	80.10	80.10	79.08	8.18
SAVER-DNN	89.36	95.98	90.08	4.70
SSAE-SVM	99.35	99.01	NA	0.13
SAE-DNN	99.71	99.72	99.74	0.30

ตารางที่ 2.7 แสดงประสิทธิภาพของ SAE-DNN เมื่อเทียบกับวิธีอื่น ๆ บนชุดข้อมูล UNSW-NB15

วิธีการ	ความแม่นยำ	อัตราตรวจจับ	F1	FPR
TSDL	89.13	NA	NA	0.749
Random tree + NB Tree	89.24	83.90	NA	NA
CASCADE-ANN	86.40	86.74	NA	13.1
SAVAER-DNN	93.01	91.94	93.54	5.67
SAE-DNN	99.98	99.99	99.98	0.17

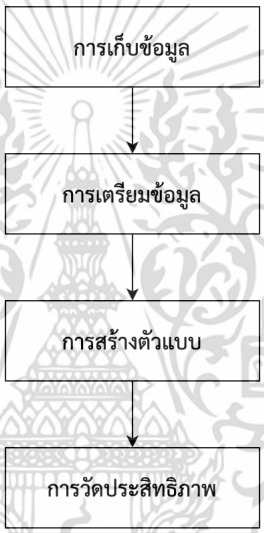
3. งานวิจัยของ Y. Xiao และคณะ [11] ได้นำเสนอวิธีการนำการลดมิติข้อมูลและโครงข่ายประสาทเทียมแบบคอนโวลูชันมาใช้เพื่อปรับปรุงประสิทธิภาพของตัวแบบสำหรับระบบตรวจจับการบุกรุกระบบเครือข่าย โดยวิธีการลดมิติข้อมูลได้นำเอาวิธีการสองแบบมาทดสอบได้แก่ตัวเข้ารหัสอัตโนมัติและ PCA ผลการทดลองปรากฏว่าตัวแบบที่งานวิจัยนำเสนอมีความแม่นมากขึ้นและใช้เวลาการทำงานลดลงเมื่อเปรียบเทียบกับวิธีในอดีต

งานวิจัยในกลุ่มนี้แสดงให้เห็นว่าการทำวิศวกรรมคุณลักษณะ (Feature Engineering) เช่น การลดมิติของข้อมูลก่อนที่จะนำไปฝึกสอนตัวแบบ ทำให้ประสิทธิภาพของตัวแบบมีความแม่นยำเพิ่มมากขึ้น โดยที่งานวิจัยของ K. Narayana Rao และคณะ ยังแสดงให้เห็นเพิ่มเติมว่าเมื่อนำเทคนิคการทำวิศวกรรมคุณลักษณะเข้ามาผสมกับการใช้โครงข่ายประสาทเทียมแบบลึก สามารถทำให้

ประสิทธิภาพของโครงข่ายประสาทเทียมแบบลึกดีขึ้น แต่การทดลองทุกงานนั้น ในส่วนของการจำแนกยังคงใช้ตัวแบบประเภทการเรียนรู้แบบมีผู้สอน

บทที่ 3
วิธีการดำเนินงานวิจัย

ในบทนี้จะกล่าวถึงวิธีการดำเนินงานซึ่งเป็นขั้นตอนการปฏิบัติงานทั้งหมดในการศึกษานี้ โดยที่การศึกษานี้คือการศึกษาการประยุกต์นำเอาการเรียนรู้เชิงลึกประเภทตัวเข้ารหัสอัตโนมัติซึ่งเป็นการเรียนรู้แบบที่ไม่มีผู้สอน มาใช้จำแนกหาข้อมูลที่บ่งชี้ถึงการบุกรุกในชุดข้อมูลการจราจรของระบบเครือข่าย UNSW-NB15 โดยที่ขั้นตอนดำเนินงานวิจัยแบ่งออกเป็น 4 ขั้นตอนหลัก ได้แก่ การเก็บข้อมูล การเตรียมข้อมูล การสร้างตัวแบบ และการวัดประสิทธิภาพของตัวแบบดังรูปที่ 3.1



รูปที่ 3.1 ขั้นตอนในการดำเนินงานวิจัย

3.1 ชุดข้อมูล (Dataset)

ชุดข้อมูลที่ผู้วิจัยใช้คือชุดข้อมูล UNSW-NB15 [2] เป็นชุดข้อมูลที่ถูกพัฒนาโดย Australian Centre for Cyber Security โดยการผสมผสานระหว่างข้อมูลปกติของจริงและข้อมูลการโจมตีที่ถูกสังเคราะห์ขึ้นมา โดยที่ชุดข้อมูลนี้จะประกอบไปด้วยการโจมตี 9 ประเภทได้แก่ DoS, Analysis, Generic, Fuzzers, Back-doors, Exploits, Shellcode, Reconnaissance และ Worms ซึ่งจะประกอบไปด้วยตัวอย่างข้อมูลเรียนรู้ทั้งหมด 175,341 ตัวอย่างและตัวอย่างข้อมูลทดสอบทั้งหมด 82,332 ตัวอย่าง ซึ่งแต่ละชุดข้อมูลจะมีคุณลักษณะทั้งหมด 49 ตัวดังที่แสดงในตารางที่ 3.1 และจะมีฉลากคำตอบที่จะใช้ระบุว่าการจราจรของระบบเครือข่ายนี้เป็นแบบปกติหรือแบบไม่ปกติ โดยที่การแจกแจงของตัวอย่างจะมีค่าดังตารางที่ 3.2 การศึกษาครั้งนี้ผู้วิจัยจะรวมชุดข้อมูลทั้งชุดข้อมูลสำหรับการเรียนรู้และชุดข้อมูลสำหรับการทดสอบมาเป็นชุดข้อมูลเดียวกัน

ตารางที่ 3.1 แสดงข้อมูลคุณลักษณะของชุดข้อมูล UNSW-NB15

หลักที่	ชื่อ	ชนิด	รายละเอียด
1	srcip	Nominal	ไอพีเครื่องต้นทาง
2	sport	Integer	พอร์ตเครื่องต้นทาง
3	dstip	Nominal	ไอพีเครื่องปลายทาง
4	dsport	Integer	พอร์ตเครื่องปลายทาง
5	proto	Nominal	โปรโตคอลการทำธุรกรรม
6	state	Nominal	สถานะของโปรโตคอล เช่น ACC, CON, ECO
7	dur	Float	ระยะเวลารวมทั้งหมด
8	sbytes	Integer	ไบต์ธุรกรรมจากต้นทางถึงปลายทาง
9	dbytes	Integer	ไบต์ธุรกรรมจากปลายทางถึงต้นทาง
10	sttl	Integer	ค่าทีทีแอลของต้นทางถึงปลายทาง
11	dttl	Integer	ค่าทีทีแอลของปลายทางถึงต้นทาง
12	sloss	Integer	แพ็คเก็ตต้นทางที่สูญหาย
13	dloss	Integer	แพ็คเก็ตปลายทางที่สูญหาย
14	service	Nominal	บริการ เช่น http, ftp, smtp
15	Sload	Float	ค่าบิตต่อวินาทีของต้นทาง
16	Dload	Float	ค่าบิตต่อวินาทีของปลายทาง
17	Spkts	Integer	จำนวนแพ็คเก็ตต้นทางถึงปลายทาง
18	Dpkts	Integer	จำนวนแพ็คเก็ตปลายทางถึงต้นทาง
19	swin	Integer	ค่าขนาดช่องทางทีซีพีที่โฆษณาต้นทาง
20	dwin	Integer	ค่าขนาดช่องทางทีซีพีที่โฆษณาปลายทาง
21	stcpb	Integer	ค่าลำดับพื้นฐานของทีซีพีต้นทาง
22	dtcpb	Integer	ค่าลำดับพื้นฐานของทีซีพีปลายทาง
23	smeansz	Integer	ค่าเฉลี่ยขนาดของแพ็คเก็ตที่ถูกส่งโดยต้นทาง

ตารางที่ 3.1 แสดงข้อมูลคุณลักษณะของชุดข้อมูล UNSW-NB15 (ต่อ)

หลักที่	ชื่อ	ชนิด	รายละเอียด
24	dmeansz	Integer	ค่าเฉลี่ยขนาดของแพ็คเก็ตที่ถูกส่งโดยปลายทาง
25	trans_depth	Integer	ความลึกของท่อของการเชื่อมต่อของเอชทีทีพี
26	res_bdy_len	Integer	ขนาดของข้อมูลก่อนบีบอัดที่ถูกส่งจากตัวบริการ
27	Sjit	Float	ความช้าในการรับแพ็คเก็ตตต้นทาง
28	Djit	Float	ความช้าในการรับแพ็คเก็ตตปลายทาง
29	Stime	Timestamp	เวลาเริ่มการบันทึก
30	Ltime	Timestamp	เวลาสิ้นสุดการบันทึก
31	Sintpkt	Float	ระยะเวลามาถึงของอินเตอร์แพ็คเก็ตต้นทาง
32	Dintpkt	Float	ระยะเวลามาถึงของอินเตอร์แพ็คเก็ตปลายทาง
33	tcprtt	Float	เวลาครบรอบของการเชื่อมต่อทีซีพี
34	synack	Float	ระยะเวลาระหว่างแพ็คเก็ต SYN และ SYN_ACK
35	ackdat	Float	ระยะเวลาระหว่างแพ็คเก็ต SYN_ ACK และ SYN
36	is_sm_ips_ports	Binary	มีค่า 1 ถ้าไอพีและพอร์ตของต้นทางและปลายทางเหมือนกัน
37	ct_state_ttl	Integer	จำนวนของ state ตามช่วงของค่าทีทีแอล
38	ct_flw_http_mthd	Integer	จำนวนของกระแสที่เป็นกระบวนการเอชทีทีพี
39	is_ftp_login	Binary	มีค่า 1 ถ้าเอฟทีพีถูกใช้งานผ่านรหัสผ่าน
40	ct_ftp_cmd	Integer	จำนวนของคำสั่งที่ใช้ผ่านเอฟทีพี
41	ct_srv_src	Integer	จำนวนของการเชื่อมต่อที่ใช้บริการและที่อยู่ต้นทางเดียวกันจาก 100 การเชื่อมต่อล่าสุด

