

บทที่ 3

โครงข่ายประสาทเทียม

3.1 บทนำ

ปัญญาประดิษฐ์ (Artificial Intelligence – AI) หรือ ปัญญาของเครื่องจักร (Machine Intelligence) คือการศึกษาเกี่ยวกับความฉลาดของสิ่งมีชีวิต โดยเฉพาะมนุษย์ เพื่อที่จะสร้างความฉลาดเทียมให้กับสิ่งที่ไม่มีชีวิต เช่น หุ่นยนต์ เครื่องจักร และคอมพิวเตอร์ ขึ้นมาได้ โดยทั่วไปแล้วเราถือกันว่าปัญญาประดิษฐ์เป็นสาขาหนึ่งของศาสตร์ด้านวิทยาการคอมพิวเตอร์ แต่นักวิทยาศาสตร์หลาย ๆ ท่านเชื่อว่าการศึกษาปัญญาประดิษฐ์นั้นเมื่อมองให้ลึกซึ้งแล้ว ครอบคลุมกว้างไกลกว่าศาสตร์ด้านวิทยาการคอมพิวเตอร์ โดยได้โยงไปถึงสาขาอื่นๆ เช่น ปรัชญา จิตวิทยา และ ชีววิทยา เนื่องจากศาสตร์เหล่านี้ ได้มีการศึกษาเรื่องกระบวนการการคิด การให้เหตุผลหรือการอนุมานและการทำงานของสมองมานานแล้ว ซึ่งจริง ๆ แล้ว นับว่าความเชื่อนี้ถูกต้อง เนื่องจากแนวคิดหลาย ๆ อย่างในศาสตร์นี้ได้มาจากการปรับปรุงเพิ่มเติมจากศาสตร์อื่นๆ [18]

ในบทนี้จะกล่าวถึงเทคโนโลยีโครงข่ายประสาทเทียม(Artificial Neural Networks, ANNs) ซึ่งเป็นหนึ่งในสาขาของปัญญาประดิษฐ์ โดยเทคโนโลยีประเภทนี้จะเลียนแบบมาจากระบบประสาทของมนุษย์

ซึ่งปัจจุบันได้มีการประยุกต์นำเทคโนโลยีโครงข่ายประสาทเทียม มาใช้งานในด้านต่าง ๆ มากมาย เช่น ทางด้านโทรคมนาคมได้นำมาใช้ในการบีบข้อมูลและภาพ (Image and Data Compression) ระบบนักบินอัตโนมัติ (Auto Pilot Aircraft) การแปลภาษาพูด (Real Time Translation of spoken language) การจดจำเสียง (Voice recognition) การจดจำภาพ (Image recognition) การจดจำรูปแบบ (Pattern recognition) ซึ่งเครื่องมือและอุปกรณ์ต่าง ๆ ที่นำเอาโครงข่ายประสาทเทียมมาช่วยวิเคราะห์นั้นจะมีความฉลาดมากขึ้น และมีระบบความคิดที่มีการทำงานในลักษณะคล้ายกับมนุษย์

โครงข่ายประสาทเทียมได้ถูกพัฒนา ค้นคว้า วิจัยและนำไปใช้ในด้านต่าง ๆ มากขึ้น อาทิ วิทยาศาสตร์ด้านระบบประสาท , การเงินการธนาคาร , ด้านคณิตศาสตร์ , การแพทย์ , ระบบสื่อสาร รวมไปถึงงานทางด้านวิศวกรรมไฟฟ้าอีกด้วย จึงถือได้ว่า โครงข่ายประสาทเทียมเป็นศาสตร์แขนงหนึ่งที่มีความสนใจเป็นอย่างมากในช่วงเวลานี้ ดังแสดงในตารางที่ 3.1 โดยจะแสดงการประยุกต์ใช้โครงข่ายประสาทเทียมในปัจจุบัน สามารถกล่าวได้ว่าโครงข่ายประสาทเทียมยังคงได้รับการพัฒนาทั้งด้านทฤษฎี การออกแบบการประยุกต์ อีกต่อไปในอนาคต

ตารางที่ 3.1 แสดงการประยุกต์ใช้โครงข่ายประสาทเทียมในปัจจุบัน [19]

	Fault Diagnosis	Fraud Detection	Forecasting	Signal/Image Processing	Pattern detection in databases
Manufacturing	X	X			
Retailing		X	X		X
Finance and Insurance		X	X		X
engineering	X			X	
Production Control	X		X		
Sale & Marketing			X		X

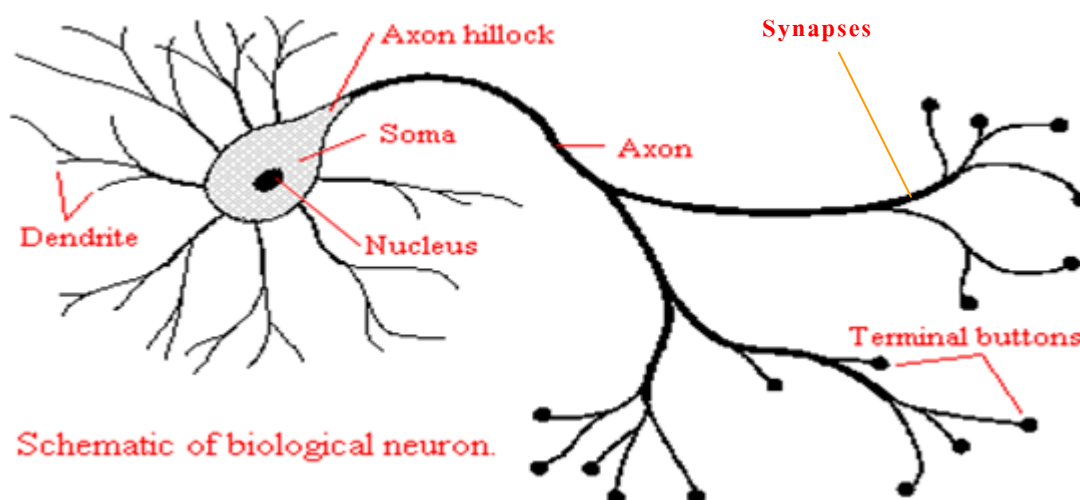
ในงานด้านไฟฟ้ากำลังนั้นได้นักวิจัยด้านไฟฟ้ากำลังหลายท่านได้นำเอาโครงข่ายประสาทเทียมมาประยุกต์ใช้ในแก้ปัญหาเกี่ยวกับไฟฟ้ากำลัง เช่น การแก้ปัญหาโหลดไหล (Load flow) , การพยากรณ์ความต้องการใช้ไฟฟ้า (Load forecasting) และปัญหาการจ่ายโหลดอย่างมีประสิทธิภาพ (Economics load dispatch) เป็นต้น ในบทนี้เป็นการอธิบายหลักการของโครงข่ายประสาทเทียม โดยจะกล่าวถึงโครงสร้างของโครงข่ายประสาทเทียม , ชนิดของหลักการเรียนรู้ของโครงข่ายประสาทเทียม รวมไปถึงโครงข่ายประสาทเทียมชนิดต่าง ๆ เพื่อนำมาประยุกต์ใช้ในการวิจัย

