บทที่ 2

ทฤษฎีการนำโครงข่ายระบบประสาทเชิงลึก มาใช้ในการทำงานของไฟร์วอลล์

2.1 เทคโนโลยีของไฟรั่วอลล์และโครงข่ายระบบประสาทเชิงลึก

2.1.1 Firewall

Firewall เป็นระบบควบคุมและรักษาความปลอดภัยของระบบเครือข่าย คัดกรองข้อมูลเข้าออกใน ช่องทางอินเตอร์เน็ต เปรียบเสมือนยามเฝ้าประตู คอยป้องกันการโจมตี สแปม ผู้บุกรุกต่างๆที่ไม่หวังคีต่อระบบ และยังสามารถใช้ควบคุมการใช้งานของโปรแกรมที่ต้องการ ในปัจจุบันมีการใช้งานได้ทั้งระบบ Hardware และ Software ขึ้นอยู่กับความเหมาะสม ผลลัพธ์ที่ออกมาจาก Firewall จะพิจารณาการกระทำของ Packet ออกมาเป็น Allow หรือ Deny

2.1.1.1 Packet Filtering

ระบบการทำงานของ Firewall ทำงานในระบบ Internet Layer และ Transport Layer ตรวจสอบ และคั่นกรอง Packet ที่เข้ามาในเครือข่าย โดยพิจารณาจาก Packet Header ตัดสินใจว่าจะทำการ Allow หรือ Deny โดยใช้กฎของ Firewall ในการอ้างอิง ซึ่ง Firewall แบ่งประเภทตามลักษณะการทำงาน ได้แก่

2.1.1.2 Stateful Filtering

Stateful Filtering จะมีเก็บสถานะ Packet ใดที่เคยถูกปล่อยผ่านและเก็บบันทึกไว้ใน State Table ทำให้การทำงานของ Firewall นี้จะถูกตรวจสอบเริ่มจากที่ State Table ก่อน ถ้าหาก Packet ที่ กำลังถูกตรวจสอบอยู่ยังไม่เคยถูกปล่อยผ่านยังไม่มีการเก็บสถานะเอาไว้ถึงจะไปพิจารณากฎของ ไฟรวอลล์เป็นอันดับถัดไป กลไกนี้จะช่วยไฟร์วอลล์ทำงานได้เร็วขึ้น เพราะช่วยลดระยะเวลาในการ ทำงานไม่ต้องเสียเวลาพิจารณาทุก Packet Header ในกลไก Packet Filtering

2.1.1.3 Application Firewall

มีชื่อเรียกได้อีกอย่างหนึ่งว่า "Application-level Firewall" หรือ "Application Gateway" เป็น Firewall ชนิดที่ติดตั้งบนเครื่องคอมพิวเตอร์แยกต่างหาก ทำให้คอมพิวเตอร์เครื่องดังกล่าวทำหน้าที่เป็น Firewall โดยเฉพาะ อย่างไรก็ตามApplication Firewall สามารถกรอง Packet ที่จะผ่านเข้ามาในเครือข่าย อีกทั้งยังตรวจสอบเนื้อหาใน Packet ได้เช่นเดียวกับ Stateful Filtering Firewall นอกจากนี้ Application Firewall ยังทำหน้าที่คล้ายกับ Proxy Server ในการให้บริการคำร้องขององผู้ใช้ได้อีกด้วย โดย ความสามารถของ Application Firewall สามารถแบ่งทำได้ดังนี้

Security

การยืนยันตัวตนด้วย AAA คือ Authentication, Authorization และ Audit โดยการสร้าง Token ไปให้ทั้งผู้รับ และผู้ส่ง มีการกำหนด Policy เพื่อการเข้าถึงข้อมูล และยังทำการเก็บ ข้อมูลการเข้าออกของ Policyนั้นๆ อีกทั้งยังมีการป้องกันด้วยการตรวจสอบข้อมูลที่ได้รับก่อน ว่าถูกต้องตามโครงสร้างที่ได้กำหนดไว้หรือไม่

Integration

การสร้างการเชื่อมต่อเข้ากับระบบต่างๆให้สามารถทำงานร่วมกันได้ เช่น ถ้าหาก ระบบที่ใช้มีโปรโตคอลที่แตกต่างกัน มันจะทำการแปลงโครงสร้างข้อความโดยการจับคู่ข้อมูล