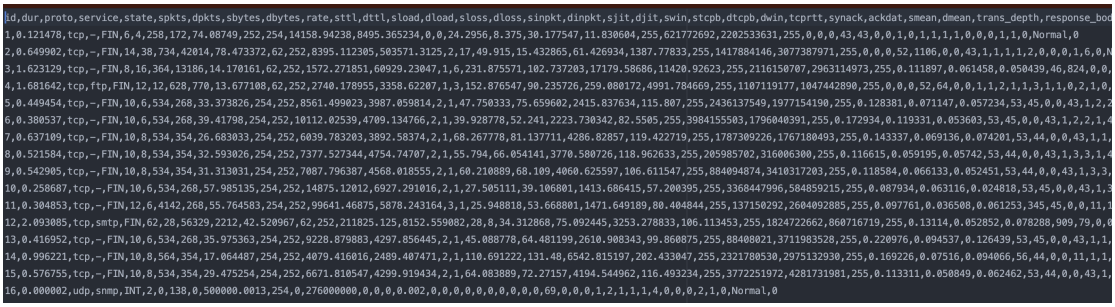
ตารางที่ 3.1 แสดงข้อมูลคุณลักษณะของชุดข้อมูล UNSW-NB15 (ต่อ)

หลักที่	ชื่อ	ชนิด	รายละเอียด
42	ct_srv_dst	Integer	จำนวนของการเชื่อมต่อที่ใช้บริการและที่อยู่ปลายทางเดียวกันจาก 100 การเชื่อมต่อล่าสุด
43	ct_dst_ltm	Integer	จำนวนของการเชื่อมต่อที่มีที่อยู่ปลายทางเดียวกันจาก 100 การเชื่อมต่อล่าสุด
44	ct_src_ ltm	Integer	จำนวนของการเชื่อมต่อที่มีที่อยู่ต้นทางเดียวกันจาก 100 การเชื่อมต่อล่าสุด
45	ct_src_dport_ltm	Integer	จำนวนของการเชื่อมต่อที่มีที่อยู่ต้นทางและพอร์ตปลายทางเดียวกันจาก 100 การเชื่อมต่อล่าสุด
46	ct_dst_sport_ltm	Integer	จำนวนของการเชื่อมต่อที่มีที่อยู่ปลายทางและพอร์ตต้นทางเดียวกันจาก 100 การเชื่อมต่อล่าสุด
47	ct_dst_src_ltm	Integer	จำนวนของการเชื่องต่อที่มีที่อยู่ต้นทางและที่อยู่ปลายทางเดียวจาก 100 การเชื่อมต่อล่าสุด
48	attack_cat	Nominal	ชื่อกลุ่มการโจมตี
49	label	Binary	จะมีค่า 0 ถ้าเป็นปกติ และมีค่า 1 ถ้าเป็นการโจมตี

ตารางที่ 3.2 แสดงการแจกแจงของข้อมูลที่มีในฐานข้อมูล UNSW-NB15

ประเภท	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ
ปกติ	56,000	37,000
โจมตี	119,341	45,332
รวม	175,341	82,332

โดยที่ชุดข้อมูลนี้จะมีรูปแบบข้อมูลเป็นแบบค่าที่ถูกแบ่งด้วยจุลภาค (CSV) ดังรูปที่ 3.2



รูปที่ 3.2 ตัวอย่างชุดข้อมูล UNSW-NB15

3.2 การเตรียมข้อมูล (Data Preparation and Preprocessing)

ในขั้นตอนนี้ผู้วิจัยได้แบ่งการดำเนินงานเป็น 3 ขั้นตอนดังรูปที่ 3.3

รูปที่ 3.3 แสดงขั้นตอนการเตรียมข้อมูล

3.2.1 การเข้ารหัสข้อมูลเชิงกลุ่ม (Categorical Data Encoding)

ชุดข้อมูลมีคุณลักษณะบางอย่างที่มีชนิดเชิงกลุ่ม ดังนั้นเพื่อที่จะสามารถนำคุณลักษณะเหล่านี้ไปเรียนรู้ผ่านตัวแบบได้ ในการศึกษานี้เลือกวิธีการสร้างตัวแปรหุ่น (Dummy Variable) มาใช้สำหรับเข้ารหัสคุณลักษณะเชิงกลุ่ม โดยตัวอย่างของการเข้ารหัสได้แสดงตามตารางที่ 3.3

ตารางที่ 3.3 แสดงตัวอย่างผลลัพธ์ของการเข้ารหัสด้วยวิธีการสร้างตัวแปรหุ่นของคุณลักษณะ state ที่มีค่า CON

Feature ที่สร้างขึ้นใหม่	ค่าของข้อมูล
state_CLO	0
state_CON	1
state_ECO	0
state_FIN	0
state_INT	0
state_PAR	0
state_REQ	0

ตารางที่ 3.3 แสดงตัวอย่างผลลัพธ์ของการเข้ารหัสด้วยวิธีการสร้างตัวแปรหุ่นของคุณลักษณะ state ที่มีค่า CON (ต่อ)

Feature ที่สร้างขึ้นใหม่	ค่าของข้อมูล
state_RST	0
state_URN	0
state_no	0

3.2.2 การทำให้ค่าเป็นมาตรฐาน (Standardization)

ช่วงขอบเขตของข้อมูลในชุดข้อมูลมีความแตกต่างกันอย่างมาก ทำให้มีความจำเป็นในการปรับช่วงขอบเขตของข้อมูลแต่ละคุณลักษณะให้อยู่ในช่วงเดียวกัน เพื่อให้เหมาะสมกับการนำไปประมวลผลต่อ ในการศึกษานี้เลือกวิธีการทำให้เป็นมาตรฐานแซด (Z-Score Normalization) มาใช้ปรับขอบเขตของข้อมูล โดยตัวอย่างของการทำให้ค่าเป็นมาตรฐานได้แสดงตามตารางที่ 3.4

ตารางที่ 3.4 แสดงตัวอย่างผลลัพธ์ของการทำให้ค่าเป็นมาตรฐาน

ชื่อคุณลักษณะ	ค่าต้นฉบับ	ค่าหลังจากทำให้เป็นมาตรฐานแล้ว
dur	0.121478	-0.188346
spkts	6	-0.101342
dpkts	4	-0.129612
sbytes	258	-0.047849
dbytes	172	-0.097232
rate	74.08749	-0.56865

3.2.3 การลดมิติข้อมูล (Dimensionality Reduction)

ชุดข้อมูลมีจำนวนคุณลักษณะเป็นจำนวนมาก ซึ่งบางคุณลักษณะอาจจะไม่มีข้อมูลที่เป็นประโยชน์ในการวิเคราะห์และจะส่งผลกระทบต่อความถูกต้องของการประมวลผล นอกจากนี้ยังทำให้สิ้นเปลืองทรัพยากรในการประมวลผล ขั้นตอนการลดมิติข้อมูลจึงมีความสำคัญในการเตรียมข้อมูล ในการศึกษานี้เลือกวิธีการการวิเคราะห์องค์ประกอบหลักมาใช้ลดมิติของชุดข้อมูล และกำหนดให้ลดจำนวนมิติของข้อมูลให้เหลือเพียง 128 ตัว

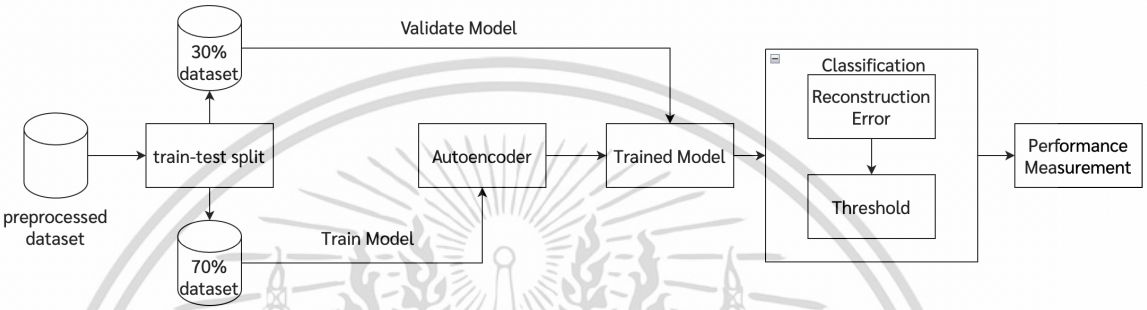
3.3 การสร้างตัวแบบ (Modeling)

ในการศึกษานี้จะสร้างตัวแบบสำหรับการจำแนกโดยการใช้การเรียนรู้แบบไม่มีผู้สอนซึ่งเป็นขั้นตอนวิธีการเรียนรู้แบบที่ไม่ต้องระบุฉลากคำตอบในชุดข้อมูล โดยที่การศึกษานี้จะเลือกใช้การ

เรียนรู้เชิงลึกประเภทตัวเข้ารหัสอัตโนมัติมาใช้ เนื่องจากชุดข้อมูลการจราจรของระบบเครือข่ายโดยปกติจะมีอัตราส่วนของข้อมูลการบุกรุกที่น้อยมากเมื่อเทียบกับข้อมูลการใช้ระบบเครือข่ายแบบปกติ แต่ชุดข้อมูลประเภทนี้จะมีจำนวนข้อมูลเป็นจำนวนมาก ทำให้ถ้าต้องมาระบุคำตอบสำหรับทุกข้อมูลที่จะใช้ฝึกสอนจะสิ้นเปลืองทรัพยากรมาก ผู้วิจัยจึงมีแนวคิดว่าการประยุกต์นำตัวเข้ารหัสอัตโนมัติมาฝึกสอนด้วยชุดข้อมูลประเภทนี้ จะสามารถเข้าช่วยแก้ปัญหาชุดข้อมูลที่ไม่มีฉลากคำตอบมาให้ได้

