บทที่ 3

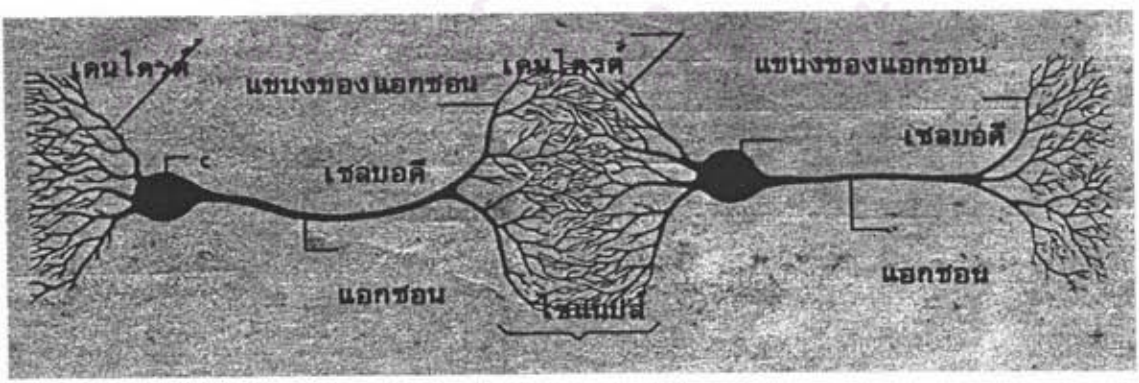
เครือข่ายประสาทเทียม

3.1 ที่มาของเครือข่ายประสาทเทียม [13-16]

เครือข่ายประสาทเทียม (Artificial Neural Network หรือ ANN) คือ เครือข่ายที่ประดิษฐ์ขึ้นเพื่อเลียนแบบพฤติกรรมการทำงานของสมองมนุษย์ ทั้งนี้เนื่องจากลักษณะอันโดดเด่นของสมองมนุษย์ซึ่งเป็นที่ต้องการในศาสตร์แห่งปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) กล่าวคือ

- มีความคงทนและทนทานต่อความผิดพร่อง เช่น เซลล์ประสาทสมองมนุษย์ตายอยู่ทุกวัน โดยไม่ส่งผลกระทบต่อสมรรถนะการทำงานภายในสมองมนุษย์แต่อย่างใด
- มีความสามารถในการจัดการความคลุมเครือ, ไม่แน่นอน, โอกาสที่เป็นไปได้, การรบกวนจากสภาพแวดล้อม หรือสิ่งที่ขัดแย้งกัน
- มีลักษณะการทำงานแบบขนานกันได้สูง
- มีขนาดเล็ก, กะทัดรัด และใช้พลังงานน้อย

เครือข่ายประสาทเทียมประกอบด้วยหน่วยจัดการ (Processing Element) ที่เชื่อมโยงระหว่างกันจำนวนมาก ซึ่งมีชื่อเรียกว่า ปม (Node) โดยทั่วไป ปมเหล่านี้จะทำงานไปพร้อมๆ กัน พฤติกรรมสะสมของเครือข่ายประสาทเทียมคล้ายกับสมองมนุษย์โดยแสดงออกด้วยความสามารถในการเรียนรู้ การฟื้นความทรงจำ และการจัดการข้อมูลที่ไม่เคยพบเห็นมาก่อนได้ โดยอาศัยรูปแบบการฝึกฝน เครือข่ายประสาทเทียมถูกสร้างขึ้นมาจากแบบจำลองทางชีวภาพของปมประสาท (Neuron) ในสมองมนุษย์ ดังนั้นหน่วยจัดการภายในเครือข่ายประสาทเทียมอาจถูกเรียกชื่อว่า "ปมประสาทเทียม" สมองมนุษย์ประกอบด้วยปมประสาทประมาณ 10^{11} ปมโดยมีชนิดที่แตกต่างกันไป แผนผังแสดงการทำงานของปมประสาททางชีวภาพแบบพื้นฐานดังรูปที่ 3.1

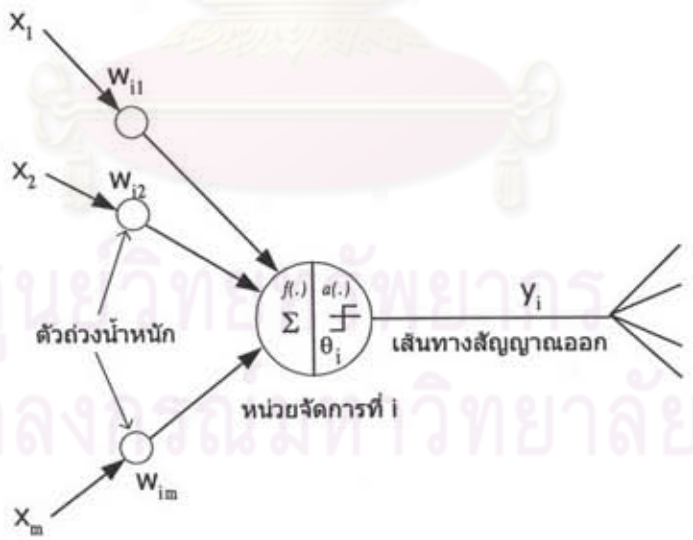


รูปที่ 3.1 แผนผังแสดงการทำงานของปมประสาททางชีวภาพแบบพื้นฐาน

จากรูปที่ 3.1 ปมประสาทปกติมีส่วนประกอบสำคัญ 3 ส่วน คือ ตัวเซลล์ (Cell body) ซึ่งภายในเป็นที่อยู่ของนิวเคลียส เดนไดรต์ (Dendrite) และแอกซอน (Axon)

เดนไดรต์เป็นเครือข่ายคล้ายต้นไม้ที่มีเส้นใยประสาทเชื่อมโยงอยู่กับตัวเซลล์ ส่วนแอกซอนนั้นเป็นทางเชื่อมทรงกระบอกมีลักษณะยาวเป็นเส้นเดียวยาวจากตัวเซลล์ และทำหน้าที่ส่งพัลส์หรือสัญญาณออกจากปมประสาท แต่ละเส้นใยจะไปรวมกันเป็นกระจุกอยู่ส่วนเล็กๆ คล้ายกระเปาะ เรียกว่า ไซแนปส์ (Synapse) ซึ่งปมประสาทจะสร้างสัญญาณไปสู่ปมประสาทอื่นๆ ข้างเคียง ด้านรับจะเป็นคล้ายรอยต่อของปมประสาทข้างเคียงซึ่งพบทั้งที่เป็นเดนไดรต์และตัวเซลล์เอง โดยปกติปมประสาทในสมองมนุษย์เรามีอยู่ประมาณ 10^4 ชุด

เดนไดรต์รับสัญญาณไฟฟ้าอิมพัลส์เมื่อมาถึงไซแนปส์ การส่งผ่านสัญญาณดังกล่าวเกี่ยวข้องกับกระบวนการทางเคมีที่ซับซ้อน ซึ่งในกระบวนการดังกล่าวสารจากตัวส่งจะถูกปล่อยมาจากปลายด้านส่งของรอยต่อ เหตุนี้เองทำให้เพิ่มหรือลดศักย์ไฟฟ้าภายในตัวเซลล์ของเซลล์รับ เซลล์รับจะทำงานหากศักย์ไฟฟ้ามีค่าสูงเกินค่าขีดเริ่ม และพัลส์หรือศักยภาพกระทำของกำลังและเวลาที่คงตัวจะถูกส่งออกไปผ่านทางแอกซอนสู่รอยต่อไซแนปส์เพื่อไปยังปมประสาทอื่นๆ ต่อไป หลังจากเซลล์รับทำงานปมประสาทต้องรอเวลาสักระยะก่อนจะเริ่มทำงานใหม่ได้อีกครั้งเรียกว่า Refractory period ไซแนปส์จะถูกกระตุ้นให้ทำงานในสองลักษณะ กล่าวคือ แบบกระตุ้นให้เซลล์รับทำงาน (Excitatory) และแบบยับยั้งไม่ให้เซลล์รับทำงาน (Inhibitory)



