**บทที่ 4**

**ผลการดําเนินงานวิจัย**

บทนี้จะเป็นการทดลองโดยปรับเปลี่ยนค่าตัวแปรต่างๆแล้วดูผลตัวชี้วัด ตัวชี้วัดจะมีอยู่   
2 ค่า ซึ่งได้แก่ ค่าความถูกต้องในการทำนายผลลัพธ์ และระยะเวลาการฝึกโมเดลให้สามารถใช้งานได้ โดยจะทำการทดลองตามวิธีการแบ่งชุดฝึกสอนที่เรากำหนด

**4.1. การเตรียมทดสอบการทำงานของการพัฒนา DNN Firewall**

**4.1.1. การกำหนดกฎไฟร์วอลล์สำหรับงานวิจัย**

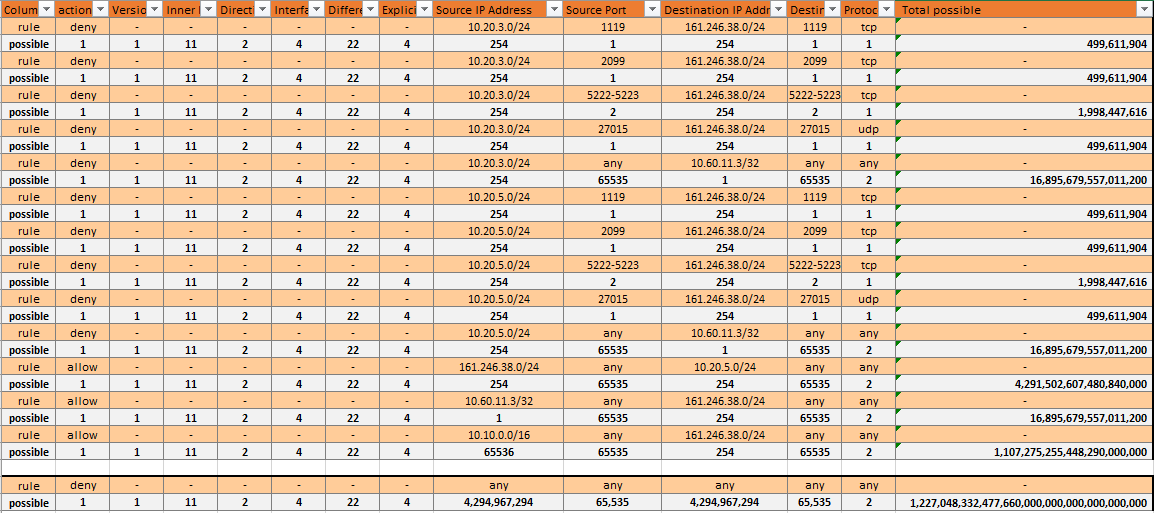
การออกแบบชุดกฎไฟร์วอลล์จะต้องแก้ไขความไม่ครอบคลุมของงานวิจัยเดิมที่มีขอบเขตความเป็นไปได้ของแพ็คเกตแต่ละกฎไฟร์วอลล์เท่ากันทั้งหมด เราต้องการให้การระบุจำนวนแพ็คเกตที่แต่ละกฎไฟร์วอลล์จะได้รับไม่ยึดติดกับกฎ สามารถใช้ได้กับกฎทุกรูปแบบและใช้งานได้อย่างมีประสิทธิภาพ ดังนั้นเราจึงเลือกใช้กฎที่มีการออกแบบมาและมีการใช้งานจริง กฎที่เรานำมาใช้มาจากกฎไฟร์วอลล์ IT Support ของคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเจ้าคุณทหารลาดกระบังมีทั้งหมด 14 กฎไฟร์วอลล์ และกฎ Default Rule อีก 1 กฎทำให้รวมเป็น 15 กฎ

Table

Description automatically generated  
การออกแบบกฎไฟร์วอลล์ที่ใช้สำหรับงานวิจัยนี้

**4.1.2. การคำนวณหาค่าจำนวนชุดฝึกสอนของแต่ละอัลกอริทึ่ม**

เพื่อให้ได้จำนวนชุดฝึกสอนที่แต่ละกฎไฟร์วอลล์จะได้รับในแต่ละอัลกอริทึ่ม จำเป็นต้องมีการคิดคำนวณหาค่า Sample Space ของแต่ละกฎ เพราะค่านี้เป็นตัวแปรที่บ่งบอกถึงขนาดขอบเขตของกฎส่งผลโดยตรงเมื่อมีการเข้าสูตรคำนวณการแบ่งจำนวนแพ็คเกต ซึ่งจากกฎที่เราได้ออกแบบ ได้ผลจากการคำนวณ Sample Space ดังนี้