Control and Managing

การควบคุมปริมาณของข้อความที่จะวิ่งเข้าไปหา Server โดยการกำหนด Policy แยก ตามประเภทของ API และประเภทของข้อมูล สำหรับการควบคุมปริมาณข้อความนี้จะเป็นการ ป้องกันการถูกผู้ไม่หวังดีโจมตีจากช่องโหว่ของระบบได้ เช่น เรามี API ที่เปิดให้ลูกค้าหรือ บุคคลอื่นๆเข้ามาใช้งานได้ ถ้าหากไม่มีการกำหนดปริมาณการเรียกใช้ API หรือเส้นทางของ ข้อมูล ก็จะเกิดช่องโหว่ของระบบที่ผู้ไม่หวังดีสามารถทำการ DOS ได้

Optimizing

การลดภาระการทำงาน ของ Server โดยการทำ SSL และนำภาระงานจากการถอดรหัส ที่ Server ไปให้ไฟร์วอลล์ทำงานแทน จะทำให้ Server มีทรัพยากรเหลือพอที่จะรองรับการ ทำงานมากขึ้น

2.1.2 Packet Header

Packet Header เป็นโปรโตคอลอินเตอร์เน็ต มาตรฐานที่ทำให้อินเตอร์เน็ตสามารถเชื่อมต่อเข้าหากัน ติคต่อสื่อสารข้อมูลได้ด้วยการกำหนดวิธีการติคต่อสื่อสารร่วมกัน ในส่วนของ Packet Header จะเป็นลำดับชั้น โครงสร้างประกอบไปด้วย Field ข้อมูลที่บ่งบอกถึงวัตถุประสงค์และลักษณะการทำงานของ Packet โดย องค์ประกอบของ Packet Header มีดังนี้

		32 bits					
Datagram Header ———		Version	Header length	Type of service	Datagram length (bytes)		
		16-bit Identifier			Flags	13-bit Fragmentation offset	
		Time-to-live		Upper-layer protocol	Header checksum		
			32-bit Source IP address				
		32-bit Destination IP address					
		Options (if any)					
		Data					

Version ส่วนที่ระบุเวอร์ชั่นโปรโตคอลของ Datagram

Header length ส่วนที่ระบุขนาดของ Datagram Header

Type of service ส่วนที่ระบุประเภทของ Datagram เช่น low delay high throughput หรือ Reliability

Datagram length ส่วนที่ระบุขนาดของ Datagram ทั้งหมดรวมถึง Datagram Header

Identifier ส่วนที่มีไว้เพื่อยืนยันตัว หากมีการทำ Fragmention

Flags ส่วนที่ระบุว่า Datagram นี้จะทำการ Fragmention หรือไม่

Fragmentation offset ส่วนที่แสดงให้เห็นถึงจำนวนของข้อมูลก่อนทำการ Fragmention

Time-to-live ส่วนที่กำหนดวงจรชีวิตของ Datagram เพื่อป้องกันไม่เกิด Loop ในเครื่อข่าย

Protocol ส่วนที่ระบุโปรโตคอลที่ใช้ใน Datagram นี้

Header checksum ส่วนที่ใช้สำหรับตรวจสอบความถูกต้อง Datagram Header

Source and destination IP addresses ส่วนที่ระบุที่อยู่ของ IP ต้นทางกับ IP ปลายทาง

Options ส่วนเพิ่มเติมที่คอยเก็บข้อมูลเช่น เส้นทางต้นทางของ Datagram เส้นทางที่ใช้โดยเก็บไว้เพื่อ ตรวจสอบการทำงาน เป็นต้น

2.1.3 Artificial Intelligent

Artificial Intelligence คือ เครื่องจักรอัจฉริยะที่มีความสามารถในการทำความเข้าใจ เรียนรู้องค์ความรู้ ต่างๆ เช่น การรับรู้ การให้เหตุผล ในการแก้ไขปัญหาต่างๆเพื่อปฏิบัติงานตามความต้องการของมนุษย์ เครื่องจักรที่มีความสามารถนี้ถูกเรียกอีกชื่อหนึ่งว่า "ปัญญาประดิษฐ์"

AI ถูกจำแนกเป็น 3 ระดับตามความสามารถดังนี้

Narrow Artificial Intelligent ปัญญาประคิษฐ์เชิงแคบ คือ AI ที่มีความสามารถเฉพาะทางได้ดีกว่า มนุษย์ เช่น เครื่องจักรที่ใช้ในการผ่าตัด