รูปที่ 3.2 แบบจำลองทางคณิตศาสตร์ของปมประสาทแบบ M-P

จากรูปที่ 3.2 แสดงให้เห็นถึงแบบจำลองทางคณิตศาสตร์ของปมประสาทที่เสนอโดย McCulloch และ Pitts จึงเรียกโดยทั่วไปว่า ปมประสาทแบบ M-P ในแบบจำลองนี้ หน่วยจัดการที่

i จะคำนวณผลรวมของสัญญาณเข้าและออก โดย y_i =1 (สั่งให้ทำงาน) หรือ 0 (ไม่สั่งให้ทำงาน) ขึ้นอยู่กับผลรวมของสัญญาณเข้าที่ถูกถ่วงน้ำหนักแล้วว่ามีค่าเกินค่าขีดเริ่ม θ_i

$$y_i(t+1) = a\left(\sum_{j=1}^{m} w_{ij}x_j(t) - \theta_i\right) \qquad (3.1.1)$$

เมื่อฟังก์ชั่นการกระตุ้น a(f) เป็นฟังก์ชั่นหนึ่งหน่วยแบบขั้นบันได

$$a(f) = \begin{cases} 1 & \text{if } f \geq 0 \\ 0 & \text{if } f < 0 \end{cases} \qquad (3.1.2)$$

ตัวถ่วงน้ำหนัก w_{ij} เปรียบเสมือนความแข็งแรงของไซแนปส์ที่เชื่อมต่อระหว่างปมประสาทต้นทาง j ไปยังปมประสาทปลายทาง i ตัวถ่วงน้ำหนักที่เป็นบวกจะสอดคล้องกับไซเนปส์แบบกระตุ้นให้เซลล์ทำงาน ขณะที่ตัวถ่วงน้ำหนักที่เป็นลบจะสอดคล้องกับไซแนปส์แบบยับยั้งไม่ให้เซลล์ทำงาน แต่หากตัวถ่วงน้ำหนักมีค่าเท่ากับศูนย์ แสดงว่าไม่มีการเชื่อมต่อระหว่างปมประสาททั้งสอง

3.2 แบบจำลองพื้นฐานและกฎการเรียนรู้ของเครือข่ายประสาทเทียม [13]

3.2.1 หน่วยจัดการ

หน้าที่การทำงานของปมประสาทแบบ M-P ใน (3.1.1) และ (3.1.2) สามารถขยายแนวคิดไปสู่แบบจำลองทั่วไปของหน่วยจัดการได้ รูปที่ 3.2 กระบวนการจัดการข้อมูลของหน่วยจัดการคือฟังก์ชั่นการรวม f ซึ่งจะรวบรวมข้อมูลจากหน่วยจัดการอื่นๆ มาเป็นสัญญาณเข้าสุทธิของหน่วยจัดการ โดยทั่วไปจะเป็นฟังก์ชั่นเชิงเส้นของสัญญาณเข้า x_i ไปยังหน่วยจัดการ ดังในกรณีของปมประสาทแบบ M-P

$$f_i = \sum_{j=1}^{m} w_{ij}x_j - \theta_i \qquad (3.2.1)$$

เมื่อ θ_i เป็นค่าขีดเริ่มของหน่วยจัดการที่ i การกำหนดฟังก์ชั่นการรวมอาจจะซับซ้อนขึ้นกว่านี้ได้ ดังตัวอย่าง เช่น

- ฟังก์ชั่นกำลังสอง

$$f_i = \sum_{j=1}^{m} w_{ij}x_j^2 - \theta_i \qquad (3.2.2)$$

- ฟังก์ชั่นทรงกลม

$$f_i = \rho^{-2}\sum_{j=1}^{m}\left(x_j - w_{ij}\right)^2 - \theta_i \qquad (3.2.3)$$

เมื่อ ρ และ w_{ij} เป็นรัศมีและจุดศูนย์กลางของทรงกลมตามลำดับ

- ฟังก์ชั่นพหุนาม

$$f_i = \sum_{j=1}^{m}\sum_{k=1}^{m} w_{ijk} x_j x_k + x_j^{\alpha_j} + x_k^{\alpha_k} - \theta_i \qquad (3.2.4)$$

เมื่อ w_{ijk} เป็นตัวถ่วงน้ำหนักของสายเชื่อมโยงระหว่างหน่วยจัดการที่ j และ k ไปยังหน่วยจัดการที่ i และ α_j และ α_k เป็นค่าคงตัวแบบจำนวนจริง (3.2.4) ยังสามารถขยายรูปแบบเป็นพจน์ที่มีอันดับสูงกว่านี้ หน่วยจัดการที่มีฟังก์ชั่นการรวมแบบพหุนาม จะเรียกว่าหน่วยซิกม่า-พาย

การกระทำลำดับที่สองของหน่วยจัดการแต่ละอัน คือ การสร้างสัญญาณออกโดยนำค่าการกระตุ้นที่เป็นฟังก์ชั่นของสัญญาณเข้าเครือข่ายมาเป็นสัญญาณเข้าผ่านไปยังฟังก์ชั่นการกระตุ้นหรือฟังก์ชั่นถ่ายโอน a(f) ซึ่งที่ใช้งานทั่วไป ได้แก่

- ฟังก์ชั่นขั้นบันได

$$a(f) = \begin{cases} 1 & \text{if } f \geq 0 \\ 0 & \text{if } f < 0 \end{cases} \qquad (3.2.5)$$

- ฟังก์ชั่นค่าขีดเริ่ม (ตัวจำกัดแบบฉับพลัน)

$$a(f) = \operatorname{sgn}(f) = \begin{cases} 1 & \text{if } f \geq 0 \\ -1 & \text{if } f < 0 \end{cases} \qquad (3.2.6)$$

- ฟังก์ชั่นแรมป์

$$a(f) = \begin{cases} 1 & \text{if } f \geq 0 \\ f & \text{if } 0 \leq f \leq 1 \\ 0 & \text{if } f < 0 \end{cases} \qquad (3.2.7)$$

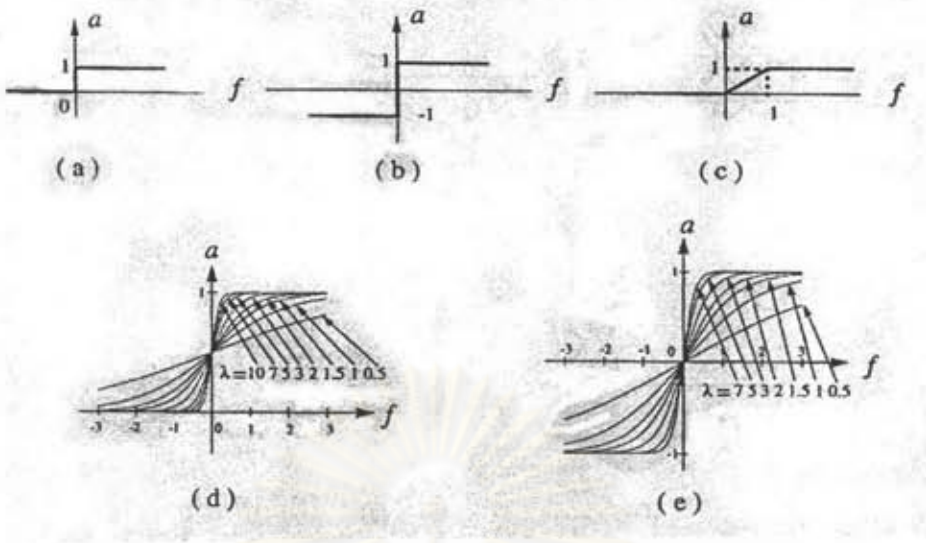
- ฟังก์ชั่นซิกมอยด์แบบขั้วเดียว

$$a(f) = \frac{1}{1+e^{-\lambda f}} \qquad (3.2.8)$$

- ฟังก์ชั่นซิกมอยด์แบบสองขั้ว

$$a(f) = \frac{2}{1+e^{-\lambda f}} - 1 = \tanh(\lambda f) \qquad (3.2.9)$$

เมื่อ $\lambda > 0$ เป็นตัวบ่งบอกถึงความชันของฟังก์ชั่นต่อเนื่อง a(f) ที่ f=0 รูปร่าง ของฟังก์ชั่นดังกล่าวแสดงในรูปที่ 3.3(d) จะสังเกตเห็นว่าเมื่อ λ มีค่ามากถึงค่าอนันต์ (3.2.8) จะลดรูปเป็น (3.2.5) ดังรูปที่ 3.3(a) ขณะที่ในรูปที่ 3.3(e) (3.2.9) จะลดรูปเป็น (3.2.6) ดังรูปที่ 3.3(b)