  
 ตารางแจกแจงความเป็นไปได้ของแต่ละ Data Field ของแต่ละกฎไฟร์วอลล์

Table

Description automatically generated  
ตารางแจกแจงจำนวนความเป็นไปได้ในแต่ละกฎไฟร์วอลล์ที่ใช้ในงานวิจัย

จากบทก่อนหน้าที่ทำให้เราได้เห็นถึงปัญหาของวิธีการแบ่งชุดฝึกสอนแบบอัตราส่วน Ratio ที่ไม่เหมาะสมที่จะนำมาใช้งานจริง เนื่องจากขอบเขตที่มากเกินไปของ Default Rule ทำให้เราไม่ตัดสินใจที่จะใช้วิธีการแบ่งแบบ Ratio ในการเปรียบเทียบ จึงทำให้เหลือเพียง 2 อัลกอริทึ่ม คือแบบ N Sample และแบบ Take log Function ซึ่งการจะเปรียบเทียบทั้ง 2 วิธีนี้ได้คือการใช้จำนวน total packet ที่จำนวนเดียวกัน โดยเราเลือกวิธีการแบ่งแบบ Take log Function เป็นตัวตั้งเนื่องจากควบคุมตัวแปรได้ยากกว่าเพราะมีหน่วยเป็นเลขฐาน หลังจากได้คำนวณจนได้จำนวนที่ต้องการของวิธีนี้แล้วให้หาผลรวมเพื่อให้ได้ total packet ที่จะฝึกสอนในโมเดล 1 ตัว ซึ่ง total packet เมื่อนำไปหารจำนวนกฎที่ออกแบบแล้วจะได้จำนวนชุดฝึกสอนในแต่ละกฎไฟร์วอลล์ตามสูตรนั่นเอง

วิธีการแบ่งแบบ Take log Function กำหนดให้ x หรือค่าฐานในงานวิจัยใช้ค่า base 16, base 12, base 10, base 8, base 6, base 4, base 3, base 2 ทำให้มีทั้งหมด 8 ค่าด้วยกัน นำไปใช้กับวิธี N Sample อีก 8 ค่าเพื่อเปรียบเทียบ และกลไกการนำ Default Rule มาคิดอีก 2 แบบคือรวมกฎ Default และไม่รวมกฎ Default ทำให้มีชุดฝึกสอนที่ต้องสร้าง แบ่งออกเป็น 32 ชุดด้วยกันดังตาราง

Table

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated  
จำนวนชุดฝึกสอนที่สร้างในแต่ละกฎ (N Sample without default)

Table

Description automatically generated  
จำนวนชุดฝึกสอนที่สร้างในแต่ละกฎ (N Sample with default)

**4.1.3. สร้างชุดฝึกสอนจากโปรแกรม Packet Generator**

สร้างแพ็คเกตจากกฎไฟร์วอลล์และเปลี่ยนรูปแบบข้อมูลเป็นเลขสองเพื่อให้สามารถใช้ได้กับโมเดล โดยชุดข้อมูลฝึกสอนจะอยู่ในรูปของไฟล์นามสกุล CSV

Background pattern

Description automatically generated  
ผลลัพธ์ที่ได้จากการสร้างแพ็คเกตฝึกสอน

จากการลองสร้างชุดฝึกสอน พบว่าไม่สามารถที่จะคำนวณการแบ่งโดยใช้ค่า log base ฐาน 2 ได้ เนื่องจากใช้ทรัพยากรเครื่องใช้ทดลองเกินกว่าจะรับได้ จึงไม่คิดการแบ่งโดยใช้ base ฐานสองในขั้นตอนถัดๆไป

**4.1.4. นำชุดฝึกสอนไปสอนโมเดล DNN Model Engine**

ชุดข้อมูลฝึกสอน 1 ชุด จะสามารถสอนโมเดลได้ 1 โมเดล ซึ่งเราได้แบ่งจำนวนชุดฝึกสอนสามารถสร้างได้ 28 โมเดล โมเดลที่ผ่านการฝึกสอนแล้วจะถูกบันทึกเป็นโครงข่ายประสาทเทียมที่พร้อมประมวลผล ไฟล์จะอยู่ในรูปของไฟล์นามสกุล .h5