3.3.1 การออกแบบขั้นตอนการฝึกสอนตัวแบบ

ขั้นตอนการฝึกสอนตัวแบบจะถูกแบ่งออกเป็น 3 ขั้นตอนหลักดังรูปที่ 3.4



รูปที่ 3.4 แสดงขั้นตอนการสร้างตัวแบบ

โดยที่ในแต่ละขั้นตอนจะมีการทำงานดังนี้

1. การแบ่งชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบ
   เป็นขั้นตอนที่จะแบ่งชุดข้อมูลเป็น 2 ชุดคือ ชุดข้อมูลฝึกฝน 70% และ ชุดข้อมูลทดสอบ 30% การแยกชุดข้อมูลเป็น 2 ชุดจะทำให้วัดประสิทธิภาพของตัวแบบได้ถูกต้องมากขึ้น
2. การฝึกสอนตัวแบบประเภทตัวเข้ารหัสอัตโนมัติ
   จากการฝึกสอนตัวแบบด้วยชุดข้อมูลที่เป็นปกติ จะทำให้ตัวแบบรู้จำถึงความสัมพันธ์ของลักษณะข้อมูล ส่งผลให้ตัวแบบมีความสามารถในการถอดรหัสข้อมูลที่เป็นปกติกลับมาโดยที่มีค่าการสูญเสียการสร้างใหม่ที่ต่ำ
3. การจำแนกชุดข้อมูลโดยการใช้ค่าการสูญเสียการสร้างใหม่
   ในขั้นตอนการจำแนกจะต้องระบุค่าเกณฑ์ (Threshold) ขึ้นมา เพื่อที่จะนำไปเปรียบเทียบกับค่าการสูญเสียการสร้างใหม่ที่ได้จากการถอดรหัส ถ้าค่าการสูญเสียการสร้างใหม่มีค่าสูงกว่าค่าเกณฑ์ให้ระบุว่าข้อมูลนั้นเป็นประเภทผิดปกติ ซึ่งในการศึกษานี้ผู้วิจัยจะใช้วิธีการดูแผนภาพฮิสโทแกรมของค่าการสูญเสียเพื่อหาค่าเกณฑ์ที่เหมาะสม

3.3.2 หาโครงสร้างของตัวเข้ารหัสอัตโนมัติที่เหมาะสม

จากการศึกษาโครงสร้างของตัวเข้ารหัสอัตโนมัติ แนวคิดของตัวเข้ารหัสอัตโนมัติคือการบังคับให้ตัวแบบบีบอัดชุดข้อมูลนำเข้าเพื่อให้ได้ข้อมูลรหัสออกมา จากนั้นตัวแบบจะถอดรหัสจากข้อมูลรหัสให้ออกมาคล้ายข้อมูลนำเข้าให้มากที่สุด ซึ่งโครงสร้างโดยทั่วไปของของตัวเข้ารหัสอัตโนมัติ

จะประกอบไปด้วย 3 ส่วนคือ ตัวเข้ารหัส ชั้นรหัส และตัวถอดรหัส โดยที่จำนวนชั้นของตัวเข้ารหัสนั้นจะต้องเท่ากับจำนวนชั้นของตัวถอดรหัส และในส่วนของจำนวนนิวรอนของชั้นซ่อนไม่ได้ถูกระบุไว้ว่าควรจะมีค่าเท่าไหร่ แต่จะมีจำนวนน้อยกว่าจำนวนนิวรอนของชั้นนำเข้าและชั้นส่งออกซึ่งจะมีลักษณะของโครงสร้างคล้ายกับการถูกบีบเป็นนาฬิกาทรายขวางดังรูปที่ 2.4 ดังนั้นการทดลองนี้จึงได้ถูกแบ่งออกตามจำนวนชั้นซ่อนของตัวเข้ารหัสอัตโนมัติได้ 5 แบบดังตารางที่ 3.5 – 3.9

ตารางที่ 3.5 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 1 ชั้น

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
1	Input layer (encoder)	-	128	-
2	Hidden layer (code)	Fully Connected	64	ReLU
3	Output layer (decoder)	Fully Connected	128	ReLU

ตารางที่ 3.6 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 3 ชั้น

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
1	Input layer (encoder)	-	128	-
2	Hidden layer (encoder)	Fully Connected	64	ReLU
3	Hidden layer (code)	Fully Connected	32	ReLU
4	Hidden layer (decoder)	Fully Connected	64	ReLU
5	Output layer (decoder)	Fully Connected	128	Linear

ตารางที่ 3.7 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 5 ชั้น

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
1	Input layer (encoder)	-	128	-

ตารางที่ 3.7 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 5 ชั้น (ต่อ)

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
2	Hidden layer (encoder)	Fully Connected	64	ReLU
3	Hidden layer (encoder)	Fully Connected	32	ReLU
4	Hidden layer (code)	Fully Connected	16	ReLU
5	Hidden layer (decoder)	Fully Connected	32	ReLU
6	Hidden layer (decoder)	Fully Connected	64	ReLU
7	Output layer (decoder)	Fully Connected	128	Linear

ตารางที่ 3.8 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้น

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
1	Input layer (encoder)	-	128	-
2	Hidden layer (encoder)	Fully Connected	64	ReLU
3	Hidden layer (encoder)	Fully Connected	32	ReLU
4	Hidden layer (encoder)	Fully Connected	16	ReLU
5	Hidden layer (code)	Fully Connected	8	ReLU
6	Hidden layer (decoder)	Fully Connected	16	ReLU
7	Hidden layer (decoder)	Fully Connected	32	ReLU

ตารางที่ 3.8 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้น (ต่อ)

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
8	Hidden layer (decoder)	Fully Connected	64	ReLU
9	Output layer (decoder)	Fully Connected	128	Linear

ตารางที่ 3.9 แสดงโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 9 ชั้น

ลำดับ	ประเภทชั้น	ชนิด	จำนวนนิวรอน	ชนิดฟังก์ชันกระตุ้น
1	Input layer (encoder)	-	128	-
2	Hidden layer (encoder)	Fully Connected	64	ReLU
3	Hidden layer (encoder)	Fully Connected	32	ReLU
4	Hidden layer (encoder)	Fully Connected	16	ReLU
5	Hidden layer (encoder)	Fully Connected	8	ReLU
6	Hidden layer (code)	Fully Connected	4	ReLU
7	Hidden layer (decoder)	Fully Connected	8	ReLU
8	Hidden layer (decoder)	Fully Connected	16	ReLU
9	Hidden layer (decoder)	Fully Connected	32	ReLU
10	Hidden layer (decoder)	Fully Connected	64	ReLU
11	Output layer (decoder)	Fully Connected	128	Linear

โดยจะกำหนด Hyperparameters ที่ใช้สำหรับการเปรียบเทียบโครงสร้างของตัวแบบดังตารางที่ 3.10 ซึ่งค่าของ Hyperparameter แต่ละประเภทที่ถูกเลือกมานั้นมีประสิทธิภาพเพียงพอที่จะทำให้เปรียบเทียบประสิทธิภาพของตัวแบบได้

ตารางที่ 3.10 แสดงค่า Hyperparameters ที่ใช้ในการหาโครงสร้างของตัวแบบ

ชื่อ	ค่าที่กำหนด
Epoch	100
Batch size	256
Optimizer	Adam
Loss function	MAE
Learning rate	0.001

3.3.3 หาค่า Hyperparameters ที่เหมาะสม

Hyperparameters ที่ต้องทดลองหาค่าที่เหมาะสมและเกี่ยวข้องกับการฝึกสอนตัวแบบมีดังต่อนี้

1. Epoch เป็นจำนวนรอบของการฝึกสอน
2. Batch size เป็นจำนวนชุดข้อมูลที่จะให้ Optimizer คำนวณใน 1 ครั้ง
3. Loss function เป็นฟังก์ชันสำหรับคำนวณค่าการสูญเสีย
4. Learning rate เป็นค่าที่ใช้กำหนดว่าการปรับค่าน้ำหนักในแต่ละรอบการฝึกสอนจะมากหรือน้อยเท่าใด

3.4 การวัดประสิทธิภาพของตัวแบบ (Performance Evaluation)

ในการศึกษานี้ชุดข้อมูลสำหรับฝึกสอนได้ถูกแบ่งอัตราส่วนระหว่างข้อมูลปกติและข้อมูลผิดปกติเป็น 4 ประเภทเพื่อจำลองสถานการณ์ของชุดข้อมูลจริงที่อาจจะเป็นไปได้ดังนี้

1. ข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%
2. ข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%
3. ข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%
4. ข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%

ซึ่งการวัดประสิทธิภาพของตัวแบบจะใช้ค่าความแม่น ค่าความเที่ยง ค่าเรียกคืน และค่าคะแนนเอฟ ถ้ามีค่าสูงแสดงว่าตัวแบบมีประสิทธิภาพในการตรวจจับสูง โดยจะพิจารณาเปรียบเทียบระหว่างชุดข้อมูลเรียนรู้ทั้ง 4 แบบข้างต้น

บทที่ 4
ผลการวิจัยและการอภิปรายผล

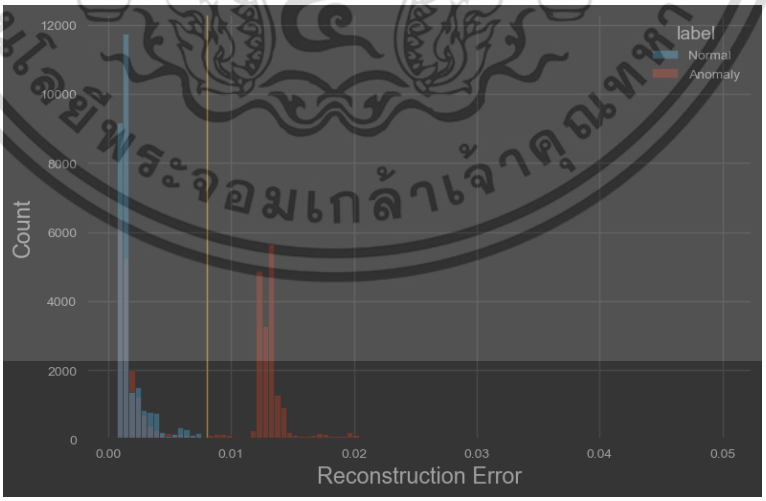
ในบทนี้จะกล่าวถึงการกำหนดโครงสร้างของตัวเข้ารหัสอัตโนมัติที่ใช้ในการทดสอบ ผลการศึกษาและทดสอบประสิทธิภาพของการนำเอาการเรียนรู้แบบไม่มีผู้สอนประเภทตัวเข้ารหัสอัตโนมัติมาฝึกสอนเพื่อใช้สร้างตัวแบบสำหรับตรวจจับการบุกรุกระบบเครือข่าย และปัญหาที่พบจากการศึกษา

