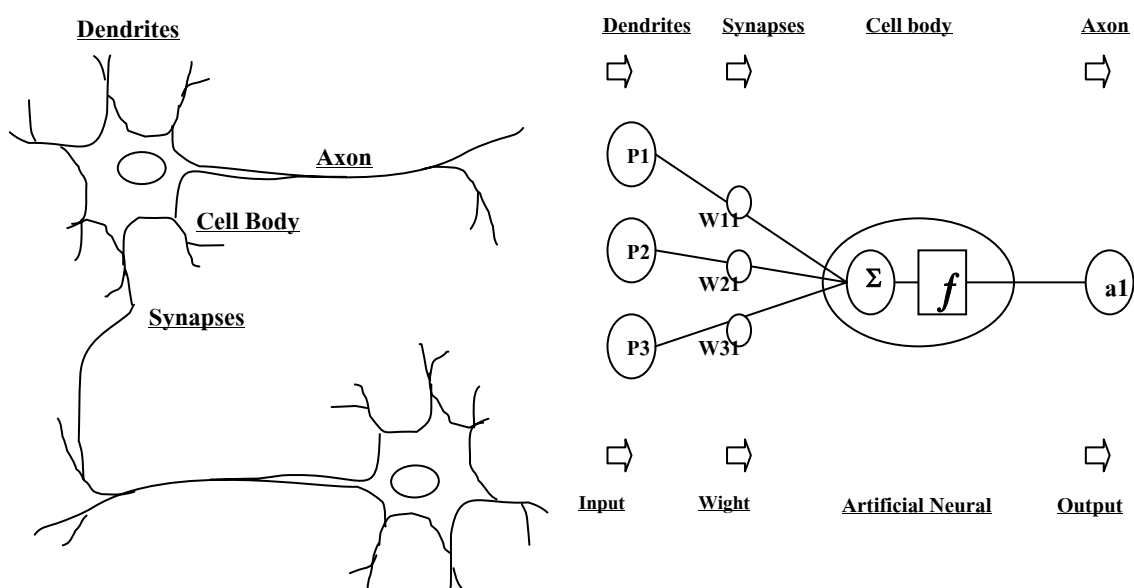


หลักการของโครงข่ายประสาทเทียม (Artificial Neural Network, ANNs)

1. ความรู้เบื้องต้นของโครงข่ายประสาทเทียม

1.1 พื้นฐานความรู้เบื้องต้นของโครงข่ายประสาทเทียม เป็นแนวคิดที่ถูกออกแบบให้ทำงานเช่นเดียวกับสมองของมนุษย์ สมองประกอบด้วยเซลล์เชื่อมต่อจำนวนมหาศาล(ประมาณ 10^{11} เซลล์) ที่สามารถเชื่อมต่อได้ 10^4 จุดต่อเซลล์ เซลล์เหล่านี้ถูกเรียกว่า เซลล์ประสาท(Neuron) การแบ่งส่วนประกอบของการทำงานสามารถแบ่งส่วนประกอบของการทำงานออกได้เป็น สาม ส่วนหลักๆ คือ เดนไดรต์ (Dendrites) เซลล์บอดี้ (Body Cell) และ แอกซอน (Axon) เดนไดรต์เป็นโครงข่ายรับข้อมูลที่มีลักษณะเป็นกิ่งก้านเหมือนต้นไม้ที่เกิดจากใยประสาท (Nerve Fiber) ซึ่งใช้เป็นทางเดินของสัญญาณทางไฟฟ้าเข้าสู่ เซลล์บอดี้ ซึ่งเซลล์บอดี้จะทำการรวมและจำกัดสัญญาณเข้านี้ แอกซอนจะเป็นเส้นใยยาวเส้นเดียวต่อหนึ่งเซลล์ประสาทที่ใช้เป็นทางเดินสัญญาณจากเซลล์บอดี้ ออกไปยังเซลล์ประสาทเซลล์อื่นๆ ที่เชื่อมโยงด้วย จุดเชื่อมต่อระหว่างแอกซอนของเซลล์หนึ่ง และเดนไดรต์ของอีกเซลล์หนึ่งถูกเรียกว่า ไซแนปส์ ซึ่งไซแนปส์นี้จะเป็นการจัดเรียงของเซลล์ประสาทโดยที่ความแข็งแรงของแต่ละไซแนปส์นั้นจะถูกกำหนดโดยกระบวนการทางเคมีที่มีความซับซ้อนและจะเป็นการสร้างฟังก์ชันของโครงข่ายประสาทขึ้นมา ภาพที่ 24 คือภาพอย่างง่ายของโครงข่ายการเชื่อมต่อของเซลล์ประสาทชีวะ(Biological Neuron) ทั้งสอง



ภาพที่ 24 แสดงการเปรียบเทียบระหว่างเซลล์ประสาทในทางชีววิทยากับเซลล์ประสาทเทียม

มีโครงสร้างของเซลล์ประสาทบางส่วนถูกกำหนดขึ้นตั้งแต่แรกเกิด และโครงสร้างของเซลล์ส่วนที่เหลือถูกพัฒนาจากการเรียนรู้ในต่อๆ มาซึ่งจะทำให้เกิดเส้นประสาทการเชื่อมต่อระหว่างเซลล์อันใหม่ และทั้งการเชื่อมต่อบางส่วนไป ซึ่งการพัฒนาสามารถจะเห็นได้ชัดที่สุดเมื่อในช่วงสภาวะแรกเกิด

โครงสร้างประสาทจะมีการเปลี่ยนแปลงตลอดชีวิตอย่างต่อเนื่อง ซึ่งการเปลี่ยนแปลงนี้ในเวลาต่อมาจะนำไปสู่การเชื่อมต่อของเส้นประสาทเกิดจุดเชื่อมต่อของไซแนปส์ (Synaptic Junction) ที่แข็งแรง (Strengthening) หรือ อ่อนแอ (Weakening) ขึ้นมา ยกตัวอย่างเช่น มีการเชื่อว่าความทรงจำใหม่จะเกิดขึ้นเมื่อมีการปรับปรุงความแข็งแรงของไซแนปส์ (Synaptic Strength) นี้ ดังนั้นกระบวนการเรียนรู้เช่นการจำหน้าเพื่อนใหม่ๆ นี้จะทำให้เกิดการเปลี่ยนแปลงอย่างมากมายของไซแนปส์

ในส่วนของโครงข่ายประสาทเทียมนั้นจะไม่เข้าถึงความยุ่งยากซับซ้อนต่างๆ ของสมองนี้ อย่างไรก็ตามมีคุณสมบัติสำคัญ สองประการที่สามารถทำได้ในระหว่างโครงข่ายประสาทชีว (Biological Neural Network) และโครงข่ายประสาทเทียม (Artificial Neural Network) ประการแรกการสร้างบล็อกของโครงข่ายทั้งสองซึ่งเป็นเครื่องมือที่ง่ายต่อการใช้คำนวณ (แม้ว่าเซลล์ประสาทเทียมจะมีความยุ่งยากน้อยกว่าเซลล์ประสาทชีวก็ตาม) ซึ่งทำให้เกิดการเชื่อมต่อระหว่างกันมากขึ้น ประการที่สองการต่อเชื่อมระหว่างเซลล์ประสาทสามารถกำหนดเป็นฟังก์ชัน (Function) ของโครงข่าย ในการกำหนดการเชื่อมต่อที่เหมาะสมจะสามารถช่วยแก้ (Solve) ปัญหาการคำนวณโดยเฉพาะได้