3.2 แรงจูงใจจากระบบชีววิทยา (Biological Motivation) [20]

โครงข่ายประสาทเทียมเป็นโมเดลทางคณิตศาสตร์ที่มีวิธีการประมวลผลข้อมูลที่ถูกพัฒนาโดยอาศัยหลักการทำงานของโครงข่ายประสาทชีววิทยาของมนุษย์ ซึ่งประกอบไปด้วยเซลล์ประสาทหรือนิวรอน (neurons) และจุดประสาท หรือ ไซแนปส์ (synapses) โดยโครงข่ายที่จำลองขึ้นมานั้นจะมีความซับซ้อนและมีการเชื่อมต่อถึงกันในแต่ละหน่วย ซึ่งแต่ละหน่วยจะมีการรับค่าข้อมูลเข้ามาทำการประมวลผลหลังจากนั้นก็จะส่งข้อมูลออกไปยังหน่วยอื่นต่อไป อันเป็นการเลียนแบบระบบประสาทที่มีการเชื่อมต่อกันระหว่างเซลล์ประสาทเข้าไว้ด้วยกัน และมีความสามารถที่จะเรียนรู้ได้ โดยโครงข่ายจำลองจะมีการเรียนรู้โดยเก็บความรู้เหล่านั้นไว้ในโครงข่ายด้วยค่า weight และค่า bias โดยค่าเหล่านี้จะสามารถปรับเปลี่ยนได้เมื่อมีการเรียนรู้สิ่งใหม่

องค์ประกอบพื้นฐานของสมองมนุษย์คือ นิวรอน (Neurons) ก็คือกลุ่มของเซลล์ประสาทที่ทำหน้าที่ในการจดจำ คิด นำความรู้ และประสบการณ์ที่บันทึกไว้ในสมองมาใช้ในการ

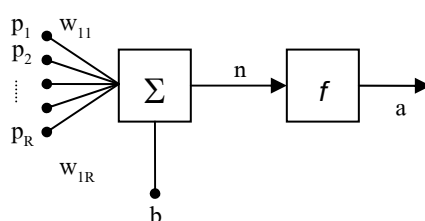
ตัดสินใจในชีวิตประจำวัน แต่ละนิวรอนอาจเชื่อมต่อกับนิวรอนอื่น ๆ มากถึง 200,000 นิวรอน พลังสมองของมนุษย์เกิดจากนิวรอนจำนวนมากและระบบการเชื่อมต่อแบบซับซ้อนของนิวรอนจำนวนมาก ในแต่ละนิวรอนจะประกอบด้วยปลายในการรับกระแสประสาท เรียกว่า "เดนไดรต์" (Dendrite) ซึ่งเป็นอินพุตและปลายในการส่งกระแสประสาทเรียกว่า "แอกซอน" (Axon) ซึ่งเป็นเหมือนเอาต์พุต ของเซลล์ โดยเซลล์เหล่านี้จะทำงานด้วยปฏิกิริยาไฟฟ้าเคมี เมื่อมีการกระตุ้นด้วยสิ่งเร้าภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรต์เข้าสู่นิวเคลียสซึ่งจะเป็นตัวตัดสินใจว่าต้องกระตุ้นเซลล์อื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็จะกระตุ้นเซลล์อื่น ๆ ต่อไปผ่านทางแอกซอนของมัน ตามโมเดลนี้ข่าขงานประสาทเกิดจากการเชื่อมต่อระหว่างเซลล์ประสาท จนเป็นเครือข่ายที่ทำงานร่วมกัน โดยกระแสประสาทจะถูกส่งผ่านจากนิวรอนหนึ่งไปสู่อีกนิวรอนหนึ่งโดยผ่านไซแนปส์ เมื่อกระแสประสาทผ่านเข้ามายังไซแนปส์ สารเคมีในไซแนปส์จะมีผลต่อกระแสประสาทในการกระตุ้นกระแสประสาทให้มีขนาดเพิ่มขึ้น หรืออาจจะยับยั้งให้กระแสประสาทมีขนาดเล็กลง เปรียบกับโครงข่ายประสาทเทียมคือการที่สัญญาณอินพุตถูกถ่วงน้ำหนักให้มีค่ามากขึ้นหรือลดลง เมื่อผ่านออกจากไซแนปส์ กระแสประสาททั้งหมดจะถูกรวมเข้าด้วยกัน หากผลรวมของกระแสประสาทมีค่ามากกว่าหรือเท่ากับค่า Threshold ของนิวรอนลำดับถัดไป กระแสประสาทดังกล่าวก็จะสามารถไหลผ่านเข้าไปในนิวรอนนั้นได้ แต่หากผลรวมของกระแสประสาทนั้นมีค่าน้อยกว่าค่า Threshold กระแสประสาทก็จะถูกยับยั้งไม่ให้ผ่านเข้าไปในนิวรอนถัดไปได้



รูปที่ 3.1 โครงสร้างของนิวรอน

3.3 โครงสร้างของโครงข่ายประสาทเทียม [21-22]

จากที่กล่าวในหัวข้อที่ผ่านมาเป็นระบบการทำงานของนิวรอนในระบบประสาทของมนุษย์ ซึ่งนิวรอนในโครงข่ายประสาทเทียมก็จะจำลองการทำงานของนิวรอนในระบบประสาทมนุษย์มา เช่นเดียวกัน ซึ่งนิวรอนในโครงข่ายประสาทเทียมนั้นจะมีโครงสร้างการทำงานที่ง่ายกว่านิวรอนในสมองมนุษย์อยู่มาก โดยแต่ละนิวรอนจะมีค่าอินพุตป้อนเข้าคือ ค่าข้อมูล ค่าถ่วงน้ำหนัก และค่าไบแอส ซึ่งรูปที่ 3.2 จะแสดงโครงสร้างการทำงานของนิวรอนจำนวน 1 นิวรอนในโครงข่ายประสาทเทียม



รูปที่ 3.2 โครงสร้างของนิวรอนในโครงข่ายประสาทเทียม [22]