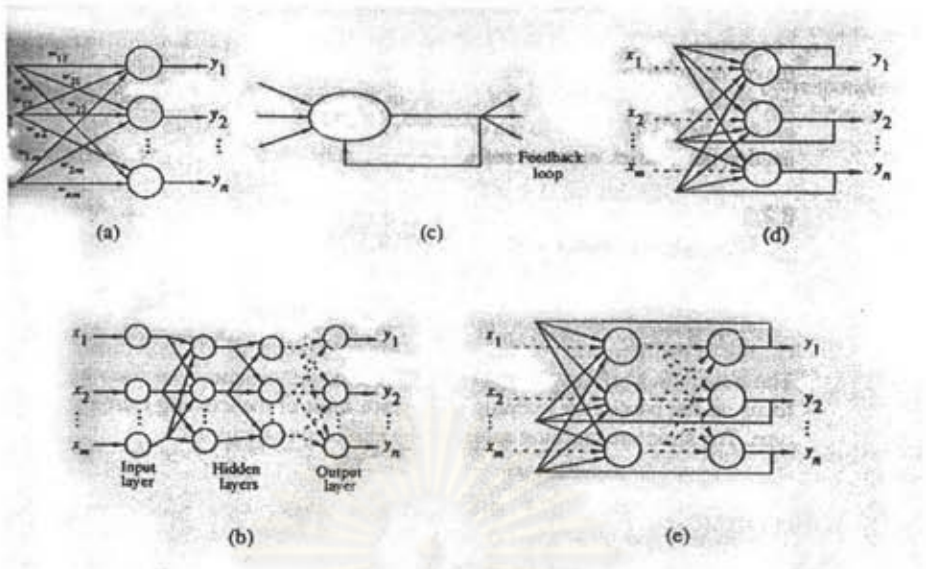
รูปที่ 3.3 ตัวอย่างกราฟฟังก์ชั่นการรวมที่ใช้กันในเครือข่ายประสาทเทียม

หน่วยจัดการใดที่มีฟังก์ชั่นการรวมเชิงเส้นและฟังก์ชั่นการกระตุ้นแบบขีดจำกัดแบบฉับพลันจะเรียกหน่วยจัดการนั้นว่า หน่วยขีดเริ่มเชิงเส้น (Linear Threshold Unit หรือ LTU) และหน่วยจัดการใดมีฟังก์ชั่นการรวมเชิงเส้นและฟังก์ชั่นการกระตุ้นแบบลาดเอียง เรียกหน่วยจัดการนั้นว่า หน่วยลาดเอียงเชิงเส้น (Linear Graded Unit หรือ LGU) ซึ่งหน่วยจัดการทั้งสองดังกล่าวนิยมใช้ในแบบจำลองเครือข่ายประสาทเทียม

ฟังก์ชั่นการรวมแบบไม่เชิงเส้นของสัญญาณเข้าจะสร้างการแปลงรูปแบบของส่วนที่ซับซ้อนของช่วงแสดงคุณลักษณะ เรียกว่า ขอบเขตการตัดสินใจ อย่างไรก็ตาม เนื่องจากฟังก์ชั่นการรวมนี้จะซับซ้อนมากขึ้น และมีความสามารถมากขึ้น ดังนั้นต้องอาศัยการกระทำระหว่างกันในเครือข่ายประสาทเทียมที่ซับซ้อนมากขึ้นด้วย ฟังก์ชั่นการกระตุ้นของหน่วยจัดการอาจแตกต่างไปจากที่ได้ แต่มีผลน้อย สิ่งที่สำคัญที่สุดในการใช้ฟังก์ชั่นการกระตุ้นที่ต้องคำนึงถึง คือ ความราบเรียบของผลตอบของหน่วยจัดการแต่ละหน่วย

3.2.2 การเชื่อมโยง

เครือข่ายประสาทเทียมประกอบด้วยกลุ่มของหน่วยจัดการที่มีการเชื่อมโยงกันสูงซึ่งสัญญาณออกของหน่วยจัดการแต่ละหน่วยต่อเชื่อมกันไปยังหน่วยจัดการอื่นๆ หรือตัวมันเองผ่านทางตัวถ่วงน้ำหนัก ขณะที่การเชื่อมอาจมีความล่าช้าหรือไม่มีก็ได้ ดังนั้นโครงสร้างที่จัดการหน่วยจัดการเหล่านี้และการเชื่อมโยงระหว่างกันควรระบุไว้ในเครือข่ายประสาทเทียมทุกครั้ง การบ่งชี้ว่าการเชื่อมโยงจะเริ่มต้นและสิ้นสุดที่ไหนก็สำคัญ นอกเหนือจากการระบุฟังก์ชั่นที่ใช้ในหน่วยจัดการแต่ละหน่วยในเครือข่ายประสาทเทียม



รูปที่ 3.4 ชนิดพื้นฐานของการเชื่อมโยงของหน่วยจัดการในเครือข่ายประสาทเทียมที่ใช้กัน

รูปที่ 3.4 แสดงชนิดพื้นฐานของการเชื่อมโยงกันของหน่วยจัดการในเครือข่ายประสาทเทียมที่ใช้กัน ซึ่งแบ่งเป็น 5 ชนิดพื้นฐาน

1. เครือข่ายป้อนไปข้างหน้าแบบชั้นเดียว
2. เครือข่ายป้อนไปข้างหน้าแบบหลายชั้น
3. ปมเดียวที่มีการป้อนกลับในตัวเอง
4. เครือข่ายที่แต่ละปมมีการป้อนกลับแบบชั้นเดียว เรียกว่าเครือข่ายรีเคอเรนท์แบบชั้นเดียว
5. เครือข่ายรีเคอเรนท์แบบหลายชั้น

เริ่มแรกพิจารณาหน่วยจัดการเพียงหน่วยเดียว จากนั้นรวมหน่วยจัดการหลายๆ หน่วยเข้าเพื่อสร้างเป็นชั้นของปมประสาท สัญญาณเข้าสามารถถูกเชื่อมโยงไปยังปมประสาทด้วยตัวถ่วงน้ำหนักที่แตกต่างกันไป ส่งผลให้ได้ชุดของสัญญาณออกโดยหนึ่งปมให้หนึ่งสัญญาณออก สิ่งนี้เองทำให้เกิดเครือข่ายป้อนไปข้างหน้าแบบชั้นเดียว ดังรูปที่ 3.4(a) และหากมีการต่อเชื่อมโยงชั้นมากขึ้น จะทำให้เกิดเครือข่ายป้อนไปข้างหน้าแบบหลายชั้นดังรูปที่ 3.4(b) ชั้นที่รับสัญญาณเข้าเรียกว่า ชั้นสัญญาณเข้า (Input layer) โดยปกติจะไม่มีหน้าที่ใดๆ นอกจากการเก็บสัญญาณเข้าไว้ สัญญาณออกของเครือข่ายจะถูกสร้างจากชั้นสัญญาณออก (Output layer) ระหว่างชั้นสัญญาณเข้าและออกถูกเรียกว่า ชั้นซ่อน (Hidden layer) เนื่องจากอยู่ภายในเครือข่ายและไม่มีการติดต่อโดยตรงกับสภาวะแวดล้อมภายนอก โดยในเครือข่ายประสาทเทียมหนึ่งๆ จำนวนชั้นซ่อนอาจไม่มีหรือมีมากมายเลยทีเดียว เครือข่ายประสาทที่สัญญาณออกทุกสัญญาณจากชั้นหนึ่งเชื่อมโยงกับปมประสาททุกปมในชั้นถัดไป จะเรียกเครือข่ายนั้นว่ามีการเชื่อมโยงแบบเต็มที่ จากรูปที่ 3.4(b) เครือข่ายดังกล่าวมีการเชื่อมโยงแบบไม่เต็มที่