สมองสามารถที่จะทำงานหลายๆ งานได้อย่างรวดเร็วกว่าเครื่องคอมพิวเตอร์ที่มีความเร็วตามมาตรฐานสากลทั่วไปถึงแม้ว่าเซลล์ประสาทชีวจะมีความล่าช้ากว่ามากเมื่อเปรียบเทียบกับวงจรไฟฟ้า (10^{-3} s เปรียบเทียบกับ 10^{-9} s) ทั้งนี้เพราะโครงสร้างการเชื่อมต่อแบบขนานอย่างจำนวนมากหลายเส้นทางของโครงข่ายประสาทชีวซึ่งทุกเซลล์ประสาทจะทำงานในเวลาเดียวกันซึ่งแตกต่างจากโครงข่ายประสาทเทียมที่จะมีการแบ่งส่วนการใช้งานโครงสร้างทางขนานนี้

ตารางที่ 1 ความสัมพันธ์ระหว่างเซลล์ประสาทชีวะกับเซลล์ประสาทเทียม

ลำดับ	เซลล์ประสาทชีวะ	เซลล์ประสาทเทียม
1	ตัวเซลล์ (Cell body)	ยูนิต (Unit)
2	เดนไดรต์ (Dendrites)	ตัวแปรป้อนเข้า (Input)
3	ไซแนปส์ (Synapse)	ค่าถ่วงน้ำหนัก (Weight)
4	แอกซอน (Axon)	ตัวแปรผลลัพธ์ (Output)
5	ความเร็วในการทำงานช้า	ความเร็วในการทำงานสูง
6	มีเซลล์จำนวนมาก (ประมาณ 10^7 ยูนิต)	มีเซลล์จำนวนน้อยกว่า (เป็นหลักร้อย)

ส่วนวิธีการประมวลผลภายในเซลล์ประสาทแต่ละเซลล์จะมีจุดเชื่อมโยงระหว่างการทำงานเป็น 2 ลักษณะ คือ

- ลักษณะการกระตุ้น (Excitatory) เป็นการทำให้สัญญาณที่ส่งผ่านเข้ามามีความถี่สูงขึ้น

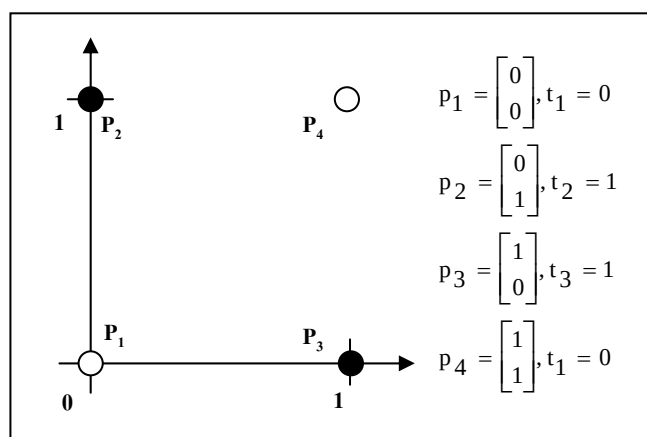
- ลักษณะการยับยั้ง (Inhibitory) เป็นการทำให้สัญญาณที่ส่งผ่านเข้ามามีความถี่ลดลง ซึ่งโครงข่ายประสาทเทียม จะมีอัตราขยายหรือหดได้เมื่อถูกกำหนดด้วยค่าถ่วงน้ำหนัก (Weight) สำหรับความสัมพันธ์ระหว่างเซลล์ประสาทชีวะกับเซลล์ประสาทเทียม แสดงไว้ในตารางที่ 1

1.2 วิวัฒนาการของโครงข่ายประสาทเทียม

การศึกษาเกี่ยวกับโครงข่ายประสาทเทียมเริ่มจากงานวิจัยของ McCulloch and Pitts ในช่วง 1940s เป็นงานวิจัยเกี่ยวกับระบบเซลล์ประสาทอย่างง่าย ในสมัยนั้นการนำแนวคิดไปใช้งานยังไม่เกิดผลเท่าที่ควร ต่อมาในปี 1949 Hebb เป็นคนแรกที่แนะนำโครงข่ายที่มีการประมวลผลจากการเรียนรู้ของเซลล์สมองซึ่งเรียกว่า Hebbian Learning ต่อมาได้นำมาใช้ในการออกแบบโครงข่ายประสาทเทียมต่างๆ อีกมากมาย การใช้เครื่องคอมพิวเตอร์ในการคำนวณโดยใช้โครงข่ายประสาทเทียมได้เริ่มต้นอย่างเป็นทางการ เมื่อปี 1956 โดย Rochester *et al.* ซึ่งเป็นการเริ่มยุคของ Artificial Intelligent (AI) อย่างแท้จริง จากนั้นได้มีนักวิจัยหลายท่านพัฒนาโครงข่าย

ประสาทเทียมขึ้นมาหลายแบบ เช่น ในปี 1958 Rosenblatt ได้พัฒนาโครงข่ายประสาทเทียมที่เรียกว่า Perceptron ขึ้น ซึ่งสามารถแก้ปัญหการแบ่งกลุ่มของตัวแปรป้อนเข้าออกเป็น 2 กลุ่มได้ โดยเป็นการแบ่งแบบเชิงเส้น โดยมีพื้นฐานการทำงานเป็นแบบ 3 ชั้น กล่าวคือ ชั้นอินพุต (Input layer) ชั้นแอสโซซิเอทีฟ (Associative layer) และชั้นเอาต์พุต (Output layer) ในปี 1960 Widrow and Hoff ได้พัฒนาโครงข่าย ADALINE (ADaptive LInear NEuron) ซึ่งต่อมามีการปรับปรุงเป็นโครงข่ายที่มีชื่อภายหลังว่า MADALINE (Multiple ADALINEs)

ในเวลาต่อมาโครงข่ายประสาทเทียมที่กลุ่มวิจัยเสนอไปนั้นไม่สามารถแก้ปัญหาคิวเฝินบางอย่างได้สำเร็จซึ่ง Minsky and Papert (1969) ได้ตีพิมพ์บทความเพื่อพิสูจน์ให้เห็นถึงข้อจำกัดของ Perceptron เช่น การแก้ปัญหา Non-linear logic (Exclusive or, XOR) ดังภาพที่ 25 ทำให้โครงข่ายประสาทเทียมหยุดชะงักเป็นเวลานาน