4.1 การกำหนดโครงสร้างของตัวเข้ารหัสอัตโนมัติที่ใช้ในการทดสอบ

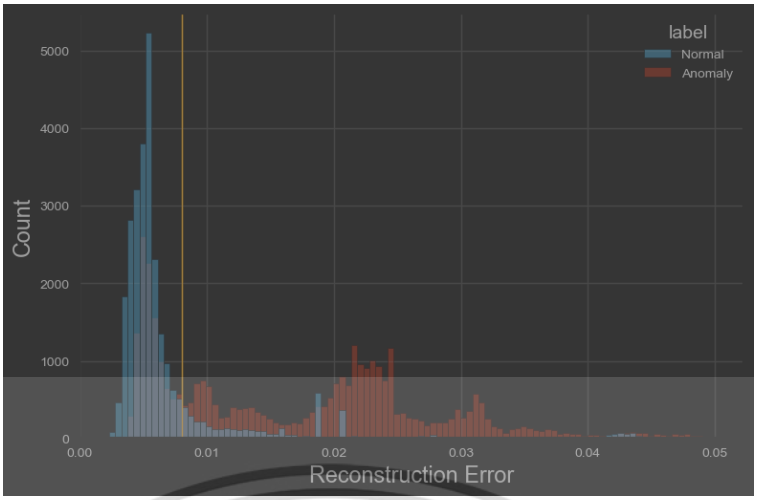
จากการทดลองปรับค่าและดูประสิทธิภาพของตัวแบบที่สร้างจากชุดข้อมูลปกติ 100% เพื่อที่จะเลือกโครงสร้างที่มีประสิทธิภาพกับการใช้จำแนกข้อมูล การทดลองถูกแบ่งเป็นสองส่วนคือ 1. การหาโครงสร้างของตัวเข้ารหัสอัตโนมัติ และ 2. การหาค่า Hyperparameters ที่เหมาะสม

4.1.1 การหาโครงสร้างของตัวเข้ารหัสอัตโนมัติ

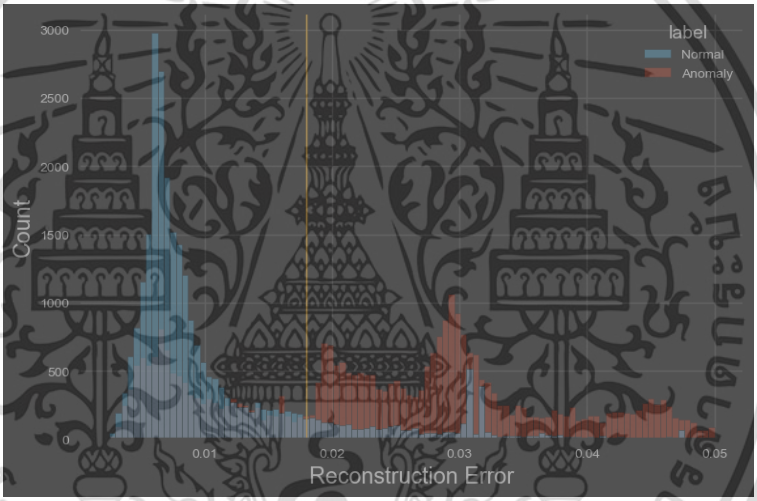
จากทดลองพบว่าตัวแบบที่มีโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้นมีความสามารถในการแยกแยะชุดข้อมูลที่ผิดปกติออกจากชุดข้อมูลที่ปกติได้ดีที่สุด โดยการเปรียบตามแผนภาพฮิสโทแกรมของค่าการสูญเสียระหว่างข้อมูลที่เป็นปกติ (สีฟ้า) และข้อมูลที่ผิดปกติ (สีแดง) ดังรูปที่ 4.1 – 4.5



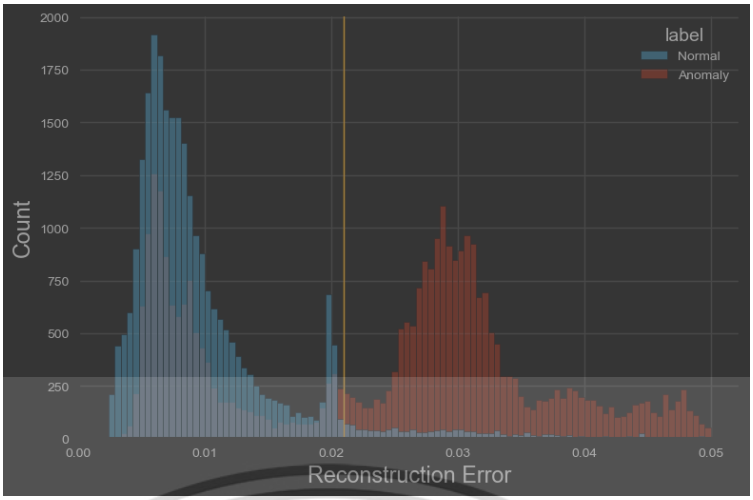
รูปที่ 4.1 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 1 ชั้น



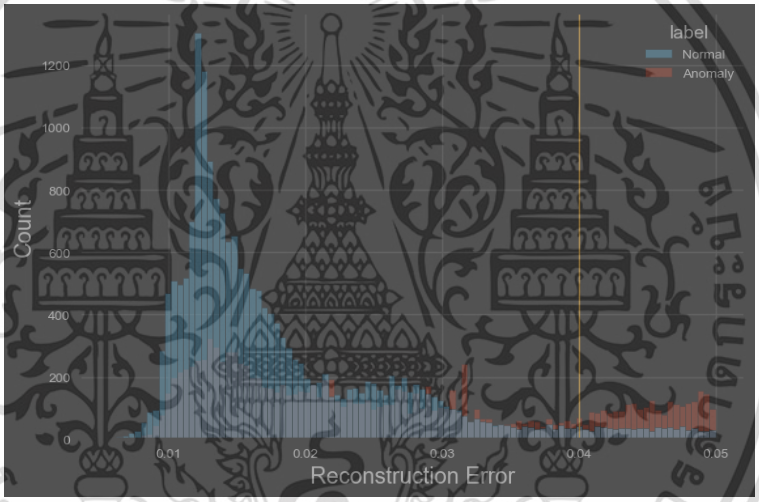
รูปที่ 4.2 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 3 ชั้น



รูปที่ 4.3 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 5 ชั้น



รูปที่ 4.4 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้น



รูปที่ 4.5 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 9 ชั้น

และเมื่อเปรียบนำค่าประสิทธิภาพของตัวแบบมาเปรียบเทียบดังตารางที่ 4.1 จะพบว่าตัวแบบที่มีโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้นมีประสิทธิภาพดีที่สุด ซึ่งสอดคล้องไปกับผลลัพธ์ที่แสดงจากฮิสโทแกรม ดังนั้นผู้วิจัยจึงเลือกโครงสร้างของตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้น ไปใช้หาค่า Hyperparameter ต่อไป

ตารางที่ 4.1 แสดงตารางเปรียบเทียบประสิทธิภาพของตัวแบบที่มีชั้นซ่อน 1-7 ชั้น

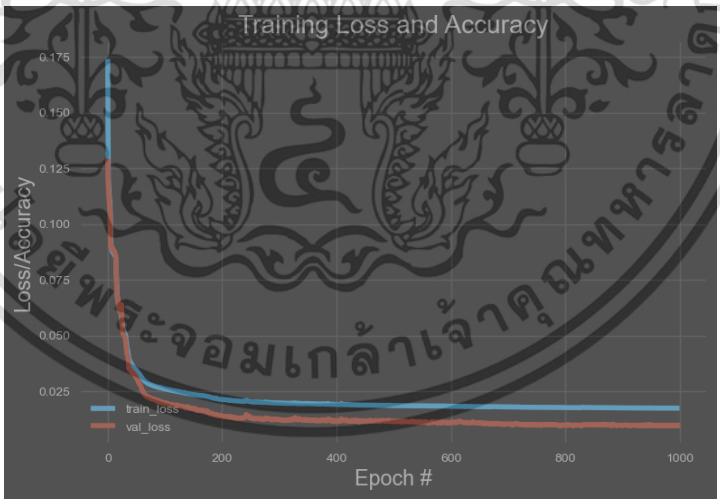
	จำนวนชั้นซ่อน				
	1	**3**	**5**	**7**	**9**
Accuracy	0.77	0.80	0.80	**0.81**	0.76
Precision	0.80	0.79	0.79	**0.81**	0.73
Recall	0.82	0.81	0.80	**0.83**	0.74
F1-Score	0.77	0.79	0.79	**0.80**	0.74

4.1.2 การหาค่า Hyperparameters ที่เหมาะสม

จากการทดลองโดยการกำหนดค่า Hyperparameters บนโครงสร้างตัวเข้ารหัสอัตโนมัติที่มีชั้นซ่อน 7 ชั้น ได้ผลได้ดังนี้

1. Epoch

   จากการทดลองพบว่า ตัวแบบสามารถฝึกสอนได้ถึง 1000 รอบโดยที่ค่าการสูญเสียได้น้อยลงตามจำนวนรอบที่ฝึกสอนดังรูปที่ 4.6 และยังไม่ก่อให้เกิดการท่องจำ (Overfitting) อีกทั้งยังใช้เวลาในการฝึกสอนไม่มากเกินไป ผู้วิจัยจึงเลือกค่า 1000 รอบมาใช้งาน

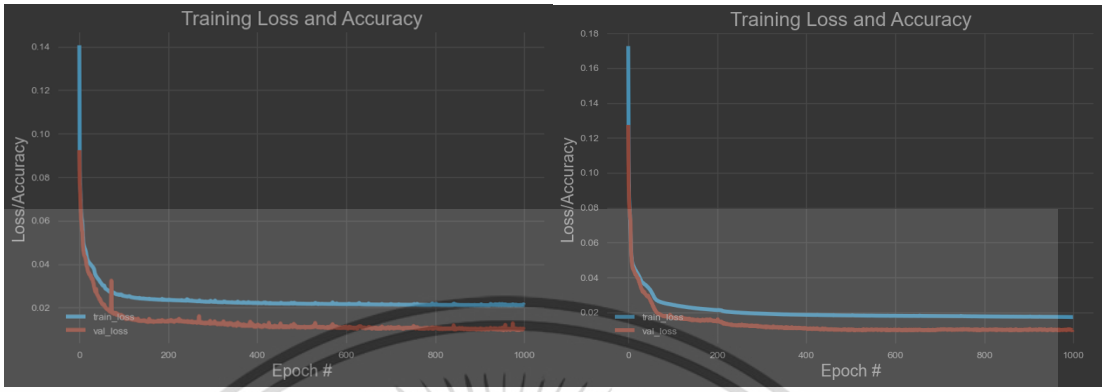


รูปที่ 4.6 แสดงค่าการสูญเสียของการฝึกสอน 0 – 1000 รอบ

2. Batch size

   จากการทดลองพบว่า ค่านี้ไม่ได้ส่งผลถึงประสิทธิภาพของการฝึกสอนอย่างมีนัยยะสำคัญ ดังรูปที่ 4.7 แต่จะช่วยให้การฝึกสอนชุดข้อมูลจำนวนมากเป็นไปได้ในเครื่องที่มี