จากรูปที่ 3.2 เราสามารถเขียนค่าเอาต์พุตที่ออกมาจากนิวรอนแทนได้ตามสมการดังนี้

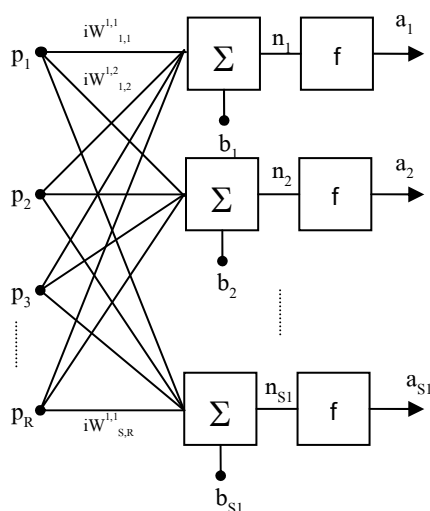
$$a = f(w * p + b) \quad (3.1)$$

เมื่อ	$p_1, p_2, p_3, \dots, p_R$	เป็นค่าอินพุตของหน่วยประสาทเทียม
	$w_{11}, \dots, w_{1R}$	เป็นค่าถ่วงน้ำหนักของข้อมูลเข้าหน่วยประสาทเทียม
	b	เป็นค่าไบแอสของหน่วยประสาทเทียม
	f	เป็นฟังก์ชันการถ่ายโอน (transfer function)
	a	เป็นผลลัพธ์ของหน่วยประสาทเทียม

เมื่อเปรียบเทียบกับระบบประสาทชีวภาพจะเห็นว่าในส่วนที่เป็นชั้นอินพุตเลเยอร์จะเปรียบได้กับเดนไดรต์ ข้อมูลอินพุตคือกระแสประสาทที่เข้าไปยังนิวรอน ค่าถ่วงน้ำหนักจะเหมือนกับไซแนปส์ที่ทำการถ่วงน้ำหนักให้กับค่าอินพุตที่เข้าสู่ นิวรอน ชั้นเอาต์พุตเลเยอร์คือแอกซอน และเอาต์พุตที่ออกมาจากนิวรอนคือกระแสประสาทที่ออกมาจากนิวรอน

3.3.1 โครงข่ายประสาทเทียมแบบชั้นเดียว (Single Layer Neural Network)

เมื่อนำเอานิวรอนมาต่อหลายๆ นิวรอนเข้าด้วยกันให้เกิดเป็นลักษณะของโครงข่ายเป็นชั้นๆ หรือที่เรียกว่าเลเยอร์ (Layer) ก็จะสามารถสร้างโครงข่ายประสาทเทียมขึ้นมาได้ โดยนิวรอนที่อยู่ในเลเยอร์เดียวกันจะไม่มี การเชื่อมต่อถึงกัน โดยโครงข่ายประสาทเทียมแบบชั้นเดียวจะประกอบไปด้วยอินพุตเลเยอร์ และเอาต์พุตเลเยอร์ ซึ่งอินพุตเลเยอร์นั้นเราจะไม่นับเป็น 1 ชั้น เนื่องจากเป็นเลเยอร์ที่ไม่มีการประมวลผลใด ๆ โครงข่ายประสาทเทียมแบบชั้นเดียวนี้ จะมีความสามารถในการคำนวณที่ค่อนข้างน้อย แต่เวลาที่ใช้ในการคำนวณนั้นจะเร็ว รูปของโครงข่ายประสาทเทียมแบบชั้นเดียวนี้จะแสดงได้ตามรูปที่ 3.3

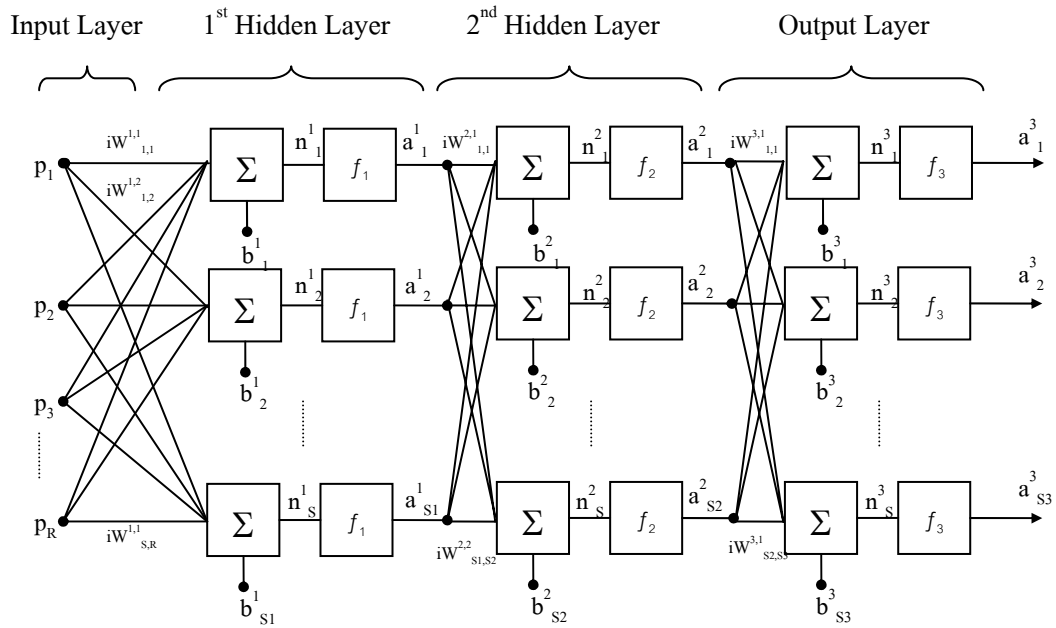


รูปที่ 3.3 โครงสร้างของโครงข่ายประสาทเทียมแบบชั้นเดียว [22]

3.3.2 โครงข่ายประสาทเทียมแบบหลายชั้น (Multi Layer Neural Network)

โครงข่ายประสาทเทียมแบบนี้จะมีชั้นที่เพิ่มขึ้นมาจากโครงข่ายประสาทเทียมแบบชั้นเดียวที่มีแค่อินพุตเลเยอร์ เอาต์พุตเลเยอร์ นั่นคือจะมีชั้นซ่อน (Hidden Layer) เพิ่มขึ้น โดยชั้นซ่อนนี้อาจจะมีได้ตั้งแต่ 1 ชั้นขึ้นไป ขึ้นอยู่กับความซับซ้อนของปัญหาที่นำมาให้โครงข่ายประสาทเทียมทำการวิเคราะห์ แต่โดยทั่วไปแล้วชั้นซ่อนนี้มักจะมีไม่เกิน 2 ชั้น เนื่องจากจะทำให้เวลาที่ใช้ในการคำนวณนานขึ้น เพราะถ้าจำนวนชั้นซ่อนมาก เวลาที่ใช้ในการคำนวณก็จะมากตามไปด้วย