**การออกแบบโครงสร้าง Deep Neural Network Model ที่ใช้ในงานวิจัย**

* Hidden layer = 3 layers
* Input node = 125 nodes
* Node in each hidden layers = 40 nodes
* Output node = 2 nodes

**4.2. ผลการดำเนินการทดลองประมวลผลชุดฝึกสอน**

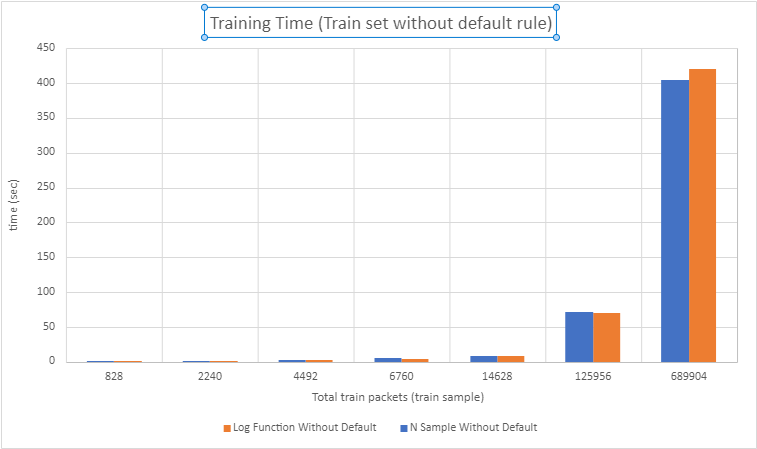
โมเดลที่ผ่านการฝึกฝนแล้วจะนำมาลองทดสอบ มี 28 โมเดล ประกอบไปด้วย

* โมเดลที่ฝึกสอนด้วยชุดฝึกสอนที่ไม่มี Default Rule เปรียบเทียบระหว่างอัลกอริทึ่ม N Sample และ Log Function โดยใช้ค่า total packet อย่างละ 7 ค่า ประกอบไปด้วย 828, 2,240, 4,492, 6,760, 14,628, 125,956, 689,904 แพ็คเกตตามลำดับ
* โมเดลที่ฝึกสอนด้วยชุดฝึกสอนที่มี Default Rule เปรียบเทียบระหว่างอัลกอริทึ่ม N Sample และ Log Function โดยใช้ค่า total packet อย่างละ 7 ค่า ประกอบไปด้วย 2,880, 6,294, 14,488, 24,192, 47,748, 519,168, 2,610,900 แพ็คเกตตามลำดับ

ชุดข้อมูลทดสอบ Test set เพื่อประมวลผลโมเดลแบ่งออกเป็น 3 ชุด ได้แก่

* ชุดข้อมูลทดสอบ InScope สร้างโดยใช้กฎไฟร์วอลล์ที่ออกแบบไม่รวม Default Rule มีทั้งหมด 14 กฎ แบ่งกฎละ 40,000 แพ็คเกต รวมเป็น 560,000 แพ็คเกต
* ชุดข้อมูลทดสอบ OutScope สร้างจาก Default Rule มี 2,000,000 แพ็คเกต
* ชุดข้อมูลทดสอบ MixScope ชุดทดสอบจาก InScope และ OutScope รวมกัน รวมเป็น 2,560,000 แพ็คเกต ข้อมูลแพ็คเกตเป็นคนละชุดกับ ชุดทดสอบก่อนหน้า