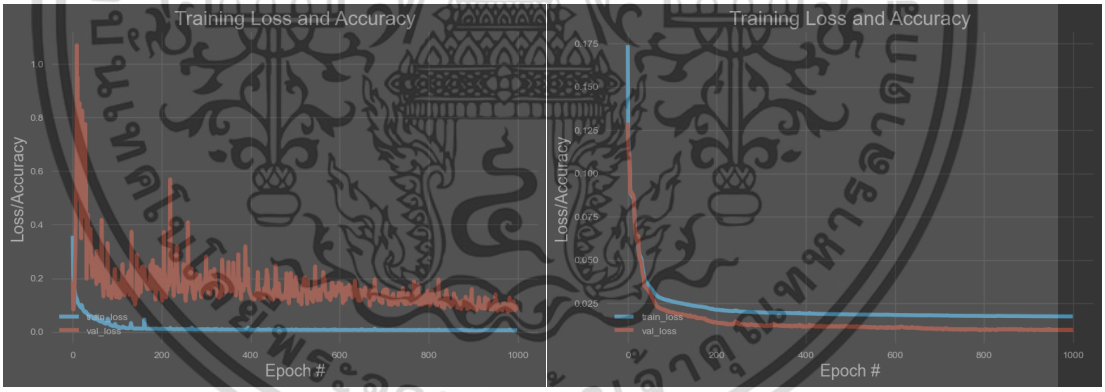
หน่วยความจำน้อย แลกกับระยะเวลาในการฝึกสอนที่มากขึ้นเนื่องจาก Optimizer ต้องคำนวณจำนวนรอบที่มากขึ้น ดังนั้นผู้วิจัยจึงเลือกค่า 256 มาใช้งาน



รูปที่ 4.7 แสดงค่าการสูญเสียของการฝึกสอนระหว่าง Batch size 128 (รูปซ้าย) และ 256 (รูปขวา)

3. Loss function

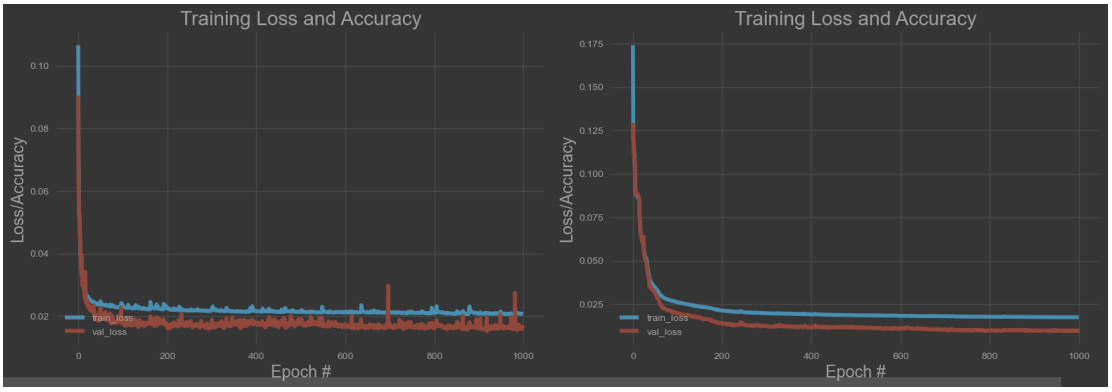
จากการทดลองพบว่า MSE มีความแปรปรวนมากเมื่อเปรียบเทียบกับ MAE ดังรูปที่ 4.8 ดังนั้นผู้วิจัยจึงเลือกค่า MAE มาใช้งาน



รูปที่ 4.8 แสดงค่าการสูญเสียของการฝึกสอนระหว่าง MSE (รูปซ้าย) และ MAE (รูปขวา)

4. Learning rate

จากการทดลองพบว่าเมื่อระบุค่าเป็น 0.01 ค่าการสูญเสียของตัวแบบมีความแปรปรวนมากเมื่อเปรียบเทียบกับค่า 0.001 ดังรูปที่ 4.9 ส่วนเมื่อระบุค่าเป็น 0.0001 ตัวแบบไม่สามารถฝึกสอนได้จนจบได้บนเครื่องที่ใช้ทดสอบ ดังนั้นผู้วิจัยจึงเลือกค่า 0.001 มาใช้งาน



รูปที่ 4.9 แสดงค่าการสูญเสียของการฝึกสอนระหว่าง Learning rate 0.01 (รูปซ้าย) และ 0.001 (รูปขวา)

จากข้อมูลข้างต้นจึงสามารถสรุปค่า Hyperparameters ที่จะนำมาใช้ในการทดลองได้ดังตารางที่ 4.2

ตารางที่ 4.2 แสดงค่า Hyperparameters ที่ใช้ในตัวแบบสำหรับการทดลอง

ชื่อ	ค่าที่กำหนด
Epoch	1,000
Batch size	256
Optimizer	Adam
Loss function	MAE
Learning rate	0.001

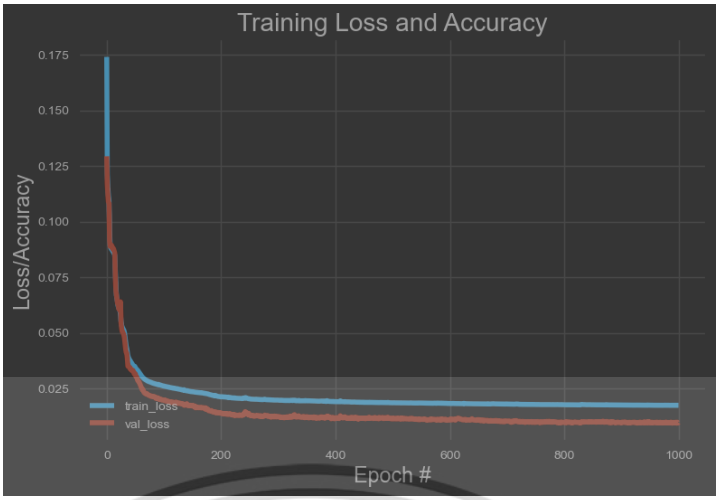
4.2 ผลการศึกษาและวัดประสิทธิภาพ

ผลการศึกษาและทดสอบประสิทธิภาพได้ถูกแบ่งตามอัตราส่วนของข้อมูลฝึกฝนดังนี้

1. ข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%
2. ข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%
3. ข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%
4. ข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%

4.2.1 อัตราส่วนข้อมูลฝึกฝน ข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%

จากการทดลองพบว่าตัวแบบสามารถเรียนรู้ชุดข้อมูลนี้ได้ดีดังรูปที่ 4.10 และสามารถแยกแยะข้อมูลที่ผิดปกติได้ดีจากค่าการสูญเสียดังรูปที่ 4.11



รูปที่ 4.10 แสดงค่าการสูญเสียของการฝึกสอนของชุดข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%

รูปที่ 4.11 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของชุดข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%

โดยที่เมื่อกำหนดค่าเกณฑ์ที่ 0.018 และนำผลลัพธ์มาสร้างเป็นเมทริกซ์ความสับสนและคำนวณประสิทธิภาพของตัวแบบดังตารางที่ 4.3 และ 4.4 จะพบว่าตัวแบบมีประสิทธิภาพในการตรวจจับข้อมูลผิดปกติที่ ความแม่น (Accuracy) 82% ความเที่ยง (Precision) 96% ค่าเรียกคืน (Recall) 76% และคะแนนเอฟ (F1-Score) 84%

ตารางที่ 4.3 แสดงเมทริกซ์ความสับสนของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%

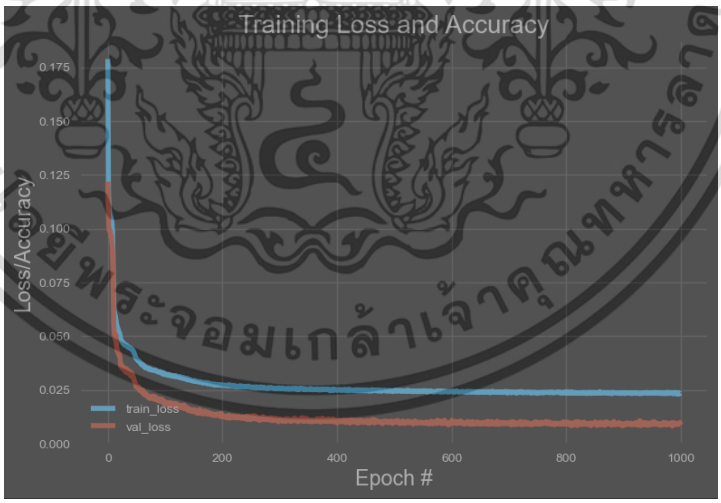
		ค่าทำนาย	
		ข้อมูลปกติ	**ข้อมูลบุกรุก**
ค่าจริง	**ข้อมูลปกติ**	26,187	1,754
	ข้อมูลบุกรุก	12,015	37,346

ตารางที่ 4.4 แสดงค่าประสิทธิภาพของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 100% ข้อมูลผิดปกติ 0%

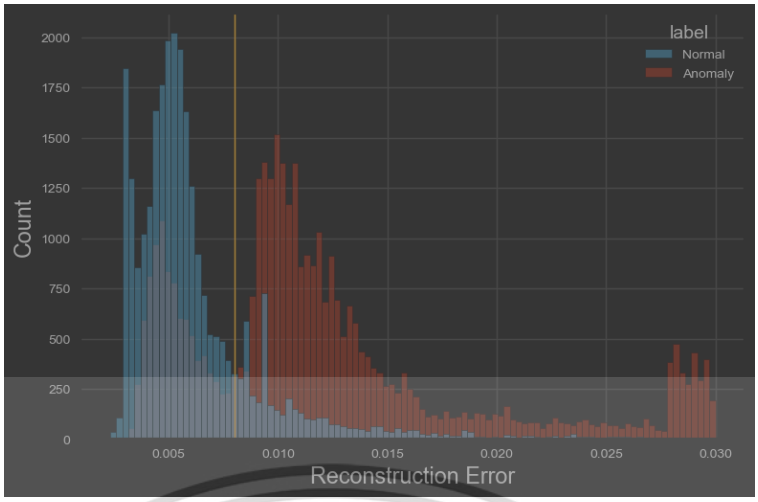
	Accuracy	Precision	Recall	F1-score	Support
ข้อมูลปกติ	0.82	0.69	0.94	0.79	27,941
ข้อมูลบุกรุก		0.96	0.76	0.84	49,361

4.2.2 อัตราส่วนข้อมูลฝึกฝน ข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%