เครือข่ายสองชนิดแรกที่กล่าวถึงก่อนหน้านี้เป็นเครือข่ายป้อนไปข้างหน้า เนื่องจากไม่มีสัญญาณออกของหน่วยจัดการใดเป็นสัญญาณเข้าของปมในชั้นเดียวกันหรือชั้นก่อนหน้านั้น แต่เมื่อใดที่สัญญาณออกสามารถถูกส่งไปเป็นสัญญาณเข้าของปมที่อยู่ชั้นเดียวหรือชั้นก่อนหน้านั้นแล้ว จะเรียกเครือข่ายนั้นว่า เครือข่ายป้อนกลับ การป้อนกลับที่มีสัญญาณออกของหน่วยจัดการส่งกลับไปเป็นสัญญาณเข้าของหน่วยจัดการในชั้นเดียวกันเรียก การป้อนกลับแบบข้าง (Lateral feedback) เครือข่ายป้อนกลับที่มีวงรอบปิดเรียก เครือข่ายรีเคอเรนท์ (Recurrent network) ชนิดการเชื่อมโยงที่มีการป้อนกลับแบบข้างที่สำคัญอันหนึ่งคือ โครงสร้างศูนย์กลางหนึ่งกระจายไปโดยรวม (On-center-off-surround) หรือ ยับยั้งแบบข้าง (Lateral-inhibition) ในโครงสร้างดังกล่าวหน่วยจัดการแต่ละหน่วยจะรับสัญญาณเข้า 2 ลักษณะที่แตกต่างกัน กล่าวคือ สัญญาณกระตุ้นให้ทำงาน (Excitatory) จากหน่วยจัดการใกล้เคียงและสัญญาณยับยั้งการทำงาน (Inhibitory) จากหน่วยจัดการที่ไกลออกไป การเชื่อมโยงของโครงสร้างยับยั้งแบบข้างแสดงดังรูปที่ 3.5 ซึ่งการเชื่อมด้วยวงกลมใสเป็นการเชื่อมแบบกระตุ้นให้ทำงาน ขณะที่การเชื่อมด้วยวงกลมทึบเป็นการเชื่อมแบบยับยั้งการทำงาน

รูปที่ 3.5 โครงสร้างศูนย์กลางหนึ่งกระจายไปโดยรวม หรือ ยับยั้งแบบข้าง

3.2.3 กฎการเรียนรู้

องค์ประกอบสำคัญอีกประการหนึ่งในการกำหนดเครือข่ายประสาทเทียม คือ กฎการเรียนรู้ โดยมีการเรียนรู้อยู่ 2 ชนิดหลักๆ ในเครือข่ายประสาทเทียม คือ การเรียนรู้พารามิเตอร์ (Parameter learning) ซึ่งเกี่ยวข้องกับการปรับปรุงค่าตัวถ่วงน้ำหนักที่เชื่อมโยงในเครือข่ายประสาทเทียม และการเรียนรู้โครงสร้าง (Structure learning) ซึ่งเน้นไปที่การเปลี่ยนแปลงโครงสร้างเครือข่ายได้แก่ จำนวนหน่วยจัดการ ชนิดการเชื่อมโยงเป็นต้น การเรียนรู้ทั้งสองชนิดนี้สามารถกระทำไปพร้อมๆ กันหรือแยกจากกันก็ได้ ในที่นี้จะกล่าวถึงเฉพาะการเรียนรู้พารามิเตอร์เท่านั้น เนื่องจากเป็นการเรียนรู้ที่ใช้อยู่ในวิทยานิพนธ์ฉบับนี้

เพื่อความเข้าใจโดยง่าย เราจะสมมติให้มีหน่วยจัดการ n หน่วยในเครือข่ายประสาทเทียม และแต่ละหน่วยที่มีตัวถ่วงน้ำหนักที่ปรับค่าได้อยู่ m ตัว เราจะนิยามเมตริกซ์ตัวถ่วงน้ำหนัก (Weight matrix) หรือเมตริกซ์การเชื่อมโยง (Connection matrix) ดังนี้

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_n^T \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{bmatrix} \qquad (3.2.10)$$

เมื่อ $\mathbf{w}_i = [w_{i1} \quad w_{i2} \quad \cdots \quad w_{im}]^T, i = 1,2,\ldots,n$ เป็นเวคเตอร์ตัวถ่วงน้ำหนัก (Weight vector) ของหน่วยจัดการที่ i และ w_{ij} เป็นตัวถ่วงน้ำหนักของการเชื่อมหน่วยจัดการที่ i (ปมต้นทาง) ไปยังหน่วยจัดการที่ j (ปมหลายทาง)

สมมติเมตริกซ์ตัวถ่วงน้ำหนัก W ประกอบด้วยองค์ประกอบที่ปรับค่าได้ทั้งหมดในเครือข่ายประสาทเทียมแล้ว ชุดของเมตริกซ์ W ที่เป็นไปได้ทั้งหมดหาได้จากชุดรูปแบบการจัดการข้อมูลที่เป็นไปได้ทั้งหมดสำหรับเครือข่ายประสาทนี้ หรือกล่าวอีกอย่างก็คือ ถ้าเครือข่ายประสาทเทียมรู้จักประสิทธิภาพการจัดการข้อมูลที่เราต้องการ เครือข่ายประสาทเทียมนั้นสามารถหาได้โดยหาเมตริกซ์ตัวถ่วงน้ำหนักที่เหมาะสมมาใช้นั้นเอง ในการเชื่อมลักษณะนี้ เราเรียกตัวถ่วงน้ำหนักเข้ารหัสความจำระยะยาว (Long-term memory หรือ LTM) และเรียกสถานะการกระตุ้นของปมประสาทเข้ารหัสความจำระยะสั้น (Short-term memory หรือ STM) ในเครือข่ายประสาทเทียม ดังนั้นสำหรับเครือข่ายประสาทเทียมที่ใช้การเรียนรู้ตัวถ่วงน้ำหนัก เราต้องพัฒนากฎการเรียนรู้เพื่อใช้หาเมตริกซ์ตัวถ่วงน้ำหนักที่มีประสิทธิภาพเพื่อให้ได้สมรรถนะเครือข่ายที่ต้องการ โดยทั่วไปแล้วกฎการเรียนรู้แบ่งออกได้ 3 ประเภทคือ การเรียนรู้แบบควบคุมดูแล (Supervised learning) การเรียนรู้แบบเสริม (Reinforcement learning) และการเรียนรู้แบบไม่ควบคุมดูแล (Unsupervised learning)

ในการเรียนรู้แบบควบคุมดูแลนั้น ที่แต่ละช่วงเวลาเมื่อมีสัญญาณเข้า เครือข่ายประสาทเทียมจะให้ผลตอบที่ต้องการ ดังนั้นเครือข่ายจะถูกสอนให้รู้ว่าสัญญาณออกที่ควรเป็นคืออะไร เพื่อความเข้าใจกระจ่างมากขึ้น ในการเรียนรู้แบบควบคุมดูแลนั้น เครือข่ายประสาทเทียมถูกสอนด้วยชุดข้อมูลการฝึกฝนดังตัวอย่างต่อไปนี้ $(\mathbf{x}_1, d_1), (\mathbf{x}_2, d_2), \ldots, (\mathbf{x}_k, d_k)$ เมื่อ $\mathbf{x}_k$ เป็นสัญญาณเข้าของเครือข่ายประสาทเทียมและ d_k เป็นสัญญาณออกที่ต้องการได้จากเครือข่ายประสาทเทียม ในรูปที่ 3.6(a) จากความแตกต่างระหว่างสัญญาณออกจริง y_k และสัญญาณออกที่ต้องการ d_k เพื่อใช้สัญญาณความผิดพลาดนี้ในการแก้ไขปรับปรุงค่าตัวถ่วงน้ำหนักเพื่อให้สัญญาณออกจริงมีค่าใกล้เคียงกับสัญญาณออกที่ต้องการมากขึ้น