รูปที่ 3.4 จะแสดงโครงสร้างของโครงข่ายประสาทเทียมที่มี 2 ชั้นซ่อน โดยโครงข่ายประสาทเทียมจะประกอบไปด้วย 4 เลเยอร์ คือ อินพุตเลเยอร์ เอาต์พุตเลเยอร์ และชั้นซ่อนจำนวน 2 เลเยอร์



รูปที่ 3.4 โครงสร้างของโครงข่ายประสาทเทียมที่มี 2 ชั้นซ่อน [22]

จากรูปจะแสดงสมการได้คือ

$$a^1 = f^1(p * IW^{1,1} + b^1) \quad (3.2)$$

$$a^2 = f^2(a^1 * LW^{2,1} + b^2) \quad (3.3)$$

$$a^3 = f^3(a^2 * LW^{3,2} + b^3) \quad (3.4)$$

$$a^3 = f^3(f^2(f^1(p * IW^{1,1} + b^1) * LW^{2,1} + b^2) * LW^{3,2} + b^3) \quad (3.5)$$

โครงข่ายประสาทเทียมแบบหลายชั้นนี้จะมีสามารถในการคำนวณที่มากกว่าโครงข่ายประสาทเทียมแบบชั้นเดียว แต่ว่าเวลาในการคำนวณนั้นก็จะใช้เวลานานกว่าโครงข่ายประสาทเทียมแบบชั้นเดียว

3.4 ชนิดของโครงข่ายประสาทเทียม [22-23]

โครงข่ายประสาทเทียมสามารถแบ่งออกได้เป็น 2 ประเภทหลักๆ ตามลักษณะการเรียนรู้ของโครงข่ายประสาทเทียม ได้แก่

1) โครงข่ายประสาทเทียมที่มีการเรียนรู้แบบ Supervised Learning ตัวอย่างของโครงข่ายประสาทเทียมชนิดนี้ เช่น

- Feed Forward Neural Networks
- Radial Basis Neural Networks

2) โครงข่ายประสาทเทียมที่มีการเรียนรู้แบบ Unsupervised Learning ตัวอย่างของโครงข่ายประสาทเทียมชนิดนี้ เช่น

- Kohonen Neural Networks
- Hopfield Neural Networks

3.4.1 การเรียนรู้แบบ Supervised Learning

เป็นการเรียนรู้แบบที่มีการตรวจคำตอบเพื่อให้โครงข่ายประสาทเทียมปรับตัว โดยชุดข้อมูลที่ใช้ฝึกสอนโครงข่ายประสาทเทียมจะมีคำตอบไว้คอยตรวจดูว่าโครงข่ายประสาทเทียมให้คำตอบที่ถูกต้องหรือไม่ ถ้าตอบไม่ถูกต้อง ก็จะนำค่าผิดพลาดที่ได้จากความแตกต่างของเอาต์พุตจากโครงข่ายประสาทเทียมและเอาต์พุตของชุดฝึกสอนป้อนกลับไปเข้าไปเพื่อให้คำนวณหาค่าถ่วงน้ำหนักใหม่ โครงข่ายประสาทเทียมก็จะสามารถปรับตัวเองเพื่อให้ได้คำตอบที่ดีขึ้นได้

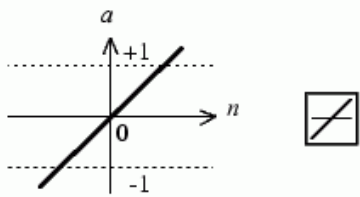
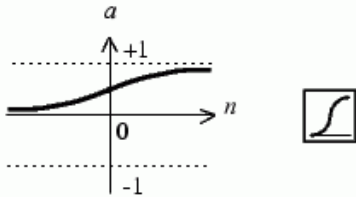
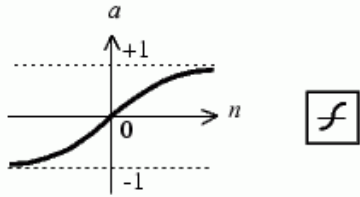
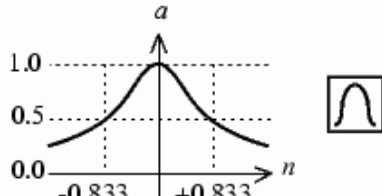
3.4.2 การเรียนรู้แบบ Unsupervised Learning

การเรียนรู้แบบ Unsupervised Neural Network หรือบางครั้งเรียกว่า Self-supervised Learning เป็นการเรียนรู้แบบไม่มีการตรวจคำตอบว่าถูกต้องหรือผิด โครงข่ายประสาทเทียมจะจัดเรียงโครงสร้างด้วยตัวเองตามลักษณะของข้อมูลและผลลัพธ์ที่ได้ การปรับค่าถ่วงน้ำหนักนั้นจะทำการวิเคราะห์ด้วยตัวของมันเองโดยไม่ต้องอาศัยปัจจัยภายนอก โครงข่ายประสาทเทียมจะดูแลแนวโน้มจากสัญญาณของอินพุต และทำการปรับปรุงตามฟังก์ชันของโครงข่าย

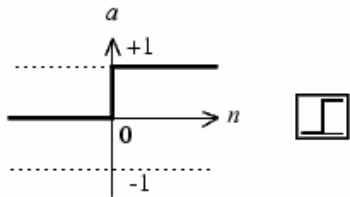
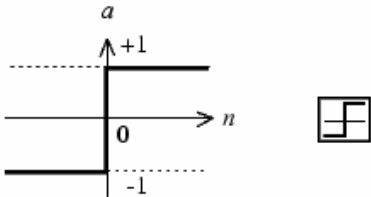
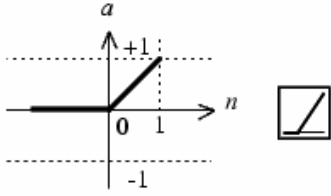
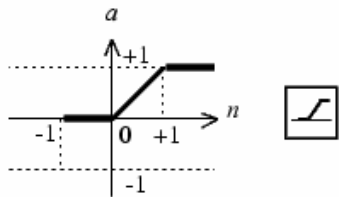
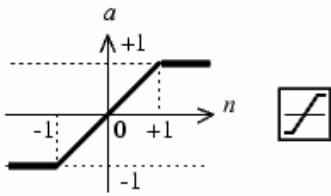
3.5 ฟังก์ชันการถ่ายโอน (Transfer Function)