เมื่อมีการเพิ่มข้อมูลที่ผิดปกติเข้าไป 1% พบว่าตัวแบบมีค่าการสูญเสียที่เพิ่มมากขึ้นในการเรียนรู้ดังรูปที่ 4.12 และเมื่อดูฮิสโทแกรมรูปที่ 4.13 จะพบว่าการกระจายตัวของค่าการสูญเสียของข้อมูลผิดปกติเริ่มเข้าใกล้ค่าการสูญเสียของข้อมูลปกติมากขึ้น



รูปที่ 4.12 แสดงค่าการสูญเสียของการฝึกสอนของชุดข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%



รูปที่ 4.13 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของชุดข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%

โดยที่เมื่อกำหนดค่าเกณฑ์ที่ 0.008 และนำผลลัพธ์มาสร้างเป็นเมทริกซ์ความสับสนและคำนวณประสิทธิภาพของตัวแบบดังตารางที่ 4.5 และ 4.6 จะพบว่าตัวแบบมีประสิทธิภาพในการตรวจจับข้อมูลผิดปกติที่ ความแม่น (Accuracy) 81% ความเที่ยง (Precision) 88% ค่าเรียกคืน (Recall) 81% และคะแนนเอฟ (F1-Score) 84%

ตารางที่ 4.5 แสดงเมทริกซ์ความสับสนของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%

		ค่าทำนาย	
		ข้อมูลปกติ	**ข้อมูลบุกรุก**
ค่าจริง	**ข้อมูลปกติ**	22,303	5,638
	ข้อมูลบุกรุก	9,159	40,202

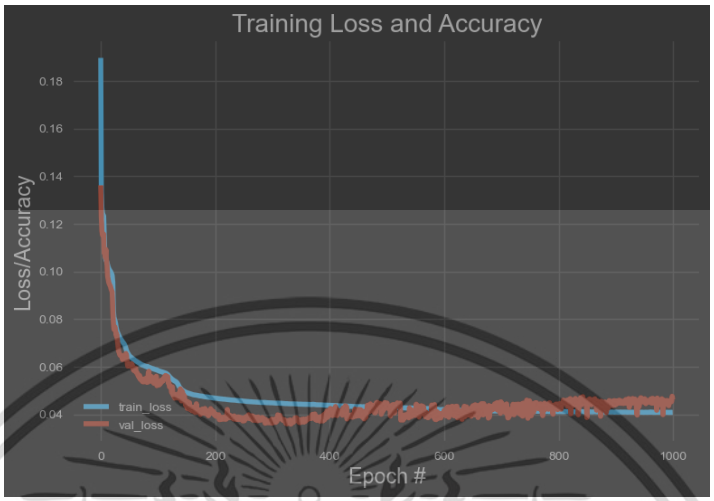
ตารางที่ 4.6 แสดงค่าประสิทธิภาพของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 99% ข้อมูลผิดปกติ 1%

	Accuracy	Precision	Recall	F1-score	Support
ข้อมูลปกติ	0.81	0.71	0.80	0.75	27,941
ข้อมูลบุกรุก		0.88	0.81	0.84	49,361

4.2.3 อัตราส่วนข้อมูลฝึกฝน ข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%

เมื่อมีการเพิ่มข้อมูลที่ผิดปกติเข้าไป 3% พบว่าตัวแบบมีค่าการสูญเสียที่เพิ่มมากขึ้นในการเรียนรู้ดังรูปที่ 4.14 และเมื่อดูฮิสโทแกรมรูปที่ 4.15 จะพบว่าการกระจายตัวของค่าการสูญเสีย

ของข้อมูลผิดปกติมีแนวโน้มเริ่มซ้อนทับค่าการสูญเสียของข้อมูลปกติ ทำให้ตั้งข้อสังเกตได้ว่า จำนวนของข้อมูลผิดปกติที่นำไปฝึกสอนตัวแบบน่าจะมีผลต่อประสิทธิภาพในการจำแนกของตัวแบบ



รูปที่ 4.14 แสดงค่าการสูญเสียของการฝึกสอนของชุดข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%

รูปที่ 4.15 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของชุดข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%

โดยที่เมื่อกำหนดค่าเกณฑ์ที่ 0.032 และนำผลลัพธ์มาสร้างเป็นเมทริกซ์ความสับสนและคำนวณประสิทธิภาพของตัวแบบดังตารางที่ 4.7 และ 4.8 จะพบว่าตัวแบบมีประสิทธิภาพในการตรวจจับข้อมูลผิดปกติที่ ความแม่น (Accuracy) 76% ความเที่ยง (Precision) 82% ค่าเรียกคืน (Recall) 80% และคะแนนเอฟ (F1-Score) 81%

ตารางที่ 4.7 แสดงเมทริกซ์ความสับสนของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%

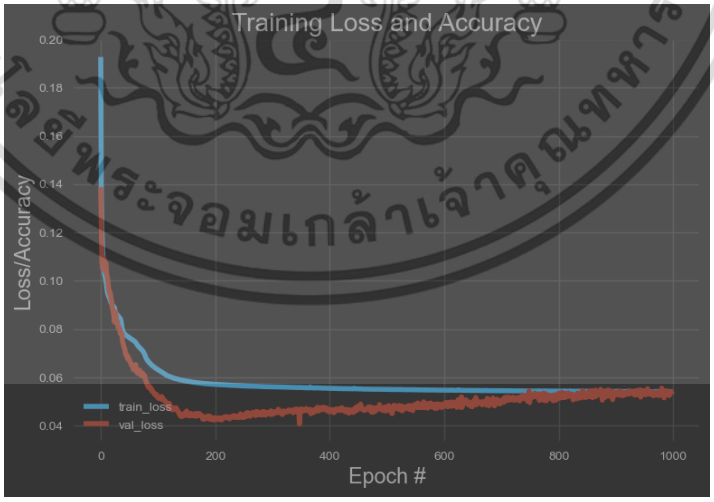
		ค่าทำนาย	
		ข้อมูลปกติ	**ข้อมูลบุกรุก**
ค่าจริง	**ข้อมูลปกติ**	19,239	8,702
	ข้อมูลบุกรุก	9,874	39,487

ตารางที่ 4.8 แสดงค่าประสิทธิภาพของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 97% ข้อมูลผิดปกติ 3%

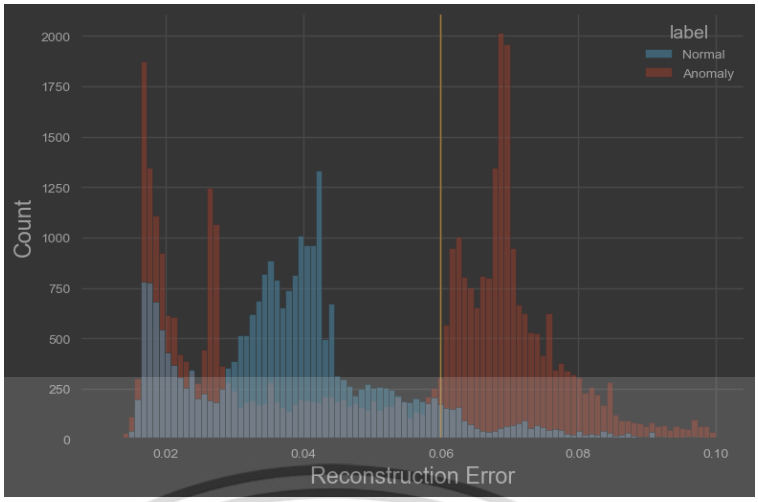
	Accuracy	Precision	Recall	F1-score	Support
ข้อมูลปกติ	0. 76	0.66	0.69	0.67	27,941
ข้อมูลบุกรุก		0.82	0.80	0.81	49,361

4.2.4 อัตราส่วนข้อมูลฝึกฝน ข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%

เมื่อมีการเพิ่มข้อมูลที่ผิดปกติเข้าไป 5% จะพบว่าตัวแบบมีค่าการสูญเสียที่เพิ่มมากขึ้น และมีแนวโน้มสูงขึ้นต่อเนื่องจากการฝึกสอนหลังจากรอบที่ 200 ดังรูปที่ 4.16 และเมื่อดูฮิสโทแกรม รูปที่ 4.17 จะพบว่าข้อมูลผิดปกติจำนวนหนึ่งมีค่าการสูญเสียที่ต่ำ ส่งผลให้การจำแนกมีความผิดพลาดมากขึ้น จากข้อมูลข้างต้นจะสรุปได้ว่า จำนวนของข้อมูลผิดปกติที่นำไปฝึกสอนตัวแบบมีผลต่อประสิทธิภาพในการจำแนกของตัวแบบอย่างเห็นได้ชัด



รูปที่ 4.16 แสดงค่าการสูญเสียของการฝึกสอนของชุดข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%



รูปที่ 4.17 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของชุดข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%

โดยที่เมื่อกำหนดค่าเกณฑ์ที่ 0.06 และนำผลลัพธ์มาสร้างเป็นเมทริกซ์ความสับสนและคำนวณประสิทธิภาพของตัวแบบดังตารางที่ 4.9 และ 4.10 จะพบว่าตัวแบบมีประสิทธิภาพในการตรวจจับข้อมูลผิดปกติที่ ความแม่น (Accuracy) 70% ความเที่ยง (Precision) 87% ค่าเรียกคืน (Recall) 63% และคะแนนเอฟ (F1-Score) 73%

ตารางที่ 4.9 แสดงเมทริกซ์ความสับสนของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%

