**บทที่ 6**

**สรุปผลและข้อเสนอแนะ**

**6.1 สรุปผลการดำเนินงานวิจัย**

จากการทดลองในเชิงเปรียบเทียบโดยใช้ชุดข้อมูลฝึกสอนที่สร้างขึ้นเองโดยอิงตามสมมติฐานที่ตั้งไว้พบว่า จำนวนของชุดข้อมูลฝึกสอนที่มากขึ้นส่งผลให้โมเดลสามารถทำนายได้แม่นยำมากขึ้นเรื่อยๆ แค่ผลความแม่นยำจะเริ่มตกลงเมื่อมีข้อมูลฝึกสอนมากถึงจุดหนึ่ง อีกประเด็นหนึ่งคือความแม่นยำในการทำนายผลลัพธ์ซึ่งวิธีการแบ่งชุดข้อมูลฝึกสอนแบบ N Sample และแบบอัตราส่วน Ratio สามารถให้ผลลัพธ์ได้แม่นยำขึ้นทั้งคู่ แต่ถ้าหากอ้างอิงถึงเรื่องของประสิทธิภาพและอัตราการเรียนรู้ของโมเดลแล้ว แบบวิธีการแบ่งชุดข้อมูลฝึกสอนแบบ N Sample จะเห็นผลได้ดีกว่า ต่างกับการแบ่งชุดข้อมูลฝึกสอนแบบ Ratio ที่ความสามารถในการเรียนรู้ลดลงเรื่อยๆและทำนายผลลัพธ์ได้ผิดพลาดมากขึ้นถ้าหากมีข้อมูลมากเกินไป

ในงานวิจัยถัดไปจะเป็นการลงรายละเอียดเกี่ยวกับการพัฒนาแบ่งชุดข้อมูลฝึกสอนด้วยอัลกอริทึ่มแบบใหม่ ซึ่งเราได้คาดเดาว่าวิธีนี้จะเป็นการแก้ไขปัญหาวิธีการแบ่งชุดข้อมูลที่เป็นแบบอัตราส่วน โดยประเด็นปัญหาที่สามารถเห็นได้ชัดคือ การแบ่งข้อมูลฝึกสอนที่มีความแตกต่างกันทางด้านกฎของไฟร์วอลล์มากเกินไปจนทำให้ไม่สามารถทำนายชุดข้อมูลที่มีหลายเงื่อนไขได้ หรืออาจเพิ่มประเด็นวิจัยเพื่อเพิ่มความแม่นยำในการทำนายผล เช่น การปรับโมเดลหรือเปลี่ยนแปลงโครงสร้างของชุดข้อมูลฝึกสอน เป็นต้น

**6.2 ปัญหาและอุปสรรคที่พบในงานวิจัย**

* การพัฒนางานวิจัยใช้เวลานานมากกว่าที่คาดเอาไว้ เนื่องจากต้องพัฒนาโปรแกรมทั้งระบบควบคู่กับการทำทดลองไปด้วย ซึ่งการทดลองปัญญาประดิษฐ์ในเชิงเปรียบเทียบจำเป็นต้องทดลองซ้ำหลายรอบเพื่อให้ได้ผลลัพธ์ที่แม่นยำและวิเคราะห์ได้ ถ้าหากมีเวลาสำหรับการทดลองมากขึ้น อาจทำให้ได้ผลลัพธ์ที่แม่นยำและมีรายละเอียดที่น่าพึงพอใจมากขึ้น
* โปรแกรมในการสร้างชุดข้อมูลฝึกสอนมีข้อจำกัดหลายอย่าง เพราะเป็นเพียงการจำลองข้อมูลจาก Packet Header เพียงอย่างเดียว ยังไม่ได้ลงรายละเอียดในส่วนของ Data Field และยังจำเป็นต้องลดความเป็นไปได้ของ Possible Packet เนื่องจากมีปัญหาที่เครื่องคอมพิวเตอร์ที่ใช้ประมวลผลไม่สามารถรับภาระแบนด์วิดธ์ที่มากเกินไปได้