General Artificial Intelligent ปัญญาประดิษฐ์ทั่วไป คือ AI ที่มีความสามารถระดับเดียวกับมนุษย์ สามารถทำทุกอย่างในประประสิทธิภาพที่ใกล้เคียงกับมนุษย์

Strong Artificial Intelligent ปัญญาประดิษฐ์แบบเข้ม คือ AI ที่มีความสามารถมากกว่ามนุษย์ใน หลายๆด้าน

และจากการนำปัญญาประดิษฐ์มาประยุกต์ใช้ในการแก้ไขปัญหา มุมมองต่อ AI ที่แต่ละคนมีอาจไม่เหมือนกัน ขึ้นอยู่กับว่า เราต้องการความฉลาดโดย คำนึงถึงพฤติกรรมที่มีต่อสิ่งแวดล้อมหรือคำนึงการคิดได้ของ ผลผลิต AI ดังนั้นจึงมีคำนิยาม AI ตามความสามารถที่มนุษย์ต้องการ ให้มันแบ่งได้ 4 กลุ่ม ดังนี้

Thinking humanly (การคิดคล้ายมนุษย์)

natural language processing สื่อสารกับ มนุษย์ได้ด้วยภาษาที่มนุษย์ใช้ เช่น ภาษาอังกฤษ เป็น การประมวลผลภาษาธรรมชาติ

computer vision มีประสาทรับสัมผัสคล้ายมนุษย์ เช่นคอมพิวเตอร์วิทัศน์ รับภาพได้โดยใช้ อุปกรณ์รับสัญญาณภาพ machine learning เพื่อปรับให้เข้ากับสถานการณ์ใหม่และ ตรวจจับและคาคการณ์รูปแบบ

Thinking rationally (คิดอย่างมีเหตุผล)

คิดอย่างมี เหตุผล หรือคิดถูกต้อง โดยใช้หลักตรรกศาสตร์ในการคิดหาคำตอบอย่างมีเหตุผล เช่น ระบบผู้เชี่ยวชาญ

Acting humanly (การกระทำคล้ำยมนุษย์)

การคิดคล้าย มนุษย์ ก่อนที่จะทำให้เครื่องคิดอย่างมนุษย์ได้ ต้องรู้ก่อนว่ามนุษย์มีกระบวนการ คิดอย่างไร ซึ่งการวิเคราะห์ลักษณะการคิดของมนุษย์เป็นศาสตร์ด้าน cognitive science เช่น ศึกษา โครงสร้างสามมิติของเซลล์สมอง การแลกเปลี่ยนประจุไฟฟ้าระหว่างเซลล์สมอง วิเคราะห์การ เปลี่ยนแปลงทางเคมีไฟฟ้าในร่างกายระหว่างการคิด ซึ่งจนถึงปัจจุบันเราก็ยังไม่รู้แน่ชัดว่า มนุษย์เรา คิด ได้อย่างไร

Acting rationally (การกระทำอย่างมีเหตุผล)

กระทำอย่างมีเหตุผล เช่น agent (agent เป็นโปรแกรมที่มีความสามารถในการกระทำ หรือเป็น ตัวแทนในระบบอัตโนมัติต่าง ๆ) สามารถกระทำอย่างมีเหตุผลคือ agent ที่กระทำการเพื่อบรรลุ เป้าหมายที่ได้ตั้งไว้ เช่น agent ใน ระบบขับรถอัตโนมัติที่มีเป้าหมายว่าต้องไปถึงเป้าหมายในระยะทาง ที่สั้นที่ สุด ต้องเลือกเส้นทางที่ไปยังเป้าหมายที่สั้นที่สุดที่เป็นไปได้จึงจะเรียกได้ ว่า agent กระทำอย่าง มีเหตุผล อีกตัวอย่างเช่น agent ใน เกมหมากรุกมีเป้าหมายว่าต้องเอาชนะคู่ต่อสู้ ต้องเลือกเดินหมากที่จะ ทำให้คู่ต่อสู้แพ้ให้ได้ เป็นต้น

2.1.4 Machine Learning

Machine Learning คือ ส่วนการเรียนรู้ของเครื่อง ถูกใช้งานเสมือนเป็นสมองของปัญญาประดิษฐ์ในการ สร้างความฉลาด มักจะใช้เรียกโมเคลที่เกิดจากการเรียนรู้ของปัญญาประดิษฐ์ โดยมนุษย์มีหน้าที่เขียนโปรแกรม ให้เรียนรู้จากชุดข้อมูลฝึกสอนหรือ Training set และอาศัยกลไกที่เป็นโปรแกรม หรือเรียกว่า Algorithm ที่มี หลากหลายแบบ โดยมี Data Scientist เป็นผู้ออกแบบ หนึ่งใน Algorithm ที่ได้รับความนิยมสูง คือ Deep Learning ซึ่งถูกออกแบบมาให้ใช้งานได้ง่าย และประยุกต์ใช้ได้หลายลักษณะงาน อย่างไรก็ตาม ในการทำงาน จริง Data Scientist จำเป็นต้องออกแบบตัวแปรต่างๆ ทั้งในตัวของ Deep Learning เอง และต้องหา Algorithm

อื่นๆ มาเป็นคู่เปรียบเทียบ เพื่อมองหา Algorithm ที่เหมาะสมที่สุดในการใช้งานจริง โดยตามหลักแล้วจะแบ่ง ประเภทของ Machine Learning ได้ดังนี้

2.1.4.1 Supervised

การทำให้เครื่องคอมพิวเตอร์สามารถเรียนรู้ได้จากชุดข้อมูลฝึกสอนหรือ Training set ก่อนที่จะ ประมวลผล โดยมนุษย์จะเป็นผู้กำหนดคุณลักษณะ ความสัมพันธ์เฉพาะของข้อมูลที่ต้องการให้เครื่อง คอมพิวเตอร์เรียนรู้ หรือที่เรียกว่า Label และเมื่อโมเดลผ่านการเรียนรู้แล้ว จะสามารถแยกแยะประเภท มีวิธีการคิดที่เริ่มมีเหตุผล เมื่อข้อมูลที่ต้องการวิเคราะห์มีจำนวนที่มากขึ้นจำเป็นต้องมีข้อมูลที่เป็น Training set มากขึ้นเช่นเดียวกัน โดยการเรียนรู้แบบ Supervised Learning นี้จะประกอบไปด้วยดังนี้

2.1.4.1.1 Classification

คือการสอนโมเคลให้สามารถแบ่งหรือแยกประเภทกลุ่มข้อมูลได้ โดยอ้างอิงจาก ความสัมพันธ์และน้ำหนักของข้อมูลแต่ละ Label ตัวอย่างเช่น การแยกกลุ่มผู้ป่วยว่าเป็นเนื้อ งอกในสมอง ซึ่งจะมีปัจจัยต่างๆมากมายไม่ว่าจะเป็น ขนาด, รูปร่าง, ตำแหน่ง หรือ สีผิว ซึ่งถ้า หากมีข้อมูลเพียงแค่ Label เคียว ไม่สามารถพิสูจน์หรือแบ่งกลุ่มได้

2.1.4.1.2 Regression

การสอนโมเคลโดยอิงจากผลลัพธ์ที่ผ่านมา โดยผลลัพธ์จะเป็นการประมาณค่าความ เป็นไปได้ที่จะเกิดขึ้นต่อ ทำให้เหมาะแก่การวิเคราะห์ความสัมพันธ์ของตัวแปรที่อยู่ใน รูปกราฟ เช่น การหาความสัมพันธ์ระหว่างขนาดของบ้านและราคา การประเมินราคาหุ้น

2.1.4.2 Unsupervised

รูปแบบการเรียนรู้ที่ไม่จำเป็นต้องใช้ชุดข้อมูลฝึกสอน แต่เป็นการป้อนข้อมูล Test set ไป ประมวลผลเพียงอย่างเคียว ทำให้ผลลัพธ์ที่ออกมาไม่รู้ผลลัพธ์แน่ชัด ซึ่งอัลกอรึที่มจะวิเคราะห์และหา โครงสร้างของข้อมูลเอง

2.1.4.2.1 Clustering

เป็นการกำหนดให้เครื่องคอมพิวเคอร์หาวิธีแบ่งกลุ่มหรือจัดกลุ่มข้อมูลเอง เปรียบเสมือนการลด Label ของข้อมูลที่มีปริมาณมาก จัดกลุ่มหาข้อมูลที่มีความสัมพันธ์ ใกล้เคียงกัน ผลลัพธ์ที่ได้ออกมาจะมีปริมาณ Label ที่น้อยลงเป็นอย่างมาก

2.1.4.2.2 Dimensionality Reduction

เป็นการกลไกการบีบอัดและลดมิติข้อมูลจำนวนมากให้มีจำนวนลดลงโดยที่ข้อมูลยัง ครบถ้วน และยังสามารถนำไปจำแนกข้อมูลได้เหมือนเดิม

2.1.4.3 Reinforcement Learning

เป็นการเรียนรู้ด้านหนึ่งของ Machine Learning มักใช้พัฒนาหุ่นยนต์หรือการเรียนรู้ที่อยู่ภายใน เกมคอมพิวเตอร์ เช่น การลองผิดลองถูกไปเรื่อยเพื่อหาผลลัพธ์ที่ดีที่สุดประเมินออกมาเป็นคะแนน โดย ชุดข้อมูลทดสอบจะเป็นสภาพแวดล้อมโดยรอบขึ้นอยู่กับความต้องการของผู้พัฒนา

2.1.5 Deep Learning

Deep learning คือ อัลกอริทึมการเรียนรู้เชิงลึกโดยใช้หลักการ Artificial Neural Networks ที่มีรูปแบบ การทำงานคล้ายคลึงกับเซลล์ประสาทที่เชื่อมต่อกันเป็นโครงข่ายประสาทในสมองมนุษย์ เหมาะกับการ วิเคราะห์ข้อมูลขนาดใหญ่ที่มีความซับซ้อน เช่น การจำแนกรูปภาพ การจำแนกใบหน้า ประกอบไปด้วย โครงสร้างของหน่วยประมวลผลจำนวนมากคือเซลล์ประสาท หรือ Neuron โดยอัลกอริที่มนี้จะประกอบไปด้วย ชั้นต่างๆ ดังนี้

Input Layer มีหน้าที่รับข้อมูลเข้ามาประมวลผลและส่งต่อให้ Hidden Layer

Hidden Layer มีหน้าที่คำนวณและประมวลผลข้อมูลโดยสามารถมีได้หลายชั้น หลายขนาดขึ้นอยู่กับ ความซับซ้อนของข้อมูล

Output Layer มีหน้าที่ส่งผลลัพธ์ข้อมูลที่ผ่านการประมวลผลแล้วออกมา

เมื่อเริ่มการฝึกฝนจะเริ่มจากการสุ่มค่าถ่วงน้ำหนัก (Weight) และจะเริ่มปรับผลลัพธ์เอามาคูณกับค่าถ่วงน้ำหนัก แล้วบวกด้วยค่าความเอนเอียงของข้อมูล (Bias) หลังจากนั้นจะนำผลลัพธ์ที่ได้มาในแต่ละขาของ Neural Network มารวมกันแล้วมาผ่านพึงก์ชั่นส่งต่อให้ลำดับชั้นถัดไปประมวลผลมีการใช้วิธีการประมวลผลทาง คณิตศาสตร์ (Activation Function) โดยทุกวันนี้มีการะประยุกต์ใช้อย่างแพร่หลาย แบ่งชนิดโครงข่ายประสาท ออกเป็นดังนี้

2.1.5.1 โครงข่ายประสาทแบบป้อนไปข้างหน้า (Forward Propagation)

Feed-forward neural networks ถือเป็นโมเคลที่มีโครงสร้างที่เรียบง่ายที่สุด เพราะว่า การ คำเนินการของข้อมูลจะเป็นไปในทิศทางเคียว ก็คือ รับข้อมูลจาก input layer แล้วส่งไปต่อไปยัง hidden layer เลื่อยๆ จนกระทั้งถึง output layer ก็จะหยุด สังเกตุได้ว่าจะไม่มีวงวน หรือ loop เกิดขึ้นเลย

2.1.5.2 โครงข่ายแบบวนซ้ำ (Recurrent neural networks : RNN)

Recurrent neural networks คือ neural networks หลายเลเยอร์ที่สามารถเก็บข้อมูล information ไว้ที่ node จึงทำให้มันสามารถรับข้อมูลเป็นแบบลำคับ (data sequences) และ ให้ผลลัพธ์ออกเป็นลำคับ ของข้อมูล ได้ อธิบายอย่างง่ายๆ RNN ก็คือ neural network เชื่อมต่อกันหลายๆอันและยังสามารถต่อกัน เป็นวงวนหรือ loop ได้นั่นเอง เพราะฉะนั้น RNN จึงเหมาะสมในการประมวลผลข้อมูลที่เป็นลำคับ อย่างมาก

2.2 ทบทวนวรรณกรรม

2.2.1 การนำเอาความสามารถของ GPU มาใช้ในการคำนวณ

การที่เราเลือกใช้ GPU ในการทำ Machine Learning เนื่องจากตัว GPU นั้นมีหน่วยความจำที่ให้ค่า แบนด์วิดธ์ที่สูง และตัว GPU เองยังออกแบบให้สามารถแก้สมการทางคณิตสาสตร์ได้อย่างรวดเร็ว นอกจากนี้ยัง มีจำนวนหน่วยประมวลผลที่มีมากกว่า CPU หลายเท่าตัว จึงทำให้มีอัตราการประมวลผลที่สูงกว่า CPU และยังมี แพลตฟอร์มของ Nvidia ที่รองรับอย่าง CUDA ซึ่งเป็น Parallel Computing แพลตฟอร์มเพื่อช่วยให้นักพัฒนา สร้าง Tools ในการเรียกใช้การประมวลผลของ GPU และยังมี library อย่าง NVDIA cuDNN ซึ่งรองรับการทำ Deep Neural Network โดยตัว cuDNN ได้อำนวยการปรับแต่งขั้นสูงสำหรับการทำงานของ DNN เช่น forward และ backward convolution pooling normalization activation layers เป็นต้น

2.2.2 ทฤษฎี Rule of Thumb ในการหาจำนวนของ Hidden Layer

การตัดสินใจเลือกจำนวน Neurons ใน Hidden Layers นั้นถือเป็นส่วนสำคัญในการตัดสินภาพรวมของ สถาปัตยกรรมโครงข่ายประสาทเทียม โดย Hidden Layers นั้นจะ ไม่ค่อยมีผลกับองค์ประกอบภายนอกแต่จะมี ผลอย่างมากกับผลลัทธ์ที่จะออกมา จึงทำให้การกำหนดจำนวน Hidden Layers และ จำนวน Neurons ใน Hidden Layers นั้นต้องพิจาราณาอย่างระมัดระวัง เพราะถ้าเราใช้จำนวน Neurons น้อยเกินไปผลลัพท์ก็จะเกิดปัญหาที่ เรียกว่า Underfitting โดยจะเกิดขึ้นเมื่อมีจำนวน Neurons ใน Hidden Layers น้อยเกินไปผนไม่สามารถตรวจจับ สัญญาณในข้อมูลที่ซับซ้อนได้อย่างเพียงพอ แต่ในทางกลับกันหากเราใช้จำนวน Neurons มากเกินก็จะเกิด ปัญหาหลายอย่างตามมาโดยอย่างแรกก็คือ Ovefitting โดยจะเกิดขึ้นเมื่อความจุของข้อมูลที่จะประมวลผลมีมาก เกินไป ซึ่งจะไปจำกัดข้อมูลที่จะอยู่ในชุดฝึกสอนทำให้ไม่เพียงพอต่อการเรียนรู้ของ Neurons ใน Hidden Layer ปัญหาที่สองนั้นก็สามารถเกิดขึ้นมาได้แม้จะมีการเรียนรู้ของข้อมูลเพียงพอแล้วก็ตาม เนื่องด้วยจำนวน Neurons ที่มากเกินไปนั้นจะทำใช้เวลาในการเรียนรู้เพิ่มขึ้น ซึ่งเวลาในการเรียนรู้ที่เพิ่มขึ้นนั้นสามารถเพิ่มไปถึงจุดที่ทำ ให้การเรียนรู้ไม่สามารถทำใด้อย่างเพียงพอ ดังนั้นทำให้ต้องการกำหนดจำนวน Neurons ที่ไม่น้อยเกินไปหรือ มากเกินไป โดยมีหลักการอย่างง่ายในการกำหนดจำนวน Neurons ตามนี้

จำนวน Neurons ควรอยู่ในช่วงขนาดของ Input Layer และ Output Layer

จำนวน Neurons ควรมีขนาดเป็น 2 : 3 ของขนาด Input layer รวมกับ Output layer

จำนวน Neurons ควรมีขนาดน้อยกว่า 2 เท่าของขนาด Input Layer

โดยกฎทั้งสามที่ยกมานั้นเป็นเพียงส่วนหนึ่งในตัวเลือกให้สามารถนำไปใช้เพื่อให้ไม่ต้องมาสุ่มจำนวน Neurons ใหม่ซึ่งเท่าทำให้ไม่เสียเวลาที่ต้องนำไปทดลองกับจำนวน Neurons ที่สุ่มขึ้นใหม่