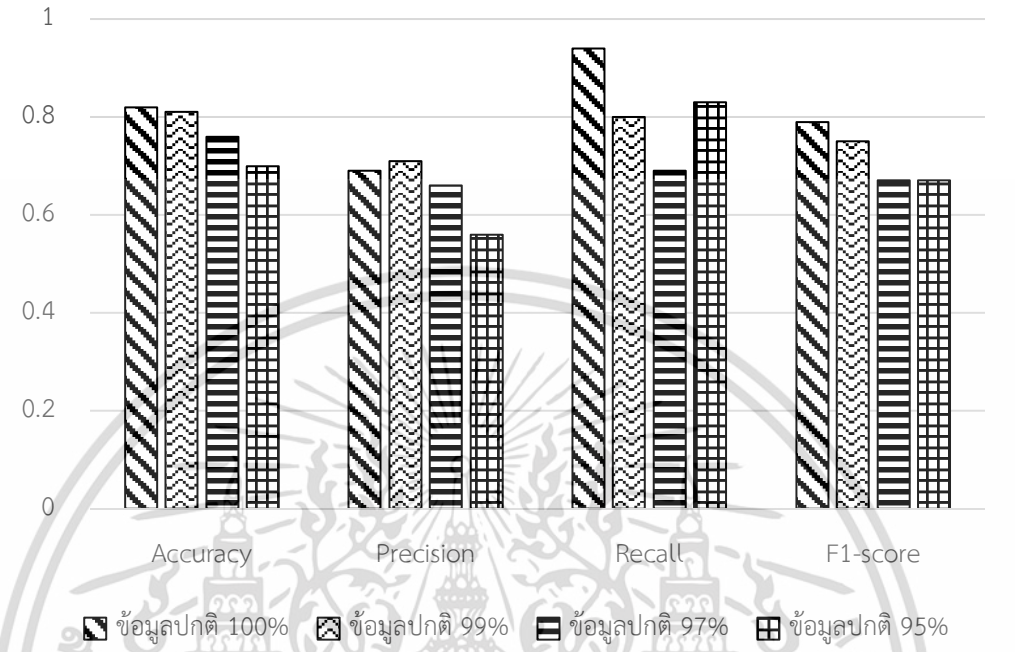
		ค่าทำนาย	
		ข้อมูลปกติ	**ข้อมูลบุกรุก**
ค่าจริง	**ข้อมูลปกติ**	23,120	4,821
	ข้อมูลบุกรุก	18,137	31,224

ตารางที่ 4.10 แสดงค่าประสิทธิภาพของตัวแบบที่ฝึกสอนจากชุดข้อมูลปกติ 95% ข้อมูลผิดปกติ 5%

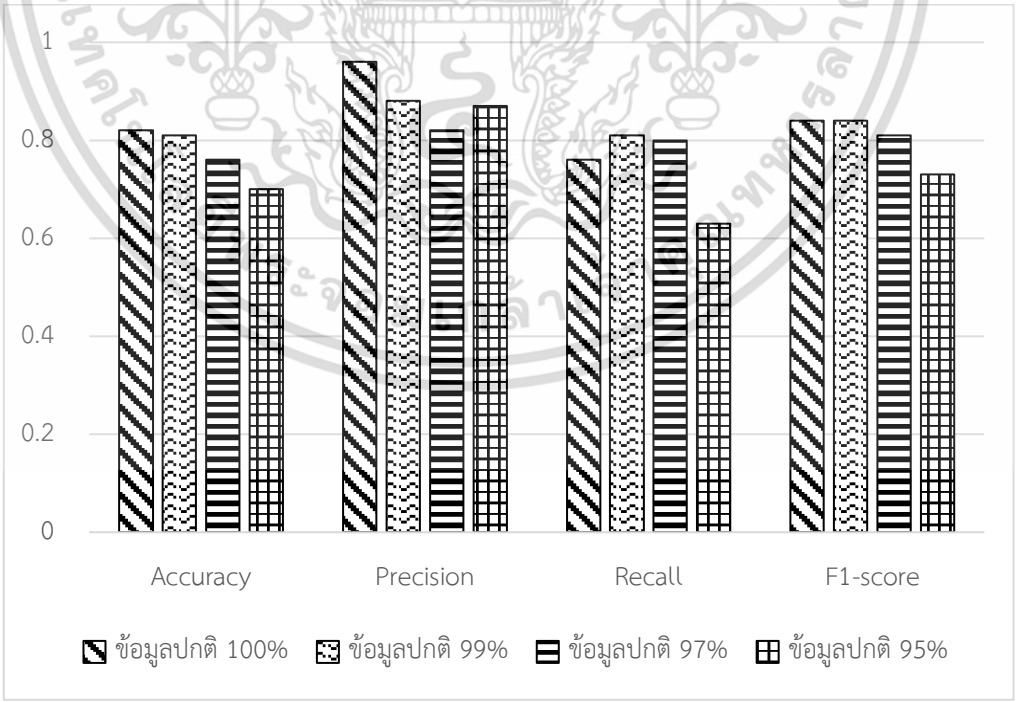
	Accuracy	Precision	Recall	F1-score	Support
ข้อมูลปกติ	0.70	0.56	0.83	0.67	27,941
ข้อมูลบุกรุก		0.87	0.63	0.73	49,361

เมื่อนำข้อมูลที่ได้จากการทดลองทั้ง 4 ชุดมาเขียนแผนภาพเพื่อเปรียบเทียบประสิทธิภาพดังรูปที่ 4.18 และ 4.19 จะเห็นได้ว่าประสิทธิภาพของการจำแนกทั้งสองกลุ่มนั้นเป็นไปในทิศทางเดียวกัน โดยที่ประสิทธิภาพที่ได้จะขึ้นอยู่กับจำนวนของอัตราส่วนข้อมูลปกติที่ใช้ในการฝึกสอนตัวแบบ ถ้า

อัตราส่วนของข้อมูลปกติมาก จะทำให้ตัวแบบมีประสิทธิภาพในการจำแนกสูง และประสิทธิภาพจะค่อย ๆ ลดลงเมื่อเพิ่มอัตราส่วนของข้อมูลผิดปกติให้มากขึ้น



รูปที่ 4.18 แสดงแผนภาพเปรียบเทียบค่าประสิทธิภาพของการจำแนก**ข้อมูลปกติ**จากการทดลอง



รูปที่ 4.19 แสดงแผนภาพเปรียบเทียบค่าประสิทธิภาพของการจำแนก**ข้อมูลผิดปกติ**จากการทดลอง

4.3 ปัญหาที่พบจากการศึกษา

4.3.1 ตัวแบบไม่สามารถตรวจจับการบุกรุกบางประเภทได้

จากผลการศึกษาของตัวแบบที่เรียนรู้จากข้อมูลปกติ 100% ข้อมูลผิดปกติ 0% เมื่อนำผลการจำแนกจากชุดข้อมูลทดสอบมาวิเคราะห์ดูประสิทธิภาพ โดยการนำผลการจำแนกมาจัดกลุ่มตามประเภทการบุกรุกดังตารางที่ 4.11 โดยที่การบุกรุกแต่ละประเภทจะมีคำอธิบายดังตารางที่ 4.12

ตารางที่ 4.11 แสดงผลการจำแนกโดยการจัดกลุ่มตามประเภทการบุกรุก

ประเภทการบุกรุก	ผลการจำแนก	จำนวนข้อมูล	อัตราส่วน
Analysis	ผิด	114	14%
	ถูก	675	86%
Backdoor	ผิด	60	8%
	ถูก	646	92%
DoS	ผิด	445	8%
	ถูก	4,489	92%
Exploits	ผิด	1,857	11%
	ถูก	11,643	89%
Fuzzers	ผิด	4,599	59%
	ถูก	2,609	41%
Generic	ผิด	117	1%
	ถูก	17,351	99%
Reconnaissance	ผิด	2,744	62%
	ถูก	1,493	38%
Shellcode	ผิด	432	88%
	ถูก	40	12%
Worms	ผิด	1	2%
	ถูก	46	98%

ตารางที่ 4.12 คำอธิบายของการบุกรุกแต่ละประเภท

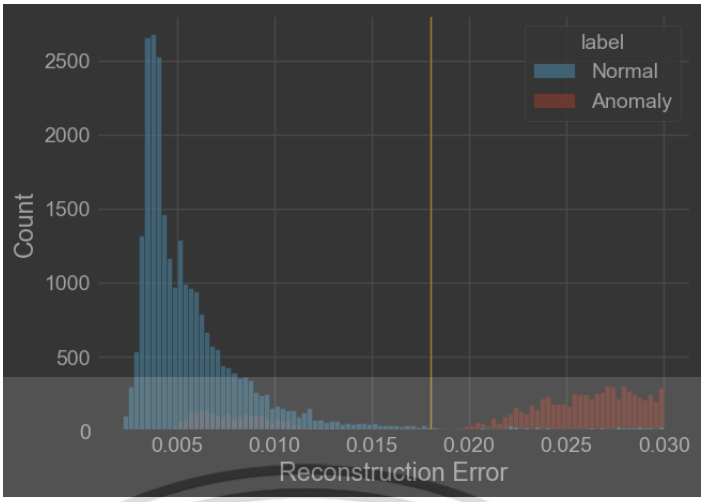
ประเภทการบุกรุก	คำอธิบาย
Analysis	เป็นการโจมตีโปรแกรมประยุกต์บนเว็บโดยมีพื้นฐานจากพอร์ต

ตารางที่ 4.12 คำอธิบายของการบุกรุกแต่ละประเภท (ต่อ)

ประเภทการบุกรุก	คำอธิบาย
Backdoor	เป็นวิธีการในการเข้าถึงระบบโดยที่ไม่ได้รับอนุญาต
DoS	เป็นการโจมตีที่จะทำให้ระบบปฏิเสธการเข้าถึงทรัพยากรต่าง ๆ ของระบบที่จำเป็นสำหรับการประมวลผล
Exploits	เป็นการโจมตีโดยใช้ประโยชน์ของช่องโหว่ของระบบที่เกิดจากความบกพร่องในระบบปฏิบัติการหรือซอฟต์แวร์
Fuzzers	เป็นการโจมตีที่ทำให้ระบบค้างโดยการป้อนข้อมูลสุ่มเป็นจำนวนมาก
Generic	เป็นเทคนิคการทำลายการเข้ารหัสโดยที่ไม่ต้องคำนึงถึงโครงสร้างของการเข้ารหัส
Reconnaissance	เป็นการโจมตีที่จะเข้าถึงข้อมูลเกี่ยวกับระบบซึ่งอาจจะทำให้รู้ถึงความบกพร่องของระบบได้
Shellcode	เป็นการโจมตีโดยใช้ชุดคำสั่งขนาดเล็กเพื่อทำให้ระบบมีช่องโหว่
Worms	เป็นการโจมตีระบบโดยใช้โปรแกรมที่มีความสามารถในการทำซ้ำตัวเองและแพร่กระจายไปยังระบบอื่น ๆ ผ่านทางช่องทางต่าง ๆ

พบว่าการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode ไม่สามารถถูกจำแนกได้อย่างมีประสิทธิภาพ เนื่องจากการโจมตีทั้งสามแบบนี้จะต้องมีการวิเคราะห์ถึงข้อมูลที่ถูกส่งมากับกลุ่มข้อมูล (Packet) ในระบบเครือข่ายด้วย ซึ่งชุดข้อมูลที่ใช้ในการศึกษานี้ไม่ได้มีคุณลักษณะดังกล่าว จึงเป็นข้อจำกัดที่ทำให้ไม่สามารถจำแนกการบุกรุกประเภทนี้ได้

ดังนั้นจึงได้มีการนำตัวแบบและชุดข้อมูลนี้มาทำการทดสอบอีกครั้ง โดยครั้งนี้การทดลองจะลบข้อมูลที่เป็นการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode ออกไป เพื่อดูประสิทธิภาพของตัวแบบในกรณีที่ไม่มีการบุกรุกประเภทดังกล่าว ซึ่งหลังจากทดลองพบว่า ตัวแบบมีความสามารถในการแยกแยะข้อมูลที่ผิดปกติได้ดีขึ้นดังรูปที่ 4.20



รูปที่ 4.20 แสดงแผนภาพฮิสโทแกรมของค่าการสูญเสียของชุดข้อมูลที่ไม่มีการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode

โดยที่เมื่อกำหนดค่าเกณฑ์ที่ 0.018 และนำผลลัพธ์มาสร้างเป็นเมทริกซ์ความสับสนและคำนวณประสิทธิภาพของตัวแบบดังตารางที่ 4.13 และ 4.14 จะพบว่าตัวแบบมีประสิทธิภาพในการตรวจจับข้อมูลผิดปกติที่ ความแม่น (Accuracy) 92% ความเที่ยง (Precision) 94% ค่าเรียกคืน (Recall) 92% และคะแนนเอฟ (F1-Score) 93%

ตารางที่ 4.13 แสดงเมทริกซ์ความสับสนของตัวแบบที่ฝึกสอนจากชุดข้อมูลที่ไม่มีการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode

		ค่าทำนาย	
		ข้อมูลปกติ	**ข้อมูลบุกรุก**
ค่าจริง	**ข้อมูลปกติ**	25,829	2,185
	ข้อมูลบุกรุก	2,828	34,537

ตารางที่ 4.14 แสดงค่าประสิทธิภาพของตัวแบบที่ฝึกสอนจากชุดข้อมูลที่ไม่มีการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode

	Accuracy	Precision	Recall	F1-score	Support
ข้อมูลปกติ	0.92	0.90	0.92	0.91	28,014
ข้อมูลบุกรุก		0.94	0.92	0.93	37,365

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของการตรวจจับการบุกรุกระบบเครือข่ายโดยการใช้วิธีการเรียนรู้แบบไม่มีผู้สอน โดยการสร้างตัวแบบสำหรับจำแนกข้อมูลจากโครงข่ายประสาทเทียมประเภทตัวเข้ารหัสอัตโนมัติ และใช้ชุดข้อมูล UNSW-NB15 สำหรับฝึกสอนตัวแบบ ซึ่งมีจำนวนข้อมูลทั้งหมด 257,673 ตัวอย่าง โดยที่ชุดข้อมูลจะถูกนำมาลดมิติโดยการใช้การวิเคราะห์องค์ประกอบหลักให้เหลือคุณลักษณะ 128 ตัว แล้วจึงนำมาฝึกสอนตัวแบบโครงข่ายประสาทเทียมประเภทตัวเข้ารหัสอัตโนมัติ

การศึกษานี้ได้ออกแบบการทดลองเป็น 4 ชุดตามจำนวนอัตราส่วนข้อมูลฝึกฝนระหว่างข้อมูลประเภทปกติและประเภทผิดปกติ โดยที่แต่ละการทดลองได้ผลลัพธ์ดังตารางที่ 5.1 และ 5.2

ตารางที่ 5.1 แสดงค่าประสิทธิภาพของการจำแนก**ข้อมูลปกติ**จากการทดลอง

อัตราส่วนข้อมูล	Accuracy	Precision	Recall	F1-score
ข้อมูลปกติ 100%	0.82	0.69	0.94	0.79
ข้อมูลปกติ 99%	0.81	0.71	0.80	0.75
ข้อมูลปกติ 97%	0.76	0.66	0.69	0.67
ข้อมูลปกติ 95%	0.70	0.56	0.83	0.67

ตารางที่ 5.2 แสดงค่าประสิทธิภาพของการจำแนก**ข้อมูลผิดปกติ**จากการทดลอง

อัตราส่วนข้อมูล	Accuracy	Precision	Recall	F1-score
ข้อมูลปกติ 100%	0.82	0.96	0.76	0.84
ข้อมูลปกติ 99%	0.81	0.88	0.81	0.84
ข้อมูลปกติ 97%	0.76	0.82	0.80	0.81
ข้อมูลปกติ 95%	0.70	0.87	0.63	0.73

ดังนั้นจากผลการทดลองที่ได้มาจะสรุปได้ว่า โครงข่ายประสาทเทียมประเภทตัวเข้ารหัสอัตโนมัติมีความสามารถในการจำแนกข้อมูลที่ผิดปกติได้ โดยที่เมื่อฝึกสอนตัวแบบด้วยชุดข้อมูลฝึกฝนที่มีจำนวนข้อมูลปกติมาก ตัวแบบที่ได้ก็มีประสิทธิภาพในการจำแนกสูง แต่เมื่อฝึกสอนด้วยชุดข้อมูลฝึกฝนที่มีจำนวนข้อมูลผิดปกติเพิ่มมากขึ้น ตัวแบบที่ได้จะมีประสิทธิภาพลดลง นอกจากนี้เมื่อ

วิเคราะห์ผลเพิ่มเติมจะพบว่าการบุกรุกประเภท Fuzzers Reconnaissance และ Shellcode ยังไม่สามารถถูกจำแนกได้อย่างมีประสิทธิภาพเพราะเป็นการบุกรุกที่มีอัตราการจำแนกถูกต้องต่ำกว่า 50% เนื่องจากการโจมตีทั้งสามแบบนี้จะต้องมีการวิเคราะห์ถึงข้อมูลที่ถูกส่งมากับกลุ่มข้อมูลในระบบเครือข่ายเพิ่มเติม

5.2 ข้อจำกัด

ในการศึกษาในครั้งนี้ผู้วิจัยได้พบข้อจำกัดในการศึกษาดังนี้

1. คุณลักษณะที่ได้มาจากชุดข้อมูล UNSW-NB15 ประกอบไปด้วยข้อมูลการเชื่อมต่อเบื้องต้น แต่ไม่ได้รวมถึงข้อมูลที่แท้จริงที่ส่งมาพร้อมกับกลุ่มข้อมูล (Data Packet) ของการเชื่อมต่อนั้น ๆ ทำให้วิธีการที่นำเสนอยังไม่สามารถตรวจจับการบุกรุกบางประเภทได้อย่างมีประสิทธิภาพ

2. เนื่องจากข้อจำกัดทางด้านเวลาในการศึกษาทำให้ยังต้องมีการกำหนดค่าเกณฑ์สำหรับการระบุผลการจำแนกด้วยตัวเอง

5.3 ข้อเสนอแนะและงานในอนาคต

1. การลดมิติชุดข้อมูลแบบไม่เป็นเส้นตรง (Nonlinear) อาจจะเพิ่มประสิทธิภาพของการจำแนกได้

2. นำวิธีการที่นำเสนอไปทดลองใช้กับชุดข้อมูลสำหรับทดสอบระบบตรวจจับการบุกรุกระบบเครือข่ายอื่น ๆ เพื่อดูประสิทธิภาพของตัวแบบเมื่อต้องทำงานกับชุดข้อมูลอื่น

3. ปรับแต่ง Hyperparameters ของตัวแบบแบบละเอียดอาจจะทำให้ประสิทธิภาพของการจำแนกดีขึ้น

4. ศึกษาหาขั้นตอนวิธีที่ทำให้ได้ค่าเกณฑ์สำหรับการระบุผลการจำแนกแบบอัตโนมัติ และไม่ต้องพึ่งพาฉลากคำตอบในการกำหนดค่าเกณฑ์

5. ควรนำผลที่ได้ไปเปรียบเทียบประสิทธิภาพกับวิธีการจำแนกในงานวิจัยอื่น ๆ เพื่อหาข้อบกพร่องและพัฒนาวิธีการเพิ่มเติม

เอกสารอ้างอิง

[1] Denning, D.E. 1987. "An Intrusion-Detection Model." *IEEE Transactions on Software Engineering.* 13(2) : 222 – 232.

[2] Moustafa, N. and Slay, J. 2015. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)." 1-6. in **Military Communications and Information Systems Conference (MilCIS).** Canberra : IEEE

[3] Xiao, Y. Xing, C. Zhang, T. and Zhao, Z. 2019. "An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks." *IEEE Access*. 7(1) : 42210 – 42219.

[4] Anderson, D. and McHeill, G. 1992. **"Artificial Neural Network Technology."** New York : Kaman Sciences Corporation

[5] Ciaburro, G. and Venkateswaran, B. 2017. **Neural Networks with R.** Birmingham : Packt Publishing

[6] Goodfellow, I. Bengio, Y. and Courville, A. 2016. **Deep Learning.** [Online]. Available : https://www.deeplearningbook.org/contents/autoencoders.html

[7] Sakurada, M. and Yairi, T. 2014. "Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction." 4-11. In Rahman, A. and Deng, J. and Li, J. **Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis.** New York : Association for Computing Machinery

[8] Choi, H. Kim, M. Lee, G. and Kim, W. 2019. "Unsupervised Learning Approach for Network Intrusion Detection System Using Autoencoders." *The Journal of Supercomputing*. 75(9) : 5597 – 5621.

[9] Zong, W. Chow, Y. and Susilo, W. 2019. "Dimensionality Reduction and Visualization of Network Intrusion Detection Data." 441-455 in Jang-Jaccard, J. and Guo, F. **Information Security and Privacy.** Christchurch : Springer Cham

[10] Narayana Rao, K. Venkata Rao, K. and P.V.G.D., P.R. 2021. "A Hybrid Intrusion Detection System Based on Sparse Autoencoder and Deep Neural Network." *Computer Communications*. 180(C) : 77 – 88.

[11] Syarif, I. Prugel-Bennett, A. Wills, G.. 2012. "Unsupervised Clustering Approach for Network Anomaly Detection." 135-145. In Benlamri, R. **Networked Digital Technologies, 2012, Volume 293.** Berlin, Heidelberg : Springer

ประวัติผู้เขียน

ชื่อ	นายวิทยา ทศพิทักษ์กุล
วัน เดือน ปีเกิด	18 กุมภาพันธ์ พ.ศ. 2534
ที่อยู่ปัจจุบัน	กรุงเทพมหานคร
ประวัติการศึกษา	(2556) วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ย 3.76 มหาวิทยาลัยเกษตรศาสตร์
ทุนการศึกษาที่ได้รับ	ไม่มี
ผลงานทางวิชาการ	ไม่มี