**6.3 ข้อเสนอแนะและแนวทางการพัฒนางานวิจัยในอนาคต**

* อาจมีวิธีการแก้ไขปัญหาการแบ่งอัตราส่วนชุดข้อมูลที่มีจำนวนต่างกันมากเกินไป อาจมีการใช้สูตรทางคณิตศาสตร์มาช่วยในการคำนวณหาจำนวนชุดข้อมูลที่เหมาะสมได้
* พัฒนาโปรแกรมสร้างชุดข้อมูลฝึกสอนให้มีประสิทธิภาพมากขึ้น ให้สามารถออกแบบได้ใกล้เคียงกับข้อมูล Packet ในเครือข่ายจริง และประมวลผลสร้างชุดข้อมูลได้รวดเร็วขึ้น
* พัฒนาเครื่องมือโมเดลโครงข่ายประสาทเทียมให้มีประสิทธิภาพมากขึ้น อาจลองศึกษาความสัมพันธ์ของตัวแปรที่ส่งผลต่อการเรียนรู้ของโมเดล ซึ่งประกอบไปด้วย จำนวนรอบที่เรียนรู้ จำนวนโหนดและวิธีการประมวลผลในรูปแบบต่างๆ และสังเกตว่าค่าเหล่านี้มีผลกับความแม่นยำและเวลาที่ใช้ในการฝึกสอนของชุดข้อมูลฝึกสอนที่สร้างไว้อย่างไร
* มีการเพิ่มสมมติฐานขึ้นใหม่ให้ใกล้เคียงกับเครือข่ายจริงมากขึ้น เช่น การเพิ่มกฎไฟร์วอลล์ที่มีความกระชับ หรือกำหนดให้มีข้อมูลที่จะพิจารณามากขึ้น เพิ่มจำนวน Rule set หรืออาจลองนำข้อมูลฝึกสอนจาก Application Layer มาใช้ควบคู่ด้วย

**บรรณานุกรม**

[1] TensorFlow Teams. “Essential Documentation**”** [Online]. เข้าถึงได้จาก :

https://www.tensorflow.org/guide. 2560

[2] สมาคมโปรแกรมเมอร์แห่งประเทศไทย. “Artificial Intelligent**”** [Online]. เข้าถึงได้จาก :

https://www.thaiprogrammer.org/2018/12/whatisai/

[3] Garry Fairhurst. “IPv4 Packet header Datagram**”** [Online]. เข้าถึงได้จาก :

https://networklessons.com/cisco/ccna-routing-switching-icnd1-100-105/ipv4-packet-header

[4] Sci-kit learn developers. “scikit classification model**”** [Online]. เข้าถึงได้จาก : <https://scikit->

learn.org/stable/search.html?q=classification

[5] พื้นฐาน Deep Learning**.** [Online]. เข้าถึงได้จาก : https://www.tensorflow.org/guide. 2560

[6] [Sinlapachai Lorpaiboon](https://medium.com/@sinlapachai.hon?source=post_page-----2f5049640e70--------------------------------). “การใช้ Pandas ในการจัดระเบียบข้อมูลใน Python**”** [Online]. เข้าถึงได้จาก :

<https://medium.com/@sinlapachai.hon/เรียนรู้วิธีการใช้งาน-Pandas-ใน-Python>

[7] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference,  
 and Prediction, Second edition, Springer, 2009

[8] Saishruthi Swaminathan, Logistic Regression Detailed Overview. 2018, From  
 <https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc>

[9] Anas Al-Masri, What Are Overfitting and Underfitting in Machine Learning. 2019, From  
 <https://towardsdatascience.com/what-are-overfitting-and-underfitting-in-machine-learning-a96b30864690>

[10] Will Koehrsen, Overfitting vs. Underfitting: A Complete Example. 2018, From  
 <https://towardsdatascience.com/overfitting-vs-underfitting-a-complete-example-d05dd7e19765>