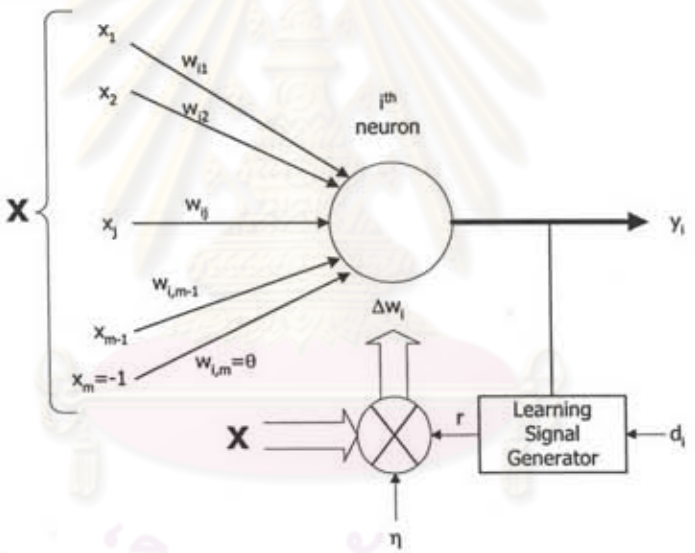
รูปที่ 3.6 กฎการเรียนรู้ของเครือข่ายประสาทเทียม

ในการเรียนรู้แบบควบคุมดูแล สิ่งที่ต้องรู้คือค่าสัญญาณออกที่ต้องการสำหรับแต่ละรูปแบบของสัญญาณเข้า แต่ในบางครั้งรายละเอียดของข้อมูลมีอยู่น้อย ยกตัวอย่างเช่น เครือข่ายประสาทเทียมอาจเพียงแต่ถูกสอนให้เรียนรู้ว่าสัญญาณออกจริงมีค่าความถูกต้องเพียง 50% หรือมากกว่านั้น ในกรณีสุดโต่งมีเพียงบิทของข้อมูลป้อนกลับเพื่อบ่งบอกว่าสัญญาณออกถูกหรือผิดเท่านั้น การเรียนรู้ที่อ้างอิงชนิดข้อมูลวิจารณ์แบบนี้ถูกเรียกว่า การเรียนรู้แบบเสริม และข้อมูลการป้อนกลับในกรณีนี้ถูกเรียกว่า สัญญาณเสริม ดังรูปที่ 3.6(b) การเรียนรู้แบบเสริมเป็นรูปแบบหนึ่งที่คล้ายคลึงกับการเรียนรู้แบบควบคุมดูแล เนื่องจากเครือข่ายยังคงรับการป้อนกลับบางอย่างจากสภาพแวดล้อมของมัน แต่การป้อนกลับเป็นเพียงแต่การประเมินผลหรือวิจารณ์แทนที่จะสอนโดยตรง กล่าวคือมันบอกเพียงแต่ว่าสัญญาณออกที่ได้ดีมากน้อยเพียงใด แต่ไม่ได้ให้คำแนะนำว่าคำตอบที่ถูกต้องควรเป็นอะไร สัญญาณเสริมภายนอกจะถูกประมวลผลด้วยตัวสร้างสัญญาณวิจารณ์เพื่อสร้างสัญญาณวิจารณ์ที่มีข้อมูลเพิ่มเติมสำหรับเครือข่ายประสาทเทียมในการปรับค่าตัวถ่วงน้ำหนักที่เหมาะสมด้วยความหวังที่ว่าจะได้การป้อนกลับวิจารณ์ที่ดีมากขึ้นในอนาคต สรุปได้ว่า การเรียนรู้แบบเสริมอาจเรียกได้ว่าเป็นการเรียนรู้ที่มีการวิจารณ์ ตรงกันข้ามกับการเรียนรู้ที่มีครูบอก ซึ่งอธิบายการเรียนรู้แบบควบคุมดูแล นั่นเอง

ในการเรียนรู้แบบไม่ควบคุมดูแลนั้น จะไม่มีครูคอยสอนให้ข้อมูลป้อนกลับ กล่าวคือ ไม่มีการป้อนกลับจากสภาพแวดล้อมภายนอกเพื่อบอกว่าสัญญาณออกควรเป็นอะไรหรือสัญญาณออกถูกต้องหรือไม่ เครือข่ายต้องค้นหารูปแบบ คุณลักษณะ กฎเกณฑ์ ความสัมพันธ์กัน หรือชนิดในข้อมูลสัญญาณเข้า และเข้ารหัสออกมาเป็นสัญญาณออกเอง เครือข่ายจะต้องกระทำผ่านการเปลี่ยนแปลงค่าพารามิเตอร์ กระบวนการดังนี้ เรียกว่า การจัดระบบตัวเอง (Self-organizing) ตัว

อย่างที่เป็นแบบอย่างก็คือ การจำแนกวัตถุโดยไม่มีการให้ข้อมูลเกี่ยวกับชนิดข้อมูลจริงๆ จัดกลุ่มที่เหมาะสมโดยการค้นหาความคล้ายคลึงและแตกต่างระหว่างวัตถุเหล่านั้นเอง

จากรูปที่ 3.6 อ้างอิงถึงโหมดการเรียนรู้พื้นฐานทั้งสามชนิด ต่อไปเราจะพิจารณาโครงสร้างการฝึกฝนโดยทั่วไปสำหรับหน่วยจัดการที่แสดงในรูปที่ 3.7 ในรูปนี้สัญญาณเข้า $x_j, j=1,2,\ldots,m$ สามารถเป็นสัญญาณออกจากหน่วยจัดการอีกหน่วย หรืออาจเป็นสัญญาณเข้าจากภายนอกเองก็ได้ ข้อสังเกตอยู่ที่ค่าพารามิเตอร์ขีดเริ่ม θ_i อาจรวมอยู่ในการเรียนรู้เสมือนค่าหนึ่งของตัวถ่วงน้ำหนัก โดยคงค่าสัญญาณเข้าหนึ่งสัญญาณที่ –1 ไว้ สัญญาณที่ต้องการ d_i จะมีก็ต่อเมื่ออยู่ในโหมดการเรียนรู้แบบควบคุมดูแล หรือแบบเสริมเท่านั้นโดยที่ d_i เป็นสัญญาณเสริม ดังนั้นโหมดการเรียนรู้ทั้งสอง ตัวถ่วงน้ำหนักของหน่วยจัดการที่ i อาจถูกดัดแปลงให้สอดคล้องกับสัญญาณเข้าที่รับมา ค่าสัญญาณออก ผลตอบการสอนที่เกี่ยวข้อง อย่างไรก็ตามในโหมดการเรียนรู้แบบไม่ควบคุมดูแล หน่วยจัดการจะดัดแปลงตัวถ่วงน้ำหนักโดยอ้างอิงจากสัญญาณเข้าและ/หรือสัญญาณออกเท่านั้น



รูปที่ 3.7 กฎการเรียนรู้ค่าตัวถ่วงน้ำหนักโดยทั่วไป

ในรูปแบบทั่วไปของกฎการเรียนรู้ ตัวถ่วงน้ำหนักในเครือข่ายประสาทเทียมจะบ่งบอกว่าการเพิ่มขึ้นของเวคเตอร์ตัวถ่วงน้ำหนัก w_i ที่สร้างขึ้นแต่ละขั้นการเรียนรู้ที่เวลา t จะเป็นสัดส่วนกับผลคูณสัญญาณการเรียนรู้ r และสัญญาณเข้า x(t) กล่าวคือ

$$\Delta \mathbf{w}_i(t) \propto r\mathbf{x}(t) \text{ หรือ } \Delta \mathbf{w}_i(t) = \eta r \mathbf{x}(t) \qquad (3.2.11)$$

เมื่อ η เป็นค่าเลขจริงบวกเรียกว่า ค่าคงตัวการเรียนรู้ ซึ่งบอกถึงอัตราการเรียนรู้ และ r เป็นสัญญาณการเรียนรู้ซึ่งโดยทั่วไปเป็นฟังก์ชั่นของ w_i, x และสัญญาณที่คอยสอน d_i กล่าวคือ

$$r = f_r(\mathbf{w}_i, \mathbf{x}, d_i) \qquad (3.2.12)$$

ดังนั้นกฎการเรียนรู้ค่าตัวถ่วงน้ำหนักโดยทั่วไปใน (3.2.11) ชี้บอกให้รู้ว่าเวคเตอร์ตัวถ่วงน้ำหนัก $\mathbf{w}_i = [w_{i1} \quad w_{i2} \quad \cdots \quad w_{im}]^T$ จะมีค่าเพิ่มขึ้นเป็นสัดส่วนกับผลคูณของสัญญาณเข้า x และสัญญาณการเรียนรู้ r โดยสามารถปรับปรุงค่าตัวถ่วงน้ำหนักในช่วงการเรียนรู้ถัดไปได้โดยอาศัย (3.2.11) และ (3.2.12)