จากรูปที่ 3.3 จะเห็นได้ว่าความสัมพันธ์ระหว่างอินพุตและเอาต์พุตของนิวรอนจะขึ้นอยู่กับฟังก์ชันการถ่ายโอน (Transfer Function) หรือบางครั้งเรียกว่าฟังก์ชันการกระตุ้น (Activation Function) โดยผลบวกของอินพุตกับค่าถ่วงน้ำหนักจะกลายเป็นอินพุตของฟังก์ชันการถ่ายโอน ซึ่งจะตัวบ่งชี้ค่าเอาต์พุตที่ออกจากนิวรอนแต่ละตัว ฟังก์ชันการถ่ายโอนจะทำหน้าที่เป็นตัวจำกัดขอบเขตของเอาต์พุต ถ้าหากไม่มีฟังก์ชันการถ่ายโอนแล้วก็จะเป็นการยากในการแสดงค่าเอาต์พุตออกมาจากนิวรอน โดยตารางที่ 3.2 จะแสดงถึงฟังก์ชันการถ่ายโอนที่เป็นที่นิยมใช้ในงานโครงข่ายประสาทเทียม

ตารางที่ 3.2 ฟังก์ชันการถ่ายโอนในโครงข่ายประสาทเทียม

ชนิดฟังก์ชันการถ่ายโอน	สมการฟังก์ชันการถ่ายโอน
 $a = \text{purelin}(n)$ Linear Transfer Function	$a(n) = n$
 $a = \text{logsig}(n)$ Log-Sigmoid Transfer Function	$a(n) = \frac{1}{1 + e^{-n}}$
 $a = \text{tansig}(n)$ Tan-Sigmoid Transfer Function	$a(n) = \frac{1 - e^{(-n)}}{1 + e^{(-n)}}$
 $a = \text{radbas}(n)$ Radial Basis Function	$a(n) = e^{-n^2}$
 $a = \text{softmax}(n)$ Compet Transfer Function	$a(n_i) = \begin{cases} 1; & i = i^* \\ 0; & i \neq i^* \end{cases}$

ตารางที่ 3.2 (ต่อ)

ชนิดฟังก์ชันการถ่ายโอน	สมการฟังก์ชันการถ่ายโอน
 $a = \text{hardlim}(n)$ Hard-Limit Transfer Function	$a(n) = \begin{cases} 1; & n \geq 0 \\ 0; & n < 0 \end{cases}$
 $a = \text{hardlims}(n)$ Symmetric Hard-Limit Trans. Funct.	$a(n) = \begin{cases} 1; & n \geq 0 \\ -1; & n < 0 \end{cases}$
 $a = \text{poslin}(n)$ Positive Linear Transfer Funct.	$a(n) = \begin{cases} n; & n \geq 0 \\ 0; & n < 0 \end{cases}$
 $a = \text{satlin}(n)$ Satlin Transfer Function	$a(n) = \begin{cases} 0; & n \leq 0 \\ n; & 0 \leq n \leq 1 \\ 1; & 1 \leq n \end{cases}$
 $a = \text{satlins}(n)$ Satlins Transfer Function	$a(n) = \begin{cases} -1; & n \leq -1 \\ n; & -1 \leq n \leq 1 \\ 1; & 1 \leq n \end{cases}$

3.6 การเลือกชนิดของโครงข่ายประสาทเทียม [24]

เนื่องจากปัจจุบันมีโครงข่ายประสาทเทียมหลายแบบมาก และปัจจุบันก็ยังไม่มีความรู้ในการเลือกชนิดของโครงข่ายประสาทเทียม ดังนั้นการเลือกใช้งานโครงข่ายประสาทเทียมให้เหมาะสมกับปัญหาที่เป็นไปได้ค่อนข้างยาก ตารางที่ 3.3 จะเป็นแนวทางแนะนำว่าควรใช้งานโครงข่ายประสาทเทียมสำหรับงานประเภทไหน

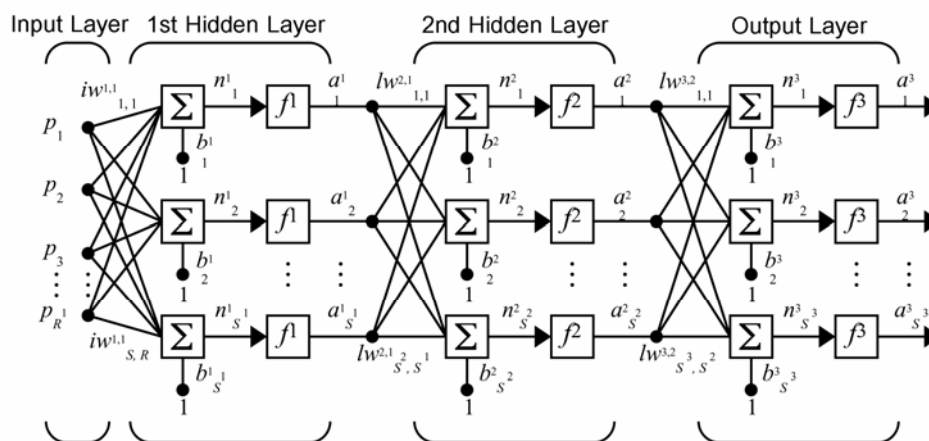
ตารางที่ 3.3 แนวทางการเลือกโครงข่ายแต่ละประเภท [24]

ประเภทของงาน	โครงข่าย
Prediction	- Feedforward Back-propagation - Delta Bar Delta - Extended Delta Bar Delta - Directed Random Search - Higher-order Neural Networks - Self Organizing Map into Back-propagation
Classification	- Learning Vector Quantization - Counter-propagation Network - Probabilistic Neural Network
Data association	- Hopfield Network - Boltzmann Machine - Hamming Network - Bidirectional Associative Memory - Spatio-Temporal Pattern Recognition
Data conceptualization	- Adaptive Resonance Network - Self Organizing Map
Data filtering	- Recirculation

ตารางที่ 3.3 จะเป็นเพียงแค่แนวทางในการเลือกชนิดของโครงข่ายประสาทเทียมสำหรับงานประเภทต่างๆ เท่านั้น ซึ่งในการปฏิบัติงานจริงอาจจะไม่เป็นไปตามตารางดังกล่าวก็ได้ อย่างเช่นโครงข่ายประสาทเทียมแบบ Feed forward Back-propagation จะเป็นโครงข่ายประสาทเทียมที่เป็นที่นิยมและสามารถใช้ได้ดีในงานหลายประเภทนอกเหนือจากการทำนายค่า