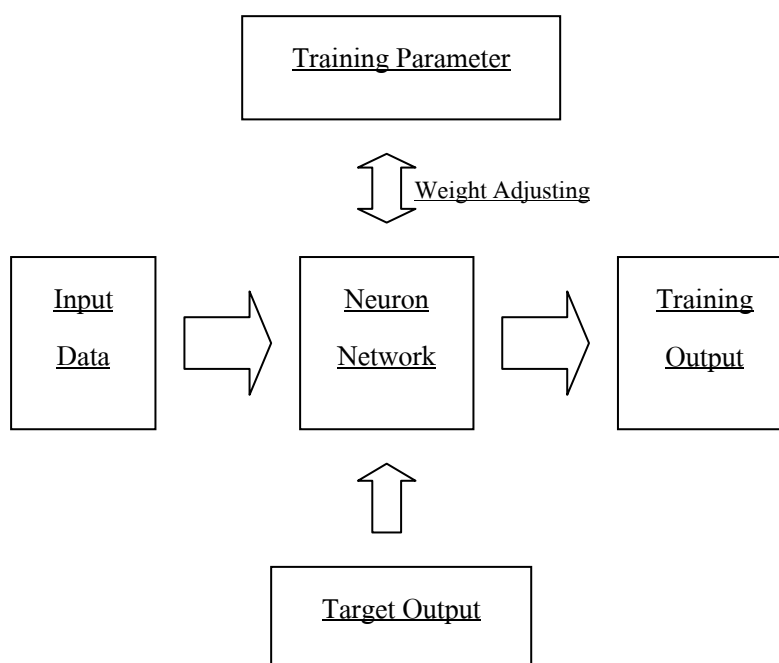
ภาพที่ 25 ลักษณะปัญหา XOR

อย่างไรก็ตาม ก็ยังมีการพัฒนาโครงข่ายประสาทเทียมอย่างต่อเนื่อง มีนักวิจัยหลายท่านให้ความสนใจเกี่ยวกับ Associative memory จนกระทั่งในปี 1982 เกิดโครงข่าย Hopfield ซึ่งเป็นโครงข่ายประเภท Recurrent และสามารถนำมาประยุกต์แก้ปัญหา Travelling salesman ในเวลานั้น วิธีแบ็คพร็อพพาร์เกชัน (Backpropagation Algorithm) ได้ถูกพัฒนาขึ้น การนำเสนอวิธีดังกล่าวทำให้งานวิจัยทางโครงข่ายประสาทเทียมกลับมาได้รับความนิยมมากขึ้นตั้งแต่ปี 1980 เมื่อเกิด

กระบวนการคิดค้น Multilayer perceptron (Rumelhart and McClland, 1986) ซึ่งทำให้ขีดจำกัดที่เคยเป็นปัญหาของ Single-layer Perceptron เช่น การแก้ปัญหา XOR หดหายไป (Martin, 1995)

1.3 กระบวนการเรียนรู้สำหรับโครงข่ายประสาทเทียม สามารถแบ่งได้บนพื้นฐานของการเรียนรู้เป็น 3 ลักษณะดังนี้ คือ

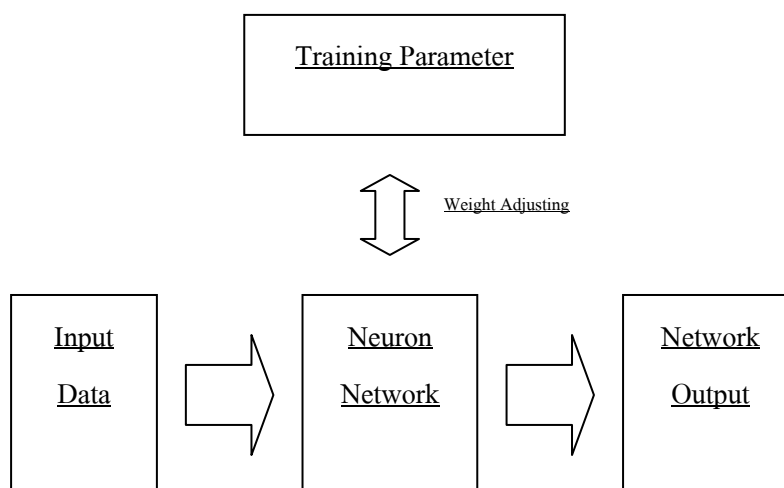
1.3.1 การเรียนรู้แบบมีครูสอน (Supervised Learning) เป็นการเรียนรู้โดยสมมติว่ามีครูคอยกำกับดูแลการทำงานของโครงข่าย ระหว่างการเรียนรู้มีการเปรียบเทียบผลการคำนวณของโครงข่ายกับเป้าหมาย ความคลาดเคลื่อนหรือความผิดพลาดที่เกิดขึ้นจะนำไปสู่การปรับปรุงตัวแปรบางตัวในโครงข่ายเพื่อเพิ่มประสิทธิภาพการทำงานของโครงข่าย รูปแบบตัวแปรป้อนเข้าอาจอยู่ในรูปของเวกเตอร์ x^p ($p=1, 2, \dots$) ซึ่งถูกส่งผ่านไปยังโครงข่ายและนำมาซึ่งรูปแบบตัวแปรผลลัพธ์ y^p ตัวแปรผลลัพธ์แต่ละรูปแบบจะถูกนำไปปรับค่าถ่วงน้ำหนักของตัวแปรป้อนเข้าแต่ละตัว เพื่อให้ค่าผิดพลาดจากการคำนวณน้อยที่สุดหรือเข้าใกล้เป้าหมายมากขึ้น ภาพที่ 26 แสดงแผนผังการเรียนรู้แบบมีครูสอน



ภาพที่ 26 แสดงแผนผังการเรียนรู้แบบมีครูสอน (Supervised Learning)

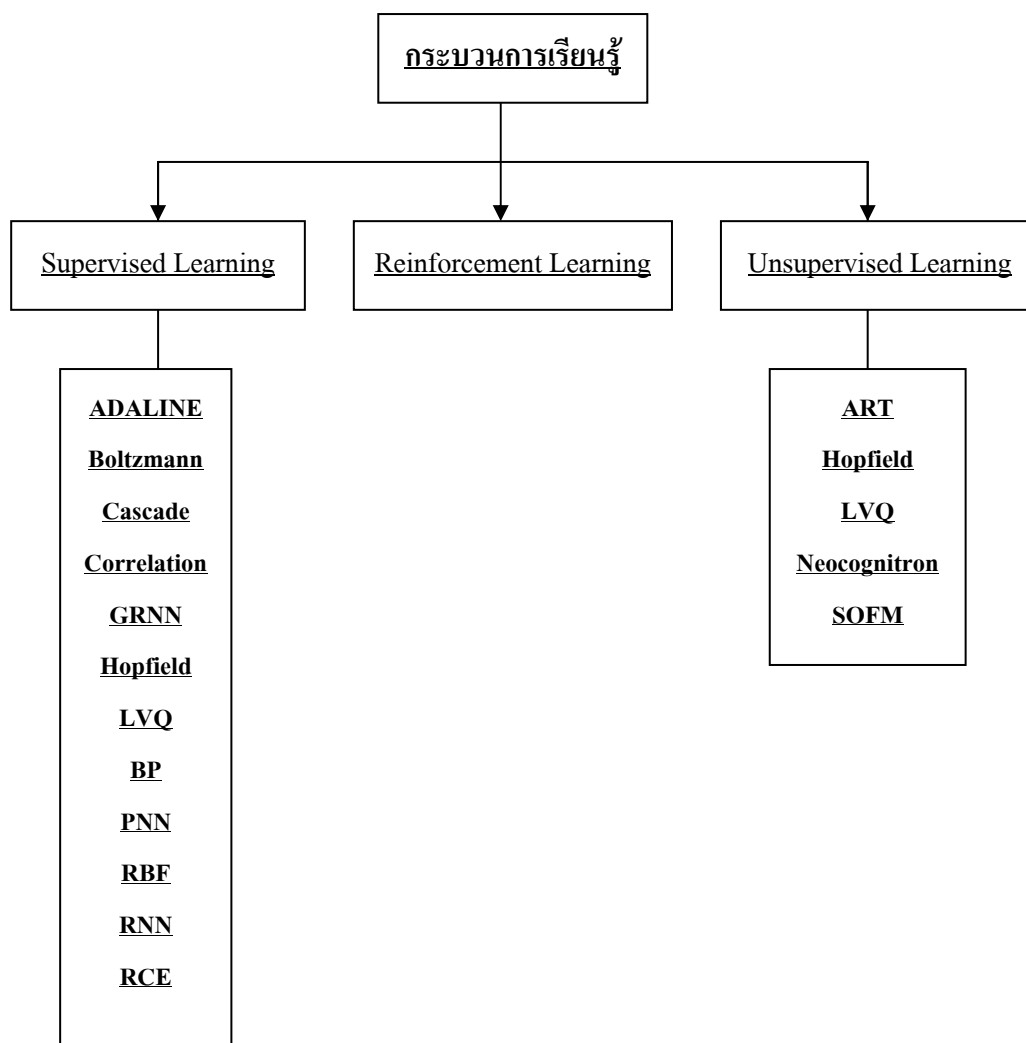
1.3.2 การเรียนรู้แบบถูกบังคับ (Reinforcement Learning) เป็นการเรียนรู้วิธีหนึ่งเป็นลักษณะที่มีครุคอยกำกับในการทำงานก็ตาม แต่คำตอบที่ถูกต้องหรือเป้าหมายไม่ได้ถูกนำมาแสดงในโครงข่าย ดังนั้นผลการคำนวณจากโครงข่ายจึงแสดงออกมาในรูปของถูกหรือผิด โครงข่ายจึงต้องใช้ประโยชน์จากข้อมูลต่างๆ เพื่อนำไปปรับปรุงการทำงานให้ดีขึ้น ถ้าผลการคำนวณถูกต้องโครงข่ายจะได้รับรางวัล (Reward) ซึ่งเป็นการเพิ่มค่าถ่วงน้ำหนักของตัวแปรป้อนเข้าบางหน่วยในทางตรงกันข้ามผลการคำนวณจากโครงข่ายออกมาผิดจะได้รับการลงโทษ (Penalty) โดยการลดค่าถ่วงน้ำหนักลง

1.3.3 การเรียนรู้แบบไม่มีครูสอน (Unsupervised Learning) โครงข่ายจะไม่สามารถรับรู้ความผิดพลาดที่เกิดขึ้นกับข้อมูลเป้าหมาย ดังนั้นโครงข่ายจะต้องเรียนรู้ให้สามารถค้นพบโครงสร้างที่เหมาะสมกับรูปแบบข้อมูลตัวแปรป้อนเข้า โดยการปรับตัวเข้าสู่การกระจายตัวเชิงสถิติของข้อมูลตัวแปรป้อนเข้า การเรียนรู้แบบนี้จะใช้วิธีการสร้างความเข้มแข็งให้กับค่าถ่วงน้ำหนักที่เชื่อมโยงกับหน่วยตัวแปรป้อนเข้าที่สนองตอบได้ดีกับตัวแปรผลลัพธ์ ในขณะเดียวกันก็จะลดความสำคัญของหน่วยตัวแปรป้อนเข้าที่อ่อนแอลง ภาพที่ 27 แสดงแผนผังการเรียนรู้แบบไม่มีครูสอน



ภาพที่ 27 แผนผังการเรียนรู้แบบไม่มีครูสอน (Unsupervised Learning)

กระบวนการเรียนรู้เปรียบเสมือน กระบวนการเปรียบเทียบ (Calibration) ในแบบจำลองคณิตศาสตร์ทั่วไป อย่างไรก็ตามโครงข่ายประสาทเทียมจะแบ่งข้อมูลที่มีอยู่เป็น 2 ส่วน ข้อมูลส่วนแรกจะใช้ในการสอนหรือการเรียนรู้โครงข่าย (Network Training) ข้อมูลส่วนที่สองจะใช้ในการทดสอบโครงข่าย (Network Testing) วัตถุประสงค์ในการเรียนรู้ข้อมูลแรกเพื่อที่จะหาค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างหน่วยในชั้นต่างๆ ของโครงข่าย ที่ซึ่งจะทำให้ผลการคำนวณจากโครงข่ายมีค่าใกล้เคียงกับข้อมูลเป้าหมาย ดังนั้นในการเรียนรู้ในช่วงนี้จึงมีความต้องการข้อมูลที่มากพอ หรือรูปแบบของข้อมูลควรมีอย่างพอเพียงด้วย สำหรับข้อมูลที่เหลือซึ่งเป็นข้อมูลในส่วนที่ 2 จะใช้ในการทดสอบโครงข่ายต่อไป



ภาพที่ 28 แผนผังโครงข่ายประสาทเทียมแบ่งตามประเภทกระบวนการเรียนรู้

ปัจจุบันมีโครงข่ายประสาทเทียมมากกว่า 22 แบบที่ใช้งานอยู่ สามารถแบ่งประเภทตามกระบวนการเรียนรู้ได้ดังภาพที่ 28

1.4 การประยุกต์ใช้โครงข่ายประสาทเทียม โครงข่ายประสาทเทียมเป็นแบบจำลองการทำงานของระบบประสาทส่วนกลางที่มีโครงสร้างเป็นลักษณะของโครงข่ายเชื่อมโยงกันระหว่างหน่วย ซึ่งสามารถที่จะรับรู้ข้อมูลและปรับตัวเข้ากับสถานการณ์หรือสิ่งแวดล้อมที่กำลังเผชิญอยู่นักวิจัยหลายท่านเชื่อว่า โครงข่ายประสาทเทียมเป็นเครื่องมือชนิดหนึ่งที่ใช้ในการสร้างระบบคอมพิวเตอร์อัจฉริยะอย่างได้ผล (Intelligent computer system) นอกจากนี้ โครงข่ายประสาทเทียมยังเป็นเครื่องมือที่มีประสิทธิภาพในการประยุกต์ใช้ในงานต่างๆ ที่เกี่ยวข้องกับการคำนวณและการจดจำ ไม่ว่าจะเป็นสาขาทางด้านวิทยาศาสตร์ ซึ่งมีอยู่ด้วยกันหลายแขนง หรือแม้กระทั่งทางด้านธุรกิจและ เศรษฐศาสตร์ก็ตาม เช่น

1.4.1 การทำนายพยากรณ์ (Forecasting) เป็นการนำโครงข่ายประสาทเทียมมาใช้ในการพยากรณ์เหตุการณ์ต่าง ๆ หรือแม้แต่ผลิตภัณฑ์ หุ่น เป็นต้น สามารถใช้ได้อย่างมีประสิทธิภาพ เช่น การพยากรณ์แนวโน้มของการขายผลิตภัณฑ์ตัวใหม่ที่เพิ่งออกมาสู่ตลาด ทำให้บริษัทผู้ผลิตสามารถวางแผนการทางตลาดได้อย่างมีประสิทธิภาพมากขึ้น การพยากรณ์สภาพอากาศ เป็นงานหนึ่งที่นิยมใช้กันในปัจจุบัน ในส่วนของการจัดการเรื่องธุรกิจการบิน สามารถประยุกต์ใช้โครงข่ายประสาทเทียม หาแนวโน้มการเดินทางของลูกค้า เพื่อใช้ในการวางแผนสนองตอบกับความต้องการในอนาคต ทางด้านธุรกิจการเงินโครงข่ายประสาทเทียมใช้ในการพยากรณ์อัตราแลกเปลี่ยนและการจัดการหลักทรัพย์เพื่อการลงทุน นอกจากนี้ทางด้านการทหารยังได้นำโครงข่ายประสาทเทียมไปใช้ในการวิเคราะห์จุดยุทธศาสตร์ที่ตั้งของข้าศึก

1.4.2 การจดจำข้อมูล (Content Addressable Memories) เป็นการนำเอาความสามารถของโครงข่ายประสาทเทียมที่สามารถเรียนรู้และจดจำรูปแบบในลักษณะของความสัมพันธ์ระหว่างตัวแปรป้อนเข้าและตัวแปรผลลัพธ์ ถ้าใช้วิธีอื่นในการแก้ปัญหาอื่นๆ เช่น สมการอนุพันธ์เชิงเส้นประมาณแบบจำลองที่เป็น Nonlinear ให้เป็น Linear ซึ่งจะมีค่า Error มากขึ้นเมื่อจุดที่ทำนายอยู่ห่างจาก Operating Point มากขึ้น แต่โครงข่ายประสาทเทียมเป็นการสร้างแบบจำลองจากข้อมูลการทดลองจริงทำให้ไม่มีผลกระทบจากความไม่เป็น Nonlinear

1.4.3 การควบคุม (Control) เป็นการนำโครงข่ายประสาทเทียมไปใช้ในการเรียนรู้ และควบคุมการขับเคลื่อนหุ่นยนต์ รวมไปถึงหน้าที่ที่ต้องช่วยในการจับจีพียูผลิตภัณฑ์หลายอย่างที่ผลิตในประเทศญี่ปุ่นได้นำโครงข่ายประสาทเทียมมาช่วยในการบริหารและจัดการด้าน ขบวนการผลิต

1.4.4 การบีบอัดข้อมูล (Data Compression) โดยโครงข่ายประเภนี้ใช้การเรียนรู้ และการคำนวณเพื่อลดจำนวนอินพุตที่ใช้ในงานที่มีข้อมูลจำนวนมาก ๆ ถูกเก็บเอาไว้ และถูก ประมวลผล ตัวอย่างเช่น กรณีของการทำงานของภาพถ่ายดาวเทียม

1.4.5 การวินิจฉัยโรค (Diagnostics) เป็นการใช้โครงข่ายประสาทเทียมในงานวิจัย ทางด้านการแพทย์ โดยนำเอาโครงข่ายประสาทเทียมมาใช้ในการวินิจฉัยโรคที่ซับซ้อน โดย โครงข่ายประสาทเทียมจะต้องเก็บความสัมพันธ์ระหว่างตัวแปรป้อนเข้า ที่เป็นลักษณะอาการ หรือ พฤติกรรมที่ผิดปกติ กับตัวแปรผลลัพธ์ที่เป็นการระบุโรคของผู้ป่วย

1.4.6 การรวมข้อมูล (Mutisensor Data Fusion) เป็นการนำโครงข่ายประสาทเทียม ไปประยุกต์ใช้ในกระบวนการประมวลผลข้อมูลต่าง ๆ จากแหล่งข้อมูลหลายแบบ ตัวอย่างเช่น การ นำไปใช้ในการจัดทำระบบความปลอดภัยข้อมูล ที่สามารถพิจารณาได้จากหน้าตาของผู้ใช้ (Image Processing)

1.4.7 การหาแนวทางที่เหมาะสม (Optimization) เป็นการแก้ปัญหาเพื่อหาความ เหมาะสมของแนวทาง เพื่อให้ได้ผลลัพธ์ของปัญหาอย่างมีประสิทธิภาพที่สุด

1.4.8 การจำได้ (Pattern recognition) เป็นการใช้โครงข่ายประสาทเทียมเพื่อการ เรียนรู้และรับความรู้สึกของปัญหาที่มีรูปแบบสลับซับซ้อน เช่น การมองเห็นวัตถุต่าง ๆ การ วิเคราะห์เสียงพูดเพื่อแปลเป็นข้อความ เป็นต้น

2. ทฤษฎีที่เกี่ยวข้องของโครงข่ายประสาทเทียม

2.1 แบบจำลองประสาท (Neuron Models)

2.1.1 เซลล์ประสาทเทียม 1 ตัวแปรป้อนเข้า (Single-Input Neuron) เซลล์ประสาทเทียมที่จำลองขึ้นแสดงดังภาพที่ 29 มีหลักการทำงานโดยนำตัวแปรป้อนเข้า (Input) p คูณกับค่าถ่วงน้ำหนัก (Wight) w ซึ่งจะได้ wp แล้วนำมาบวกกับ ผลคูณของไบแอส (Bias) b กับตัวแปรป้อนเข้าของไบแอส ที่มีค่าเท่ากับ 1 จะได้

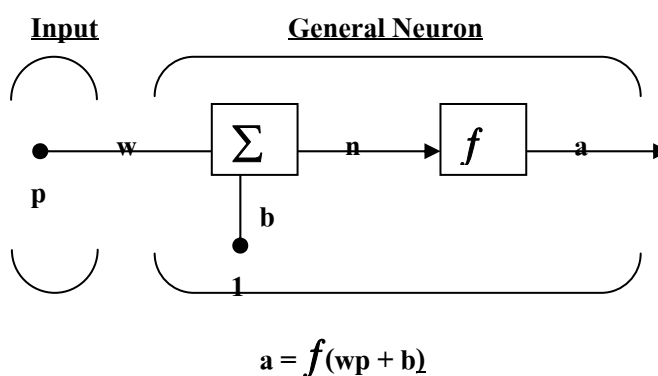
$$n = wp + b \quad (48)$$

โดย n คือ ตัวแปรป้อนเข้าสุทธิ (Net Input) และค่าตัวแปรผลลัพธ์ของเซลล์ประสาท (Neuron Output)

โดย a มีค่าเท่ากับ

$$a = f(wp + b) \quad (49)$$

โดย f คือ ทรานส์เฟอร์ฟังก์ชัน (Transfer Function)
 w และ b คือ ตัวแปรที่ปรับค่าได้ของเซลล์ประสาทเทียม



ภาพที่ 29 แสดงแบบจำลองเซลล์ประสาทเทียม

2.1.2 เซลล์ประสาทเทียมตัวแปรป้อนเข้าหลายตัว (Multiple-Input Neuron)

โดยทั่วไปเซลล์ประสาทเทียมมีตัวแปรป้อนเข้ามากกว่าหนึ่งตัวแปร ซึ่งในที่นี้ให้มิตัวแปรป้อนเข้า R ตัวแปร ดังแสดงในภาพที่ 30

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b \quad (50)$$

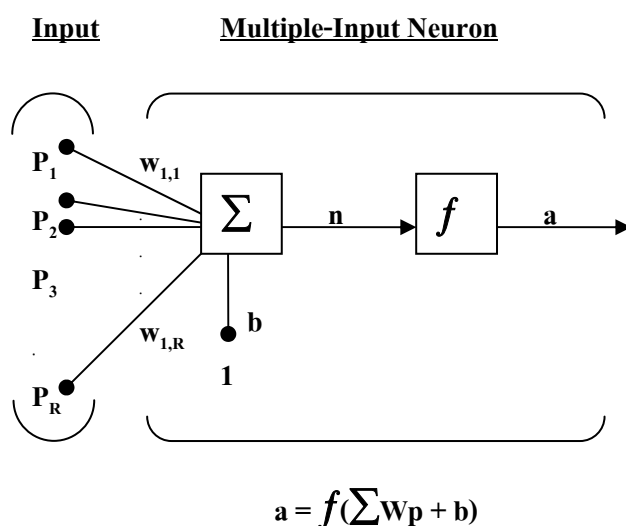
จากสมการที่ (50) แสดงเมตริกซ์ ได้ดังนี้

$$n = \sum W_p + b \quad (51)$$

โดย W คือ เมตริกซ์ของค่าถ่วงน้ำหนัก ($1 \times R$) และ p คือ เวกเตอร์ของตัวแปรป้อนเข้า ($R \times 1$)

$$a = f(\sum W_p + b) \quad (52)$$

ดังนั้นจึงได้ใช้เมตริกซ์ในการอธิบายการทำงานของ Multiple-Input Neuron ได้ง่ายขึ้น



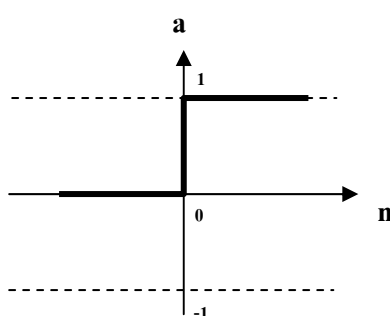
ภาพที่ 30 แสดงเซลล์ประสาทเทียมที่มีค่าอินพุตหลายค่า (Multiple-Input Neuron)

2.2 ทรานส์เฟอร์ฟังก์ชัน (Transfer Function) ในระบบโครงข่ายประสาทเทียมนั้น ชนิดของทรานส์เฟอร์ฟังก์ชัน จะขึ้นอยู่กับผู้ออกแบบและตัวแปรผลลัพธ์จริง (Actual Output) มีการใช้ทรานส์เฟอร์ฟังก์ชันอยู่หลายรูปแบบ แต่รูปแบบที่ใช้กันโดยทั่วไปมีอยู่ 3 แบบคือ

2.2.1 ฟังก์ชันฮาร์ดลิมิตทรานส์เฟอร์ (Hard Limit Transfer Function) เป็นทรานส์เฟอร์ฟังก์ชันที่แยกประเภทตัวแปรผลลัพธ์ ออกเป็น 2 ประเภทที่แตกต่างกันอย่างชัดเจน ซึ่งมีค่าตัวแปรผลลัพธ์เป็น 0 เมื่อตัวแปรป้อนเข้ามีค่าน้อยกว่า 0 และตัวแปรผลลัพธ์มีค่าเป็น 1 เมื่อตัวแปรป้อนเข้ามีค่ามากกว่า 0 ดังแสดงในภาพที่ 31 มีความสัมพันธ์ระหว่างค่าตัวแปรป้อนเข้าและตัวแปรผลลัพธ์ดังนี้

$$a = 0 \quad \text{เมื่อ } n < 0$$

$$a = 1 \quad \text{เมื่อ } n \geq 0$$



ภาพที่ 31 ฟังก์ชันฮาร์ดลิมิตทรานส์เฟอร์ (Hard Limit Transfer Function)

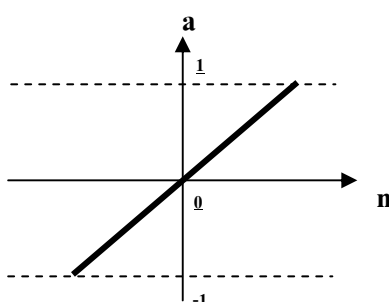
2.2.2 ฟังก์ชันลิเนียร์ทรานส์เฟอร์ (Linear Transfer Function) เป็นทรานส์เฟอร์ฟังก์ชันที่มีความสัมพันธ์ของตัวแปรป้อนเข้าและตัวแปรผลลัพธ์เป็นเส้นตรง ดังภาพที่ 32 ซึ่งค่าตัวแปรผลลัพธ์ที่ได้จะมีค่าเท่ากับค่าของตัวแปรป้อนเข้า

$$a = n$$

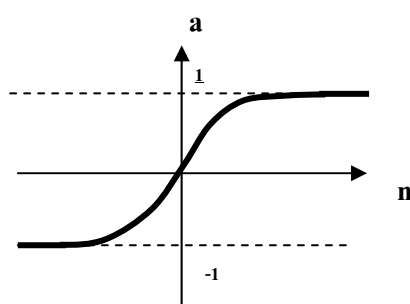
2.2.3 ฟังก์ชันไฮเปอร์โบลิคแทนเจนต์ซิกมอยด์ (Hyperbolic Tangent Sigmoid Transfer Function) เป็นทรานส์เฟอร์ฟังก์ชันที่ใช้กับตัวแปรป้อนเข้าที่มีค่าอยู่ระหว่าง $-\infty$ ถึง ∞

และจะให้ค่าตัวแปรผลลัพธ์อยู่ระหว่าง -1 ถึง 1 จะนิยมใช้ในการออกแบบโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) ซึ่งถูกสอน (Train) และเรียนรู้โดยใช้วิธีแบ็กพร็อพเพรเกชัน (Backpropagation) โดยทรานส์เฟอร์ฟังก์ชันแบบไฮเพอร์โบลิคแทนเจนต์ซิกมอยด์นี้มีลักษณะดังแสดงในภาพที่ 33 และสามารถเขียนความสัมพันธ์ออกมาเป็นสมการได้คือ

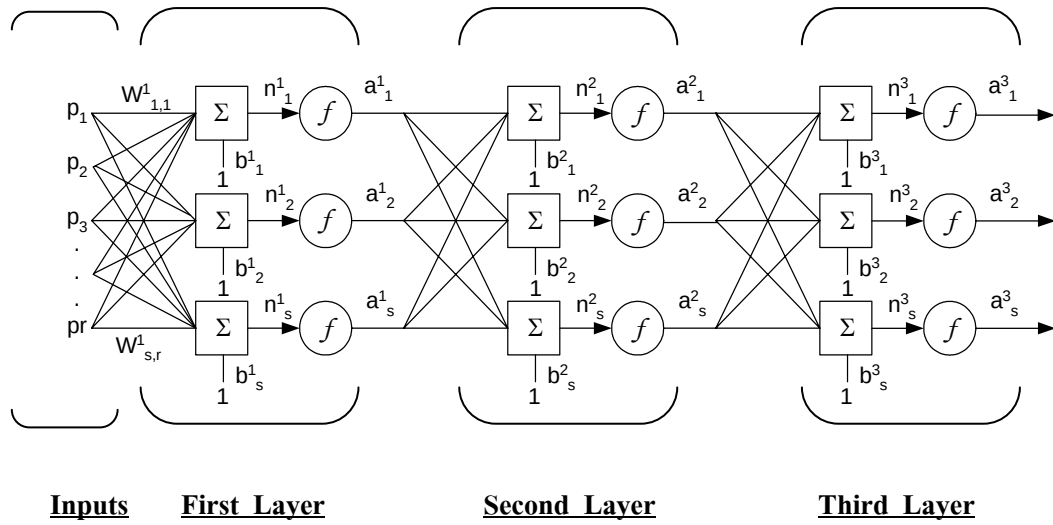
$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$



ภาพที่ 32 ฟังก์ชันลิเนียร์ทรานส์เฟอร์ (Linear Transfer Function)



ภาพที่ 33 ฟังก์ชันไฮเพอร์โบลิคแทนเจนต์ซิกมอยด์ (Hyperbolic Tangent Sigmoid Transfer Function)



ภาพที่ 34 โครงข่ายประสาทเทียมแบบ Multilayer Perceptron

2.3 Backpropagation Algorithm กระบวนการสอน (Train) คือกระบวนการสำหรับปรับน้ำหนัก และไบแอสในโครงข่ายประสาทเทียมจากภาพที่ 34 เป็นตัวอย่างโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอน (Perceptron) หลายชั้นที่มีจำนวนชั้น 3 ชั้น โดยชั้นที่ 1 และ 2 นี้เรียกว่า ชั้นซ่อน (Hidden Layer) และชั้นที่ 3 เป็นชั้นคำตอบ ด้วยความซับซ้อนของโครงข่ายประสาทเทียมแบบหลายชั้นนี้ ทำให้ต้องมีการพัฒนาวิธีการสอนสำหรับโครงข่ายประสาทเทียมแบบนี้โดยเฉพาะ ซึ่งวิธีดังกล่าวคือ วิธีแบ็คพร็อพพาร์เกชั่น โดยการปรับน้ำหนักและไบแอส เป็นไปตามสมการดังต่อไปนี้

$$W^m(k+1) = W^{m+1}(k) - \alpha s^m (a^{m-1})^T \quad (53)$$

$$b^m(k+1) = b^m(k) - \alpha s^m \quad (54)$$

เมื่อ $m = 1, 2, 3, \dots, M$

โดย s^m คือ ความไวในการตอบสนอง (Sensitivity) ในชั้นที่ m ซึ่งหาได้จาก

$$s^M = -2\dot{F}^M(\mathbf{n}^M)(\mathbf{t} - \mathbf{a}) \quad (55)$$

$$\mathbf{s}^m = \dot{\mathbf{F}}^m(\mathbf{n}^m)(\mathbf{W}^{m+1})^T \mathbf{s}^{m+1} \quad (56)$$

เมื่อ $m = M-1, \dots, 2, 1$ และ

$$\dot{\mathbf{F}}^m(\mathbf{n}^m) = \begin{bmatrix} \dot{f}^m(n_1^m) & 0 & \dots & 0 \\ 0 & \dot{f}^m(n_2^m) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \dot{f}^m(n_{s^m}^m) \end{bmatrix} \quad (57)$$

สำหรับทฤษฎีและบทพิสูจน์ที่เกี่ยวข้องกับวิธีแบ็คพรีอพาร์เกชัน สามารถศึกษาเพิ่มเติมได้จากภาคผนวก ข

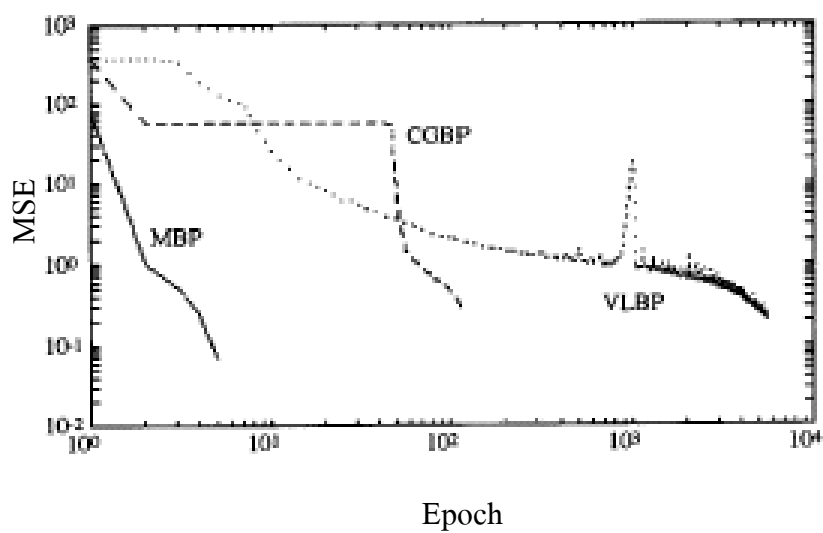
2.4 Levenberg-Marquardt Algorithm สำหรับกระบวนการสอนด้วยวิธีแบ็คพรีอพาร์เกชัน ได้มีการปรับปรุงและพัฒนาเพื่อให้การสอนด้วยวิธีนี้มีประสิทธิภาพดีขึ้น เช่น มีการใส่ค่าโมเมนตัม (Momentum) การปรับอัตราการเรียนรู้ (Learning Rate) เป็นต้น รวมทั้งการใช้วิธี Levenberg-Marquardt ก็เป็นอีกวิธีที่ทำให้การสอนมีประสิทธิภาพ ผู้เข้าสู่เป้าหมายที่ต้องการได้อย่างรวดเร็ว แต่วิธี Levenberg-Marquardt นี้จำเป็นต้องใช้หน่วยความจำและความสามารถในการคำนวณของคอมพิวเตอร์สูง (Martin, 1995)

Hagan and Menhaj (1994) ได้แสดงให้เห็นว่า วิธี Levenberg-Marquardt หรือ Marquardt Backpropagation (MBP) นั้นมีประสิทธิภาพในการสอนที่สูงกว่าเมื่อเทียบกับวิธีการปรับอัตราการเรียนรู้ (Variable Learning Rate Backpropagation, VLBP) และวิธี Conjugate gradient (Conjugate Gradient Backpropagation, CGBP) โดยปัญหาที่นำมาใช้ในการเปรียบเทียบการสอนของทั้ง 3 วิธีนี้มีอยู่ด้วยกัน 4 ปัญหา คือ

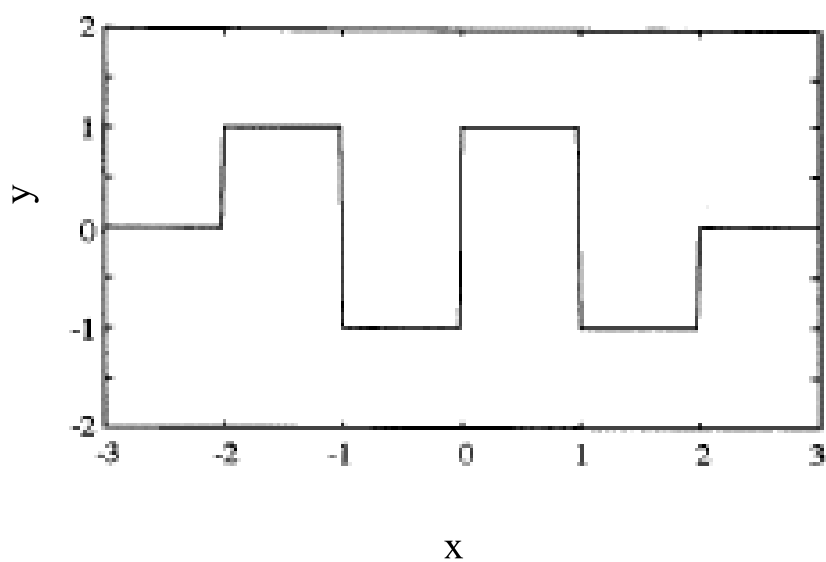
ก) ปัญหาที่ 1 การประมาณฟังก์ชัน sine ของสมการ

$$y = \frac{1}{2} + \frac{1}{4} \sin(3\pi x) \quad (58)$$

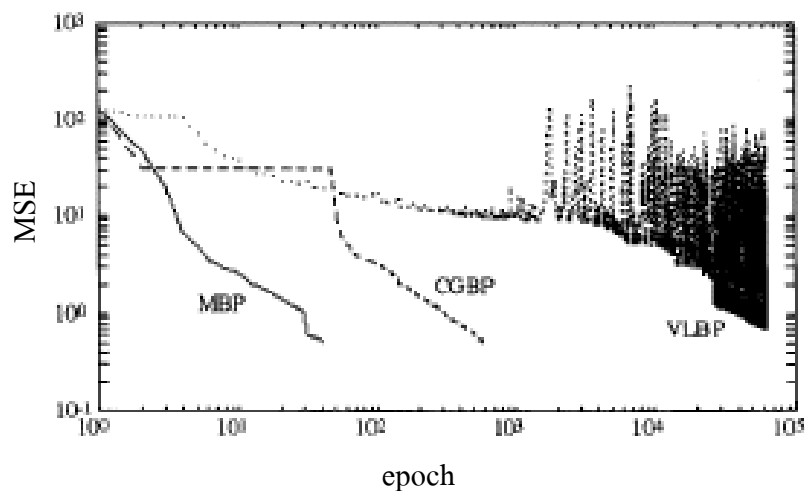
พบว่าค่าผลรวมของความคลาดเคลื่อนยกกำลังสอง (Sum of Squares Error) เทียบกับจำนวนรอบของการสอนเป็นดังภาพที่ 35



ภาพที่ 35 ผลรวมความคลาดเคลื่อนยกกำลังสองเทียบกับจำนวนรอบของการสอนของปัญหาที่ 1



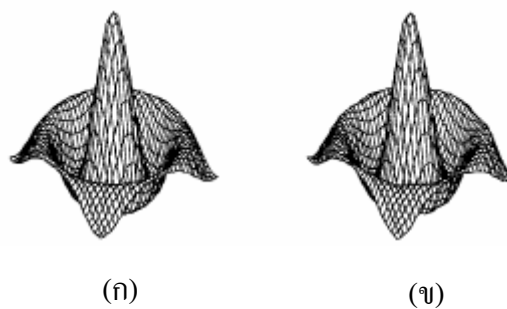
ภาพที่ 36 ฟังก์ชัน Square wave ในปัญหาที่ 2



ภาพที่ 37 ผลรวมความคลาดเคลื่อนยกกำลังสองเทียบกับจำนวนรอบของการสอนของปัญหาที่ 2

ข) ปัญหาที่ 2 การประมาณค่า Square wave ของภาพที่ 36 เมื่อเปรียบเทียบค่าของผลรวมความคลาดเคลื่อนยกกำลังสองกับจำนวนรอบการสอน พบว่าได้ผลดังแสดงในภาพที่ 37

ค) ปัญหาที่ 3 การประมาณฟังก์ชัน sine 2 มิติ ซึ่งรูปแบบของปัญหามีลักษณะดังภาพที่ 38 (ก) ส่วนภาพที่ 38 (ข) เป็นภาพที่ได้จากการสอนด้วยวิธี Levenberg-Marquardt

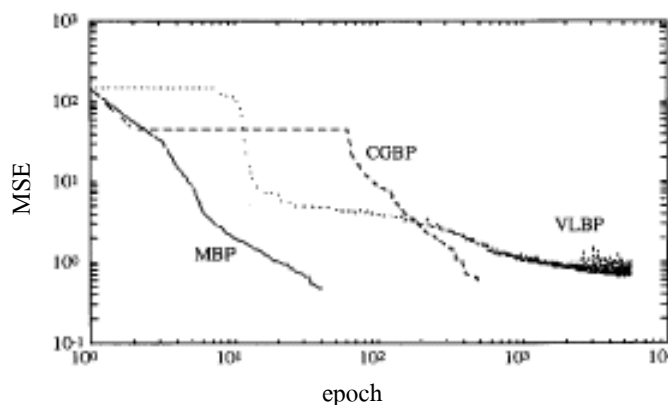


ภาพที่ 38 ฟังก์ชัน sine 2 มิติ

(ก) เป็นภาพปัญหาด้านแบบ

(ข) เป็นภาพที่ได้จากการสอนด้วยวิธี Levenberg-Marquardt

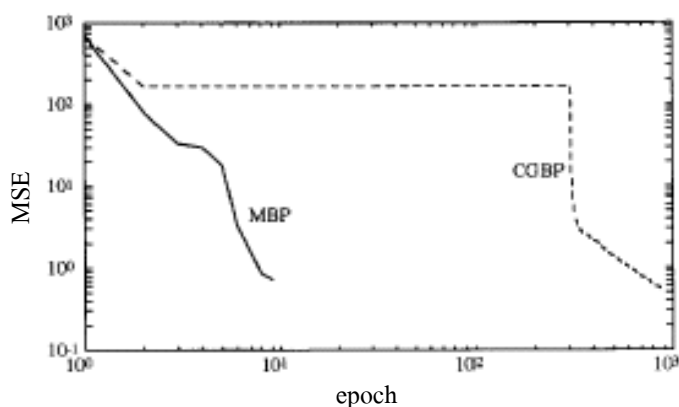
ผลที่ได้จากปัญหาที่ 3 เมื่อเปรียบเทียบทั้ง 3 วิธีได้แสดงไว้ในภาพที่ 39



ภาพที่ 39 ผลรวมความคลาดเคลื่อนยกกำลังสองเทียบกับจำนวนรอบของการสอนของปัญหาที่ 3

ง) ปัญหาที่ 4 การประมาณค่าฟังก์ชัน 4 มิติ ได้ผลดังแสดงในภาพที่ 40 โดยในปัญหาที่ 4 นี้วิธี การปรับอัตราการเรียนรู้ (VLBP) ไม่ได้ถูกนำมาพิจารณา เพราะใช้เวลานานมากเกินไป

$$y = \sin(2\pi x_1) x_2^2 x_3^2 x_4^2 e^{-(x_1 + x_2 + x_3 + x_4)} \quad (59)$$



ภาพที่ 40 ผลรวมความคลาดเคลื่อนยกกำลังสองเทียบกับจำนวนรอบของการสอนของปัญหาที่ 4

จากปัญหาทั้ง 4 จะพบว่า วิธี Levenberg-Marquardt หรือ Marquardt backpropagation (MBP) นั้น สามารถเข้าสู่เป้าหมายที่ต้องการได้เร็วกว่าวิธีอื่น ๆ โดยจะสังเกตได้จากจำนวนรอบที่ใช้จะน้อยกว่าอีก 2 วิธีที่เหลือ