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \eta f_r\left(\mathbf{w}_i(t), \mathbf{x}(t), d_i(t)\right)\mathbf{x}(t) \qquad (3.2.13)$$

ซึ่งอาจเขียนใหม่ได้ว่า

$$\mathbf{w}_i^{(t+1)} = \mathbf{w}_i^{(t)} + \eta f_r\left(\mathbf{w}_i^{(t)}, \mathbf{x}^{(t)}, d_i^{(t)}\right)\mathbf{x}^{(t)} \qquad (3.2.14)$$

เมื่อตัวห้อยบนแสดงถึงช่วงเวลาการเรียนรู้ (3.2.14) จะเกี่ยวพันกับลำดับการปรับปรุงค่าตัวถ่วงน้ำหนักเป็นช่วงๆ และแบบต่อเนื่องโดยสามารถอธิบายได้ในสมการข้างล่าง

$$\frac{d\mathbf{w}_i(t)}{dt} = \eta r \mathbf{x}(t) \qquad (3.2.15)$$

สำหรับ (3.2.14) และ (3.2.15) เราจะต้องเริ่มค่าตัวถ่วงน้ำหนักอย่างเหมาะสมก่อนเริ่มกระบวนการเรียนรู้ ยกตัวอย่างเช่น การสุ่มค่าตัวถ่วงน้ำหนัก เป็นต้น

โดยการใช้กฎการเรียนรู้ตัวถ่วงน้ำหนักทั่วไปใน (3.2.14) ได้มีการพัฒนากฎการฝึกฝนและเรียนรู้ตัวถ่วงน้ำหนักแบบควบคุมดูแลและไม่ควบคุมดูแลมากมาย ปัจจัยสำคัญที่แตกต่างระหว่างกฎเกณฑ์ทั้งสองชนิดนี้เน้นไปที่จะสร้างสัญญาณการเรียนรู้ได้อย่างไรเพื่อปรับปรุงค่าตัวถ่วงน้ำหนักในเครือข่ายประสาทเทียมให้เหมาะสมนั่นเอง

เริ่มต้นเราจะศึกษากฎการเรียนรู้ของเฮ็บบ์ (Hebb) ในรูปแบบของกฎการเรียนรู้ตัวถ่วงน้ำหนักทั่วไปก่อน เฮ็บบ์ตั้งสมมติฐานว่าเมื่อมีสัญญาณเข้ามาทางแอกซอนจากปมประสาท A ไปยังปมประสาท B แล้วทำให้ปมประสาท B ปลดปล่อยพัลส์ขึ้นมาทันทีทันใด และสถานการณ์ดังกล่าวเกิดขึ้นซ้ำๆ อย่างนี้เรื่อยไป ประสิทธิภาพของสัญญาณจากแอกซอนนั้นในเชิงความสามารถในการช่วยให้ปมประสาท B ปลดปล่อยพัลส์จะมีค่าสูงขึ้น ดังนั้นเขาจึงแนะนำว่า ความแข็งแกร่งของไซแนปส์ในสมองจะแปรผันเป็นสัดส่วนกับความสัมพันธ์กันระหว่างการปลดปล่อยพัลส์ของปมประสาทก่อนผ่านและหลังผ่านไซแนปส์ หลักการนี้เป็นสิ่งบุกเบิกในศาสตร์การคำนวณทางเครือข่ายประสาท กฎการเรียนรู้เครือข่ายประสาทมากมายส่วนใหญ่จะสะท้อนให้เห็นหลักการของเฮ็บบ์

เพื่อให้สอดคล้องกับสมมติฐานของเฮ็บบ์ที่ว่าตัวถ่วงน้ำหนักจะถูกปรับให้สอดคล้องกับความสัมพันธ์กันก่อนผ่านและหลังผ่านไซแนปส์ สัญญาณการเรียนรู้ r ในกฎการเรียนรู้ตัวถ่วงน้ำหนักทั่วไปจะถูกเขียนได้ดังนี้

$$r = a\left(\mathbf{w}_i^T \mathbf{x}\right) = y_i \qquad (3.2.16)$$

เมื่อ a(.) เป็นฟังก์ชั่นการกระตุ้นของหน่วยจัดการ ดังนั้นในกฎการเรียนรู้ของเฮ็บบ์ สัญญาณการเรียนรู้ r จะถูกเขียนในพจน์ของสัญญาณออกปัจจุบันของหน่วยจัดการ เพื่อให้สอดคล้องกับ (3.2.11) การเพิ่มขึ้นของ $\Delta \mathbf{w}_i$ ของเวคเตอร์ตัวถ่วงน้ำหนักจะเป็นดังนี้

$$\Delta \mathbf{w}_i = \eta a\left(\mathbf{w}_i^T \mathbf{x}\right)\mathbf{x} = \eta y_i \mathbf{x} \qquad (3.2.17)$$

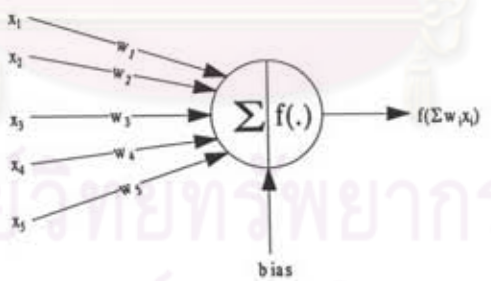
กล่าวคือองค์ประกอบของเวคเตอร์ตัวถ่วงน้ำหนักถูกปรับปรุงค่าด้วยสมการข้างล่าง

$$\Delta w_{ij} = \eta a\left(\mathbf{w}_i^T \mathbf{x}\right)x_j = \eta y_i x_j \qquad (3.2.18)$$

ดังนั้นกฎการเรียนรู้ของเฮ็บบ์จะเป็นการเรียนรู้แบบไม่ควบคุมดูแลสำหรับเครือข่ายป้อนไปข้างหน้า เนื่องจากมันใช้เฉพาะผลคูณสัญญาณเข้าและสัญญาณออกจริงเพื่อปรับค่าตัวถ่วงน้ำหนัก ไม่มีสัญญาณออกที่ต้องการมากำหนดสัญญาณการเรียนรู้เพื่อปรับปรุงค่าตัวถ่วงน้ำหนัก กฎการเรียนรู้นี้ต้องการค่าเริ่มต้นของตัวถ่วงน้ำหนักซึ่งมักเป็นตัวเลขสุ่มใกล้ศูนย์ก่อนเริ่มการเรียนรู้ (3.2.18) บอกให้รู้ว่า ถ้าความสัมพันธ์กันระหว่างสัญญาณเข้าและสัญญาณออกเป็นบวก ค่าตัวถ่วงน้ำหนัก w_i จะเพิ่มขึ้น มิเช่นนั้นจะมีค่าลดลง ยิ่งไปกว่านั้นเนื่องจากสัญญาณออกจะถูกทำให้ดีขึ้นทุกครั้งที่มีสัญญาณเข้า บ่อยครั้งรูปแบบสัญญาณจะมีอิทธิพลต่อค่าตัวถ่วงน้ำหนักและสุดท้ายจะสร้างสัญญาณออกที่มีค่าใหญ่มากสุด ดังนั้นกฎการเรียนรู้ใน (3.2.18) จะจับแนวคิดของสมมติฐานของเฮ็บบ์

3.3 ชนิดของเครือข่ายประสาทเทียม

3.3.1 เครือข่ายประสาทเทียมแบบหน่วยรับรู้ชั้นเดียว



รูปที่ 3.8 แบบจำลองของปมประสาท

หน่วยรับรู้ชั้นเดียวประกอบด้วยกลุ่มหน่วยจัดการ หรือปมประสาท (Neuron) อยู่ในชั้นเดียวกัน กระบวนการทำงานของปมประสาททั่วไปมีสองขั้นตอนซึ่งแสดงไว้ดังรูปที่ 3.8 คือ