3.7 โครงข่ายประสาทเทียมแบบ Back-propagation Neural Network [25]

โครงข่ายประสาทเทียมแบบ Back-propagation Neural Network นั้นมีโครงสร้างเป็นแบบโครงข่ายประสาทเทียมหลายชั้น (Multi Layer Neural Networks) ซึ่งประกอบไปด้วยชั้นอินพุต (Input Layer) , ชั้นซ่อน (Hidden Layer) และชั้นเอาต์พุต (Output Layer) ฟังก์ชันการถ่ายโอนที่ใช้ในโครงข่ายประสาทเทียมแบบแพร่ค่าย้อนกลับนี้จะใช้ฟังก์ชันการถ่ายโอนแบบ Linear Transfer Function , Log – Sigmoid Transfer Function และ Tan – Sigmoid Transfer Function ในวิทยานิพนธ์ฉบับนี้ได้ใช้โครงข่ายประสาทเทียมแบบแพร่ค่าย้อนกลับที่มีชั้นซ่อน 1 ชั้นและโครงข่ายประสาทเทียมแบบแพร่ค่าย้อนกลับที่มีชั้นซ่อน 2 ชั้น



รูปที่ 3.5 โครงสร้างของโครงข่ายประสาทเทียมแบบ Back-Propagation Neural Networks ที่มีชั้นซ่อน 2 ชั้น

จากรูปที่ 3.5 จะแสดงโครงข่ายประสาทเทียมที่มีชั้นซ่อน 2 ชั้น โดยในรูปที่ 3.5 ในชั้นอินพุตจะมีจำนวนอินพุตทั้งหมด R ชุด จำนวนนิวรอนในชั้นซ่อนที่ 1 และ 2 เท่ากับ S^1 และ S^2 ชุดตามลำดับ และมีเอาต์พุต S^3 ชุด สมาชิกของอินพุตเวกเตอร์ P แต่ละตัวจะถูกเชื่อมเข้ากับนิวรอนในชั้นซ่อน ด้วยค่าถ่วงน้ำหนัก w^1 (Input Weight : $iw^{1,1}$) จากนั้นก็จะทำการรวมกับค่าไบอัส (Bias : b^1) แล้วจึงนำไปผ่านฟังก์ชันการถ่ายโอนเพื่อแปลงข้อมูลให้เป็นผลลัพธ์ ซึ่งจะได้เวกเตอร์ a^1 ออกมา โดยเวกเตอร์ a^1 เป็นผลลัพธ์ของชั้นซ่อนที่ 1 ก็จะไปเป็นอินพุตให้กับชั้นซ่อนที่ 2 โดยเวกเตอร์ a^1 ก็จะถูกเชื่อมต่อเข้ากับค่าถ่วงน้ำหนัก w^2 (Layer Weight : $lw^{2,1}$) และนำไปรวมกับค่าไบอัส (Bias : b^2) แล้วนำไปผ่านฟังก์ชันการถ่ายโอนเพื่อแปลงข้อมูลให้เป็นผลลัพธ์ ซึ่งจะได้เวกเตอร์ a^2 ออกมา โดยเวกเตอร์ a^2 ก็จะไปเป็นอินพุตให้กับชั้นเอาต์พุต ซึ่งจะถูกเชื่อมต่อเข้ากับค่าถ่วงน้ำหนัก w^3 (Layer Weight : $lw^{3,1}$) จากนั้นก็จะทำการรวมกับค่าไบอัส (Bias : b^3) แล้ว

นำไปผ่านฟังก์ชันการถ่ายโอนเพื่อหาผลลัพธ์ของโครงข่ายประสาทเทียม โดยผลลัพธ์สุดท้ายก็คือเวกเตอร์ a^3

โดยเอาที่พู่ทของโครงข่ายประสาทเทียมแบบ Back-propagation Neural Network สามารถหาได้จากสมการ

$$a^3 = f^3(f^2(f^1(p * IW^{1,1} + b^1) * LW^{2,1} + b^2) * LW^{3,2} + b^3) \quad (3.6)$$

โดยขั้นตอนการปรับค่าถ่วงน้ำหนักและไบอัสของโครงข่ายประสาทเทียมแบบ Back-propagation Neural Network นั้นมีวิธีการอยู่หลายวิธีด้วยกันแต่วิธีที่นิยมใช้สำหรับการปรับค่าถ่วงน้ำหนักและค่าไบอัสนั้นได้แก่ Lavenberg-Marquardt Algorithm เนื่องจากมีขั้นตอนการปรับค่าถ่วงน้ำหนักและไบอัสที่มีประสิทธิภาพดีที่สุด โดยมีขั้นตอนการปรับค่าถ่วงน้ำหนักและค่าไบอัสดังนี้

$$X_{k+1} = X_k + \Delta X = X_k + J^T \cdot J + \mu I^{-1} \cdot g \quad (3.7)$$

$$g = J^T \cdot e \quad (3.8)$$

โดยที่ x คือค่า Weight และ Bias

J คือ Jacobian matrix ของอนุพันธ์ของค่าผิดพลาดแต่ละค่า ต่อค่า X

J^T คือ Transpose matrix ของ J

g คือ Gradient Factor

e คือ เมตริกซ์ของค่าผิดพลาด (Error) ระหว่างคำตอบกับค่าที่โครงข่ายประสาทเทียมคำนวณได้

I คือ Identity matrix

μ คือ Mu factor หรือ Learning factor

จาโคเบียนเมตริกซ์ของค่าถ่วงน้ำหนัก (w) แสดงได้ตามสมการที่ 3.9

$$J(w) = \begin{bmatrix} \frac{\partial e_1}{\partial w_{1,1}} & \frac{\partial e_1}{\partial w_{1,2}} & \dots & \frac{\partial e_1}{\partial w_{1,s^2}} \\ \frac{\partial e_2}{\partial w_{2,1}} & \frac{\partial e_2}{\partial w_{2,2}} & \dots & \frac{\partial e_2}{\partial w_{2,s^2}} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial e_{s^3}}{\partial w_{s^3,1}} & \frac{\partial e_{s^3}}{\partial w_{s^3,2}} & \dots & \frac{\partial e_{s^3}}{\partial w_{s^3,s^2}} \end{bmatrix} \quad (3.9)$$

จาโคเบียนเมตริกซ์ของค่าไบแอส (b) แสดงได้ตามสมการที่ 3.10

$$J(b) = \begin{bmatrix} \frac{\partial e_1}{\partial b_1} \\ \frac{\partial e_2}{\partial b_2} \\ \vdots \\ \frac{\partial e_{s3}}{\partial b_{s3}} \end{bmatrix} \quad (3.10)$$