**ผลการดำเนินการ เวลาที่ใช้ฝึกสอนโมเดล**



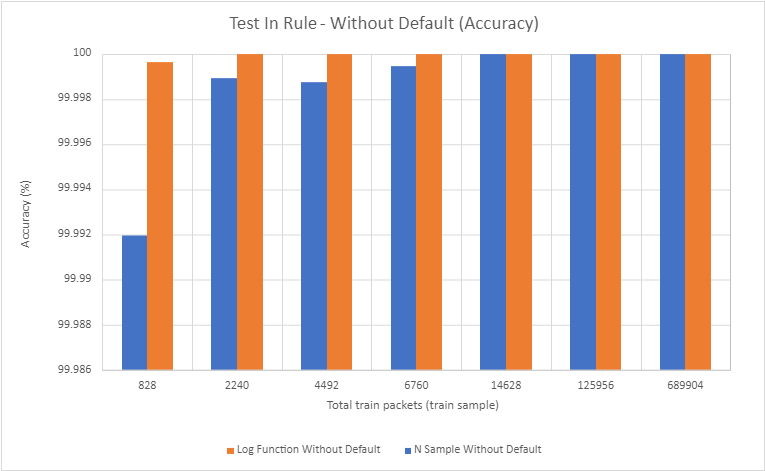
เวลาที่ใช้ในการฝึกโมเดลด้วยชุดฝึกสอน without default

Graphical user interface, application, table, Excel

Description automatically generated  
เวลาที่ใช้ในการฝึกโมเดลด้วยชุดฝึกสอน with default

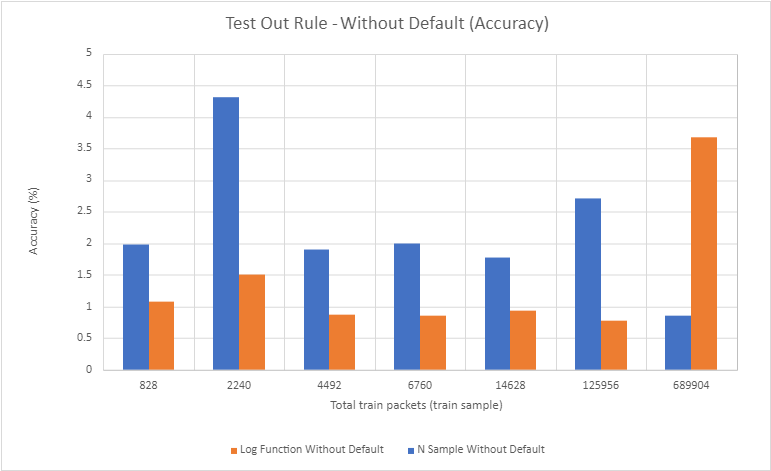
**ผลการทดลอง ความแม่นยำในการทำนายชุดทดสอบ**

การทดลองที่ 1 ประมวลผลชุดทดสอบ InScope กับโมเดลฝึกจากชุดฝึกสอนไม่รวม Default Rule



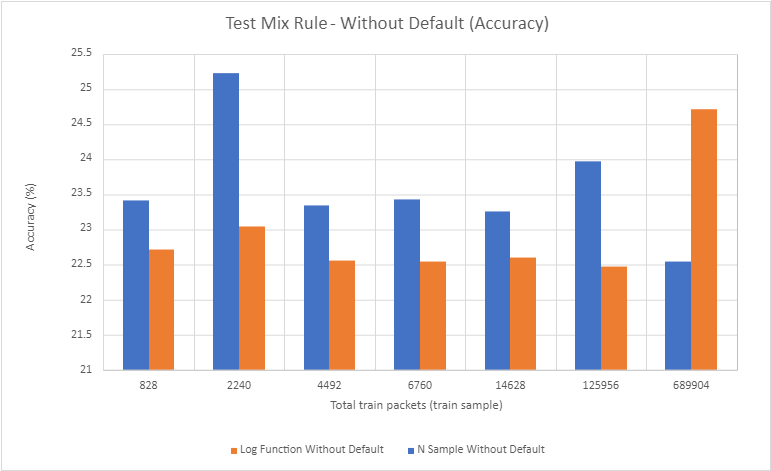
ผลลัพธ์โมเดลจากชุดฝึกสอน without default ประมวลผลชุดทดสอบ InScope

การทดลองที่ 2 ประมวลผลชุดทดสอบ OutScope กับโมเดลฝึกจากชุดฝึกสอนไม่รวม Default Rule



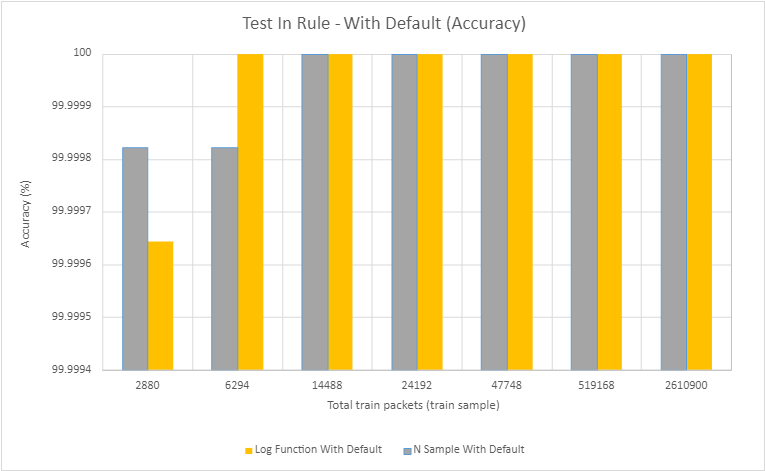
ผลลัพธ์โมเดลจากชุดฝึกสอน without default ประมวลผลชุดทดสอบ OutScope

การทดลองที่ 3 ประมวลผลชุดทดสอบ MixScope กับโมเดลฝึกจากชุดฝึกสอนไม่รวม Default Rule



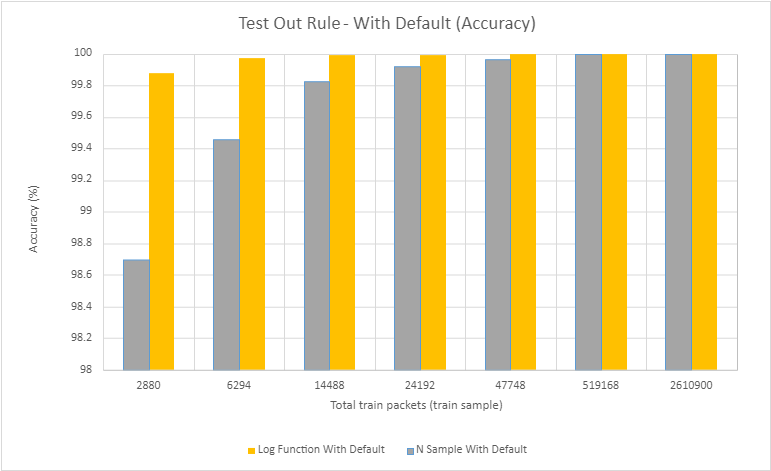
ผลลัพธ์โมเดลจากชุดฝึกสอน without default ประมวลผลชุดทดสอบ MixScope

การทดลองที่ 4 ประมวลผลชุดทดสอบ InScope กับโมเดลที่ฝึกจากชุดฝึกสอนรวม Default Rule



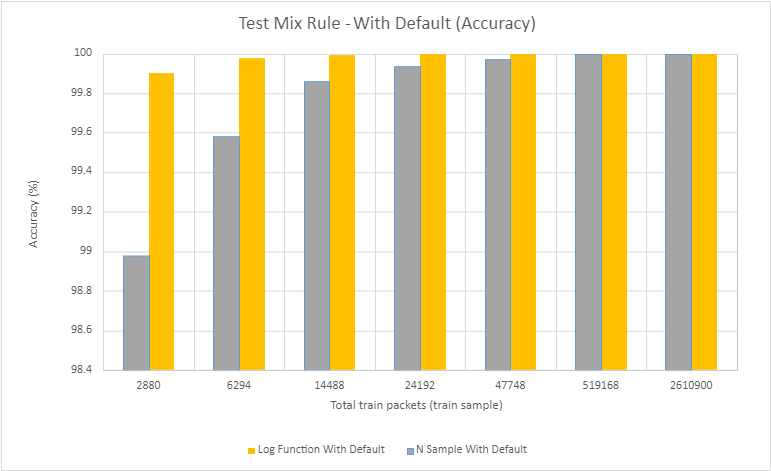
ผลลัพธ์โมเดลจากชุดฝึกสอน default ประมวลผลชุดทดสอบ InScope

การทดลองที่ 5 ประมวลผลชุดทดสอบ OutScope กับโมเดลที่ฝึกจากชุดฝึกสอนรวม Default Rule



ผลลัพธ์โมเดลจากชุดฝึกสอน default ประมวลผลชุดทดสอบ OutScope

การทดลองที่ 6 ประมวลผลชุดทดสอบ MixScope กับโมเดลที่ฝึกจากชุดฝึกสอนรวม Default Rule



ผลลัพธ์โมเดลจากชุดฝึกสอน default ประมวลผลชุดทดสอบ MixScope