1. ขั้นการรวมเชิงเส้น (Linear combination) จะรับเอาสัญญาณขาเข้ามารวมเชิงเส้นตามสมการ (3.3.1)

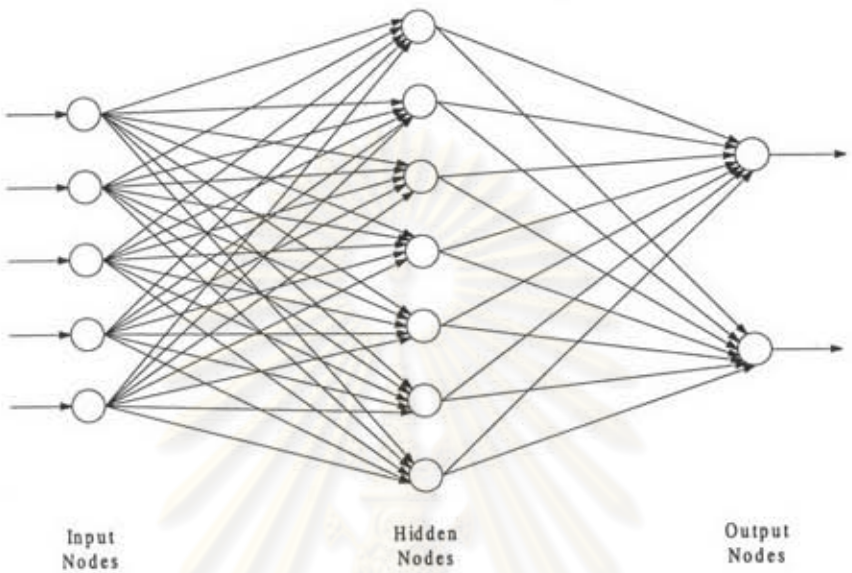
$$f_i = \sum_{j=1}^{m} w_{ij}x_j - \theta_i = \sum_{j=0}^{m} w_{ij}x_j \quad ; \; x_0 = -1 \qquad (3.3.1)$$

I20698240

2. ชั้นฟังก์ชั่นการกระตุ้น (Activation function) ทำหน้าที่รับผลจากการรวมเชิงเส้นออกมาเป็นสัญญาณขาออกของหน่วยรับรู้ชั้นเดียวตาม (3.3.2)

$$y_i = a(f_i) \qquad (3.3.2)$$

3.3.2 เครือข่ายประสาทเทียมแบบหน่วยรับรู้หลายชั้น



รูปที่ 3.9 แบบจำลองของหน่วยรับรู้หลายชั้นอย่างง่าย

หน่วยรับรู้หลายชั้นประกอบด้วยชั้นมากมายซึ่งแต่ละชั้นประกอบด้วยปมประสาทจำนวนหนึ่งเพื่อให้สามารถประมาณความสัมพันธ์แบบไม่เป็นเชิงเส้นใดๆ ได้ หน่วยรับรู้หลายชั้นแบบง่ายมีสามชั้นคือ ชั้นขาเข้า ชั้นซ่อน และชั้นขาออก ดังรูปที่ 3.9

3.3.3 เครือข่ายประสาทเทียมแบบฟังก์ชั่นมูลฐานแนวรัศมี

จากเครือข่ายประสาทเทียมแบบหน่วยรับรู้หลายชั้น พบว่าการเรียนรู้ต้องอาศัยเทคนิคออพติไมซ์แบบไม่เชิงเส้นซึ่งค่อนข้างยุ่งยาก และมีปัญหาในการลู่เข้าคำตอบช้า ทางเลือกหนึ่งที่ใช้ได้ในเครือข่ายที่มีความไม่เป็นเชิงเส้นสูง คือ การสลับตำแหน่งของส่วนที่เป็นเชิงเส้นและส่วนที่ไม่เป็นเชิงเส้นภายในเครือข่ายประสาทเทียม ทำให้ได้เครือข่ายประสาทเทียมแบบใหม่ขึ้นได้แก่ เครือข่ายประสาทเทียมแบบฟังก์ชั่นมูลฐานแนวรัศมีหรือเรียกอีกอย่างหนึ่งว่า เครือข่ายฟังก์ชั่นมูลฐานแนวรัศมี (Radial Basis Function Network หรือ RBF Network) โดยฟังก์ชั่นมูลฐานแนวรัศมีมีคุณสมบัติในการแปลงแบบไม่เป็นเชิงเส้นโดยคงค่าพารามิเตอร์ภายในฟังก์ชั่นมูลฐานแนวรัศมี ขณะที่สัญญาณออกเป็นเพียงการรวมเชิงเส้นกล่าวคือ มีเพียงการปรับค่าพารามิเตอร์หรือค่าตัวถ่วงน้ำหนักของตัวรวมเชิงเส้นเท่านั้น ทำให้ค่าพารามิเตอร์ต่างๆ สำหรับการเรียนรู้ของเครือข่ายประสาทเทียมแบบฟังก์ชั่นมูลฐานแนวรัศมีอาศัยเทคนิคออพติไมซ์แบบเชิงเส้น อย่างไรก็ตาม ประ

สิทธิภาพในการเรียนรู้ของฟังก์ชั่นมูลฐานแนวรัศมีภายในเครือข่ายประสาทเทียมนี้ขึ้นกับค่าพารามิเตอร์ต่างๆ ของฟังก์ชั่นมูลฐานแนวรัศมีแต่ละปมที่เลือกอย่างมาก ซึ่งได้กล่าวถึงวิธีการเลือกค่าพารามิเตอร์ของฟังก์ชั่นมูลฐานแนวรัศมีภายในเครือข่ายประสาทเทียมไว้ในหัวข้อ 4.2

ในเครือข่ายฟังก์ชั่นมูลฐานแนวรัศมีรูปแบบฟังก์ชั่น $\Phi(.)$ และจุดศูนย์กลาง c_i คงตัว เมื่อให้กลุ่มของสัญญาณขาเข้า $\mathbf{x}$ และสัญญาณขาออกที่สอดคล้องกันมา y ค่าตัวถ่วงน้ำหนักสามารถหาได้ด้วยวิธีความผิดพลาดกำลังสองน้อยสุดแบบเชิงเส้น (Linear Least Squares Eror method)

ในทางทฤษฎี การเลือกฟังก์ชั่นไม่เชิงเส้น $\Phi(.)$ ไม่สำคัญต่อสมรรถนะของเครือข่ายฟังก์ชั่นมูลฐานแนวรัศมี ยกตัวอย่าง เช่น ฟังก์ชั่น Thin-plate-spline ตาม (3.3.3)

$$\Phi(v) = v^2 \log(v) \tag{3.3.3}$$

จะเห็นว่า $\lim_{v\to\infty} \Phi(v) \to \infty$

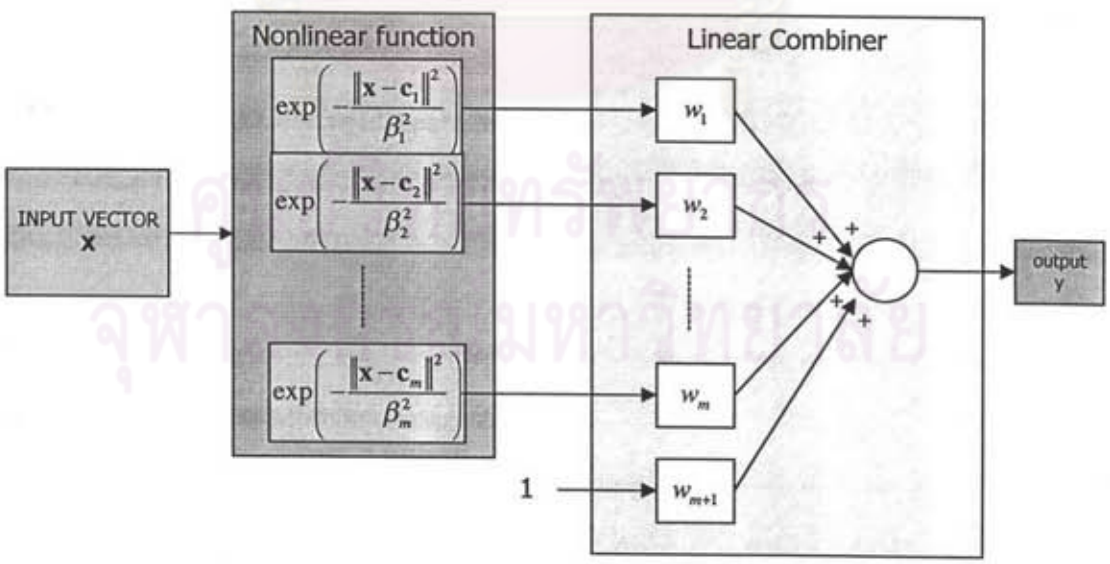
และฟังก์ชั่นเกาส์ (Gaussian function) ตาม (3.3.4)

$$\Phi(v) = \exp\left(-\frac{v^2}{\beta^2}\right) \tag{3.3.4}$$

เมื่อ β เป็นค่าคงตัว อาจเรียกว่า ตัวประกอบการกระจาย (Spread factor)