เมตริกซ์ของค่าผิดพลาดแสดงได้ตามสมการที่ 3.11

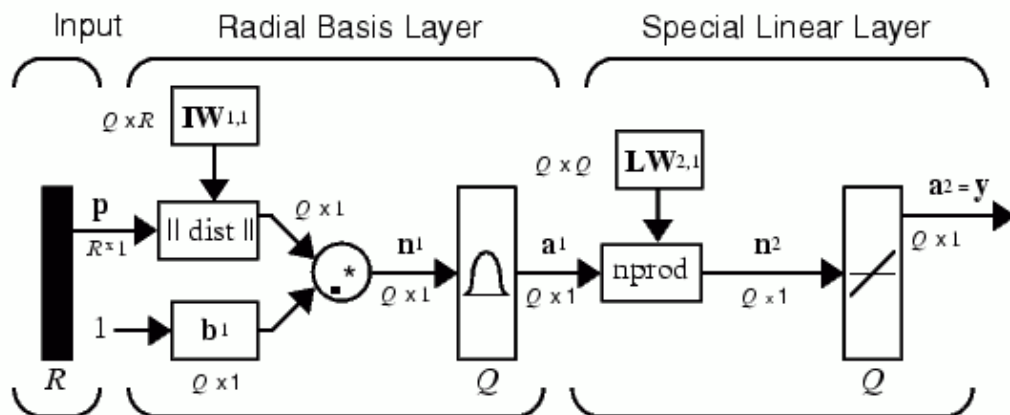
$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_{s3} \end{bmatrix} = \begin{bmatrix} t_1^3 - a_1^3 \\ t_2^3 - a_2^3 \\ \vdots \\ t_{s3}^3 - a_{s3}^3 \end{bmatrix} \quad (3.11)$$

โครงข่ายประสาทเทียมแบบ Back-propagation นั้นจะมีจำนวนนิวรอนในชั้นอินพุตและชั้นเอาต์พุตเท่ากับจำนวนอินพุตและเอาต์พุตของปัญหาที่กำลังศึกษา ส่วนจำนวนนิวรอนในชั้นซ่อนนั้นสามารถปรับเปลี่ยนได้ตามความเหมาะสมและไม่มีจำนวนที่เป็นสูตรสำเร็จ ดังนั้นในการฝึกสอนของโครงข่ายประสาทเทียมจึงต้องมีการปรับเปลี่ยนจำนวนนิวรอนในชั้นซ่อนเพื่อหาโครงสร้างที่เหมาะสม ซึ่งจำนวนนิวรอนในชั้นซ่อนนั้นมีผลต่อระยะเวลาในการฝึกสอนด้วย นั่นคือถ้ามีจำนวนนิวรอนในชั้นซ่อนมาก เวลาที่ใช้ในการฝึกสอนก็จะมากกว่าโครงข่ายประสาทเทียมที่มีจำนวนนิวรอนในชั้นซ่อนน้อยกว่า เมื่อฝึกสอนด้วยเงื่อนไขเดียวกัน แต่อย่างไรก็ตามจำนวนนิวรอนในชั้นซ่อนที่มากก็ไม่ได้เป็นตัวบ่งบอกว่าโครงข่ายจะมีประสิทธิภาพมากกว่าโครงข่ายประสาทเทียมที่จำนวนนิวรอนในชั้นซ่อนที่น้อยกว่า ซึ่งต้องทำการลองปรับเปลี่ยนดูเพื่อหาโครงข่ายที่เหมาะสมที่สุด

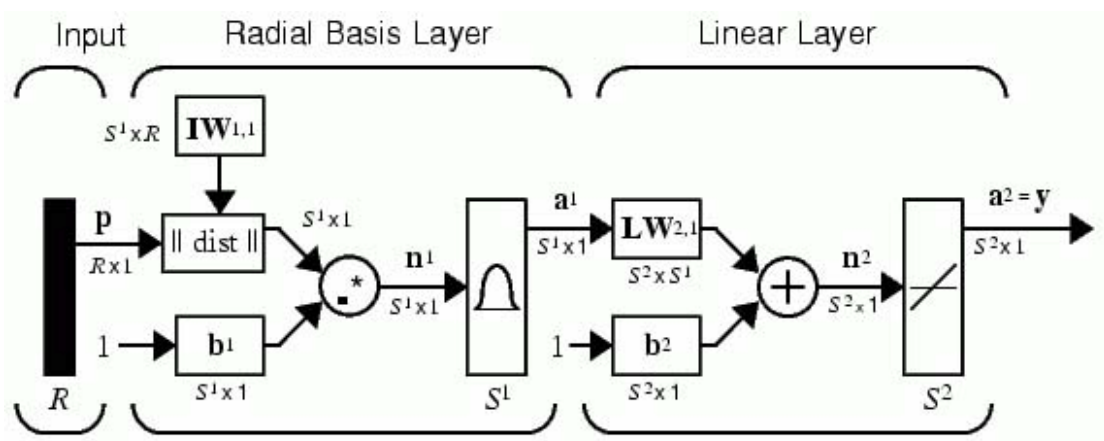
3.8 โครงข่ายประสาทเทียมแบบ Generalized Regression Neural Network [25]

โครงสร้างพื้นฐานของโครงข่ายประสาทเทียมแบบ Generalized Regression Neural Network (GRNN) นั้นจะประกอบไปด้วยชั้นของนิวรอนทั้งหมด 3 ชั้น คือ ชั้นอินพุต (Input Layer), ชั้นซ่อนหรือบางครั้งเรียกว่าชั้นฐานรัศมี (Radial Basis Layer) และชั้นเอาต์พุตหรือบางครั้งเรียกว่าชั้นเชิงเส้นพิเศษ (Special Linear Layer)

โครงข่ายประสาทเทียมแบบ Generalized Regression นั้นเป็นโครงข่ายประสาทเทียมที่มีลักษณะคล้ายกับโครงข่ายประสาทเทียมแบบฟังก์ชันฐานรัศมี (Radial Basis Function Neural Networks, RBF) ซึ่งลักษณะโครงสร้างของโครงข่ายประสาทเทียมทั้งสองแบบนี้จะมีลักษณะโครงสร้างที่คล้ายกันมาก ดังรูปที่ 3.6 และรูปที่ 3.7 โดยในรูปที่ 3.6 จะเป็นโครงสร้างของโครงข่ายประสาทเทียมแบบ Generalized Regression และรูปที่ 3.7 เป็นโครงสร้างของโครงข่ายประสาทเทียมแบบฟังก์ชันฐานรัศมี



รูปที่ 3.6 โครงสร้างของโครงข่ายประสาทเทียมแบบ Generalized Regression



รูปที่ 3.7 โครงสร้างของโครงข่ายประสาทเทียมแบบฟังก์ชันฐานรัศมี

จากรูปที่ 3.6 และ 3.7 นั้นจะเห็นว่าในชั้นซ่อนหรือชั้น Radial Basis Layer นั้นจะมีลักษณะโครงสร้างที่เหมือนกัน นั่นคือมีทั้งค่าถ่วงน้ำหนักและค่าไบอัส แต่ว่าในชั้นเอาต์พุตนั้นจะมีจุดที่แตกต่างกันระหว่างโครงสร้างของโครงข่ายประสาทเทียมทั้ง 2 ชนิดนั่นคือ ในโครงข่ายประสาทเทียมแบบฟังก์ชันฐานรัศมีนั้นจะมีทั้งค่าถ่วงน้ำหนักและค่าไบอัส ในขณะที่โครงข่าย

ประสาทเทียมแบบ Generalized Regression นั้นจะมีเพียงค่าถ่วงน้ำหนักเพียงค่าเดียว และการหาค่า n^2 ของทั้งสองโครงข่ายประสาทเทียมนั้นก็ใช้สมการที่ใช้ในการหาค่า n^2 ที่ต่างกันด้วยเช่นกัน นั่นคือ

โครงข่ายประสาทเทียมแบบ Generalized Regression นั้นจะหาค่า n^2 จากสมการ

$$n^2 = (LW^{2,1} * a^1) / \text{sum}(a^1) \quad (3.12)$$

สำหรับโครงข่ายประสาทเทียมแบบฟังก์ชันฐานรัศมีจะหาค่า n^2 จากสมการ

$$n^2 = b^2 + (LW^{2,1} * a^1) \quad (3.13)$$

ฟังก์ชันการถ่ายโอนที่ใช้ในโครงสร้างของโครงข่ายประสาทเทียมแบบ Generalized Regression นั้นจะใช้ฟังก์ชันการถ่ายโอนชนิด Radial Basis Function ในชั้นฐานรัศมี และในชั้นเอาต์พุตนั้นจะใช้ฟังก์ชันการถ่ายโอนชนิด Linear

จากรูปที่ 3.6 โครงสร้างของโครงข่ายประสาทเทียมแบบ Generalized Regression จะมีอินพุต R ชุด , นิวรอนในชั้นฐานรัศมี Q ชุด และมีนิวรอนในชั้นเอาต์พุต Q ชุด สมาชิกของอินพุตเวกเตอร์ P แต่ละตัวจะถูกเชื่อมเข้ากับนิวรอนแต่ละนิวรอนในชั้นฐานรัศมีด้วยค่าถ่วงน้ำหนัก (Input weight , $IW^{1,1}$) ซึ่งเป็นเมตริกซ์ขนาด $Q \times R$ และเวกเตอร์ a^1 จะเป็นผลลัพธ์ของชั้นฐานรัศมี โดยมีขนาด Q ก็จะไปเป็นอินพุตของชั้นเอาต์พุต สมาชิกแต่ละตัวของเวกเตอร์ a^1 ก็จะถูกเชื่อมเข้ากับนิวรอนแต่ละนิวรอนในชั้นเอาต์พุตด้วยค่าถ่วงน้ำหนัก (Layer weight , $LW^{2,1}$) ซึ่งเป็นเมตริกซ์ขนาด $Q \times Q$ และเวกเตอร์ a^2 จะเป็นผลลัพธ์ของโครงข่ายประสาทเทียมจะมีขนาด Q

3.9 การประยุกต์ใช้งานโครงข่ายประสาทเทียมกับระบบไฟฟ้ากำลัง [26-27]

ระบบไฟฟ้าในปัจจุบันเป็นระบบที่มีความต้องการคุณภาพของไฟฟ้าเพิ่มสูงขึ้น และผู้ใช้ไฟในปัจจุบันก็ต้องการระบบไฟฟ้าที่สามารถจ่ายพลังงานไฟฟ้าได้อย่างมั่นคงและมีประสิทธิภาพเพิ่มมากขึ้นเช่นกัน ดังนั้นระบบไฟฟ้าในปัจจุบันจึงมีการพัฒนาในด้านต่างๆ เพิ่มมากขึ้น แต่ในขณะเดียวกันก็ทำให้ระบบมีความซับซ้อนเพิ่มมากขึ้นด้วย จึงทำให้การวิเคราะห์ระบบไฟฟ้า การควบคุม การตรวจวัด และการวางแผนในระบบไฟฟ้ากำลังทำได้ยากขึ้น

จากความสามารถของโครงข่ายประสาทเทียมจะช่วยให้วิศวกรด้านไฟฟ้ากำลังสามารถแก้ไขและวิเคราะห์ปัญหาต่างๆ ที่เกิดขึ้นในระบบไฟฟ้าได้อย่างรวดเร็วและมีศักยภาพเพิ่มขึ้น โครงข่ายประสาทเทียมที่ถูกฝึกสอนมาอย่างดีจะสามารถจดจำรูปแบบความสัมพันธ์ของข้อมูลต่างๆ ได้ดี สามารถทำนายค่าต่างได้ในเวลาที่รวดเร็วและมีประสิทธิภาพ

ด้วยความสามารถของโครงข่ายประสาทเทียมทำให้นักวิจัยพยายามที่จะประยุกต์โครงข่ายประสาทเทียมเข้าใช้กับงานด้านระบบไฟฟ้ากำลังในหลายๆ ด้าน รวมถึงงานวิศวกรรมด้านอื่นด้วย ซึ่งโครงข่ายที่นิยมใช้สำหรับงานด้านวิศวกรรมนั้นได้แก่ โครงข่ายประสาทเทียมแบบ Back-propagation โดยตารางที่ 3.4 แสดงตัวอย่างในงานวิศวกรรมไฟฟ้าด้านต่างๆ ที่มีการนำโครงข่ายประสาทเทียมไปประยุกต์ใช้งาน

ตารางที่ 3.4 ตัวอย่างการประยุกต์ใช้โครงข่ายประสาทเทียมในงานวิศวกรรมไฟฟ้า [26]

ประเภทของงาน	สาขาของงาน
Prediction & Control	- Transient stability - Steady-state stability - Load Forecasting - Synchronous machine modeling - Dynamic load modeling - Reactive power control
Identification	- Harmonic source and load - Fault diagnosis - Protection
Classification	- Contingency analysis - Alarm processing
Optimization	- Network topological observability - Unit commitment - Economic dispatch - Maintenance scheduling - Optimal power flow - Generation expansion