จะเห็นว่า $\lim_{v\to\infty} \Phi(v) \to 0$

จากสมการทั้งสองถือเป็นฟังก์ชั่นต้นแบบที่ใช้กันทั่วไป ซึ่ง $\Phi(.)$ อาจมีรูปแบบนอกเหนือจากที่กล่าวแล้วก็ได้ ในวิทยานิพนธ์ฉบับนี้ จะใช้ฟังก์ชั่นเกาส์เป็นฟังก์ชั่นมูลฐานแนวรัศมี



รูปที่ 3.10 เครือข่ายฟังก์ชั่นมูลฐานแนวรัศมี

กลไกของเครือข่ายฟังก์ชั่นมูลฐานแนวรัศมีที่มีสัญญาณขาเข้า n ตัวและมีสัญญาณออกเพียงหนึ่งค่า เป็นดังรูปที่ 3.10 โดยมีชั้นของฟังก์ชั่นมูลฐานแนวรัศมีเป็นชั้นซ่อน ซึ่งเครือข่ายกระทำการโยงของฟังก์ชั่นดังสมการข้างล่าง

$$y = \sum_{i=1}^{m} w_i e^{\frac{\|\mathbf{x}-\mathbf{c}_i\|^2}{\beta_i^2}} + w_{m+1} \qquad (3.3.5)$$

เมื่อ $\mathbf{x} \in \mathbf{R}^n$ เป็นสัญญาณขาเข้า, $\Phi(.)$ เป็นฟังก์ชั่นที่เจาะจงซึ่งโยงค่าจำนวนจริงบวกไปยังค่าจำนวนจริงใดๆ, $\|\ .\ \|$ เป็นนอร์มยูคลิดหรือนอร์มสอง ซึ่งนิยามดัง (3.3.6) $w_i, 0 \le i \le m+1$ เป็นค่าตัวถ่วงน้ำหนักหรือพารามิเตอร์ $\mathbf{c}_i \in \mathbf{R}^n$ เป็นที่รู้จักกันในนามของจุดศูนย์กลางของฟังก์ชั่นมูลฐานแนวรัศมี และ m เป็นจำนวนจุดศูนย์กลางดังกล่าว ถึงแม้ว่าที่กล่าวมานี้เป็นเครือข่ายที่มีสัญญาณขาออกเพียงสัญญาณเดียวเท่านั้น สำหรับกรณีเครือข่ายที่มีสัญญาณขาออกมากกว่าหนึ่งสัญญาณ สามารถถูกมองแยกออกเป็นกลุ่มของเครือข่ายที่มีสัญญาณขาออกเพียงสัญญาณเดียวได้ ซึ่งเป็นวิธีที่ตรงไปตรงมา ไม่ยุ่งยาก

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}} \qquad (3.3.6)$$

3.4 ข้อดี-ข้อเสียของเครือข่ายประสาทเทียม [14,15]

3.4.1 ข้อดีของเครือข่ายประสาทเทียม

3.4.1.1 เครือข่ายประสาทเทียมมีลักษณะปรับตัวได้ (Adaptation) กล่าวคือ เครือข่ายประสาทเทียมสามารถรับข้อมูลระบบและเรียนรู้จากข้อมูลดังกล่าว ดังนั้นเครือข่ายประสาทเทียมทำหน้าที่สรุปหาพารามิเตอร์ที่แสดงความสัมพันธ์ของข้อมูลระบบ ความสามารถดังกล่าวนี้แตกต่างจากเทคนิคมาตรฐานอย่างมาก เนื่องจากเครือข่ายประสาทเทียมไม่ขึ้นกับความรู้พื้นฐานก่อนหน้านี้ของนักเขียนโปรแกรม ทำให้เครือข่ายประสาทเทียมประหยัดเวลาในการเรียนรู้ความสัมพันธ์ของระบบซึ่งอาจจะยุ่งยากในการอธิบาย

3.4.1.2 เครือข่ายประสาทเทียมสามารถทำให้มีลักษณะทั่วไป (Generalization) กล่าวคือ เครือข่ายประสาทเทียมสามารถจัดการข้อมูลที่ไม่เคยนำเสนอมาก่อนหน้านี้ได้ทำให้มีความทนทานต่อความผิดพร่อง การทำให้มีลักษณะทั่วไปดังกล่าวมีประโยชน์ในทางปฏิบัติงาน เพราะข้อมูลจริงที่มีอยู่ส่วนมากมักมีสัญญาณที่ไม่พึงประสงค์รบกวนอยู่ตลอดเวลา

3.4.1.3 เครือข่ายประสาทเทียมเป็นระบบไม่เชิงเส้น (Nonlinear) กล่าวคือเครือข่ายประสาทเทียมสามารถจับปฏิสัมพันธ์ที่ซับซ้อนระหว่างตัวแปร

สัญญาณขาเข้าในระบบได้ ในระบบเชิงเส้น การเปลี่ยนสัญญาณขาเข้าหนึ่งสัญญาณส่งผลให้มีการเปลี่ยนแปลงที่เป็นสัดส่วนในสัญญาณขาเข้า และผลของสัญญาณขาเข้าขึ้นอยู่กับค่าของสัญญาณขาเข้าที่เปลี่ยนแปลงเท่านั้น ขณะที่ในระบบไม่เชิงเส้น ผลของสัญญาณขาเข้าขึ้นอยู่กับสัญญาณขาเข้าอื่นๆ และความสัมพันธ์เป็นฟังก์ชั่นที่มีอันดับสูง ซึ่งสอดคล้องกับโลกความเป็นจริงที่ว่า ระบบเป็นระบบไม่เชิงเส้น

3.4.1.4 เครือข่ายประสาทมีลักษณะขนานสูง (Highly parallel) กล่าวคือ ปฏิบัติการที่เป็นอิสระกันและคล้ายกันจำนวนมากสามารถประมวลผลได้ในเวลาเดียวกัน

3.4.2 ข้อเสียของเครือข่ายประสาทเทียม

3.4.2.1 ยากต่อการพิจารณาถึงที่มาของผลลัพธ์ที่ได้ กล่าวคือ เครือข่ายประสาทเทียมคล้ายกับความเชี่ยวชาญของมนุษย์ และแสดงความคิดเห็นซึ่งไม่ง่ายในการอธิบาย แต่บ่อยครั้งขึ้นอยู่กับความคิดเห็นมนุษย์ ถึงแม้จะขาดความแน่นอนที่เหมาะสมไปก็ตาม

3.4.2.2 ไม่เข้าใจวิธีการฝึกฝนได้อย่างสมบูรณ์ กล่าวคือ มีกฎเกณฑ์ไม่มากนักสำหรับการเลือกสถาปัตยกรรมที่พอเหมาะ และไม่มีวิธีที่ชัดเจนในการหาคำตอบที่ดีที่สุดขึ้นอยู่กับว่าต้องการความแม่นยำของข้อมูลในการฝึกฝนเท่าใด อย่างไรก็ตาม มีการวิจัยอย่างต่อเนื่องในการหาวิธีออกแบบและฝึกฝนเครือข่ายประสาทเทียมที่พอเหมาะ

3.4.2.3 ใช้เวลามากในการฝึกฝน เช่นเดียวกันมีการวิจัยอย่างต่อเนื่องในการหาวิธีใหม่ๆ ในการฝึกฝนเพื่อลดเวลาการฝึกฝนลง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย