บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 คุณภาพน้ำ

ปัญหามลพิษทางน้ำในเรื่องน้ำเสียเกิดขึ้นพร้อมๆ กับการเจริญเติบโตของชุมชน เนื่องจากน้ำเสียเกิดขึ้นจากการใช้น้ำเพื่อวัตถุประสงค์ต่างๆ ในสมัยก่อนปริมาณน้ำเสียที่เกิดขึ้นมีจำนวนไม่มากเมื่อระบายลงสู่แหล่งน้ำสาธารณะ ธรรมชาติจะสามรถทำความสะอาดน้ำเสียได้ทัน อย่างไรก็ตามเมื่อมีการขยายตัวของชุมชน และมีการพัฒนาอุตสาหกรรมเพิ่มขึ้น น้ำเสียก็มีปริมาณเพิ่มขึ้นจนถึงจุดที่การทำความสะอาดน้ำเสียที่เกิดขึ้นตามวิธีการทางธรรมชาติไม่ได้ผล การเน่าเหม็นของเสียก็ปรากฏขึ้นทำให้จำเป็นต้องมีการบำบัดน้ำเสียด้วยวิธีต่างๆ ในกระบวนการบำบัดน้ำเสีย น้ำเสียก่อนการบำบัดที่ถูกรวบรวมและควบคุมอัตราการไหลเข้าสู่ระบบบำบัด เรียกว่า น้ำเสียก่อนการบำบัด (Influent) และเมื่อน้ำเสียนั้นผ่านการบำบัดแล้วระบายออกสู่สิ่งแวดล้อม เรียกว่า น้ำทิ้ง (Efluent)

น้ำเสียตามพระราชบัญญัติส่งเสริมและรักษาสิ่งแวดล้อมแห่งชาติ หมายถึง ของเสียที่อยู่ในสภาพที่เป็นของเหลวรวมทั้งมวลสารที่ปะปนหรือปนเปื้อนอยู่ในของเหลวนั้น ลักษณะของน้ำเสียแบ่งออกได้ 3 ด้าน คือ ด้านกายภาพ ด้านเคมี และด้านชีวภาพ

1.ลักษณะของน้ำเสียทางกายภาพ เช่น

- ปริมาณของแข็งที่ละลายน้ำได้ทั้งหมด หมายถึง ปริมาณของแข็งที่ละลายน้ำและสามารถไหลผ่านกระดาษกรองใยแก้ว แล้วนำน้ำไปที่กรองได้ ไประเหยจนแห้ง แล้วจึงนำไปอบ
- ของแข็งแขวนลอย หมายถึงปริมาณของแข็งแขวนลอยที่เหลือค้างบนกระดาษกรองใยแก้ว
- ความขุ่น หมายถึง สมบัติทางแสงของสารแขวนลอยซึ่งทำให้แสงกระจาย และดูดกลืนมากกว่าที่จะอมให้แสงผ่านเป็นเส้นตรง ความขุ่นของน้ำเกิดการมีสารแขวนลอยต่างๆ เช่น ดิน ดินตะกอน

2.ลักษณะของน้ำเสียทางเคมี เช่น

- ออกซิเจนละลาย การหาดีโอ(DO)หรือออกซิเจนละลาย สามารถทำได้ทั้งวิธีทางเคมี และใช้เครื่องวัดโดยตรง
- บีโอดี(BOD)หมายถึง ปริมาณออกซิเจนที่ใช้ในการย่อยสลายสารอินทรีย์โดยแบคทีเรีย
- ซีโอดี(COD)หมายถึง ปริมาณออกซิเจนทั้งหมดที่ต้องการใช้เพื่อออกซิเดชันสารอินทรีย์ในน้ำด้วยสารเคมีซึ่งมีอำนาจในการออกซิไดส์สูงในสารละลายที่เป็นกรด ให้เป็นก๊าซคาร์บอนไดออกไซด์ในน้ำ ค่าซีโอดีมีความสำคัญในการวิเคราะห์คุณภาพน้ำทิ้ง การควบคุมระบบบำบัดน้ำทิ้ง การควบคุมระบบบำบัดน้ำเสีย ค่า COD นี้มีหน่วยเป็น มิลลิกรัม/ลิตร

- ค่าความกรด-ด่าง(pH)มีความสำคัญในการควบคุมคุณภาพน้ำและน้ำเสียควบคุมให้เหมาะสมกับการเจริญเติบโตของสิ่งมีชีวิต เพื่อไม่ให้เกิดการกัดกร่อนของท่อ เพื่อใช้ในการควบคุมสารเคมีที่ใช้บำบัดน้ำเสียให้ทำงานได้อย่างมีประสิทธิภาพ โดยทั่งไปน้ำมีค่า pH อยุ่ในช่วง 5-8 ค่า pH เป็นค่าที่แสดงปริมาณความเข้มข้นของอนุภาคไฮโดรเจนในน้ำ
- ไนโตรเจน เป็นธาตุที่มีความสำคัญในการสังเคราะห์โปรตีน ทำให้พืชน้ำมีการเจริญเติบโตได้อย่างรวดเร็ว
- สารโลหะหนักชนิดต่างๆขึ้นอยู่กับชนิดของอุตสาหกรรม สารโลหะหนักยอมให้มีได้ในน้ำในปริมาณที่น้อยมากเนื่องจากบางตัวให้ความเป็นพิษสูง แต่บางชนิดหากทีปริมาณไม่มากนักจะมีผลดีต่อสิ่งมีชีวิตในน้ำ

3.ลักษณะของเสียทางชีวภาพ เช่น

- แบคทีเรีย คือ จุลินทรีย์เซลล์เดียว มีขนาดเล็ก ไม่สามารถมองเห็นได้ด้วยตาเปล่า เป็นผู้ย่อยสลายในแหล่งน้ำ
- รา เป็นจุลินทรีย์ที่มีหลายเซลล์ ไม่มีคลอโรฟีลล์ รามีความสำคัญในการย่อยสลายพวกคาร์บอนที่มีค่า pH ต่ำ รามีบทบาทสำคัญในการย่อยสลายสารอินทรีย์ในระบบบำบัดน้ำเสียบางระบบ

โดยแหล่งกำเนิดน้ำเสียสามารถ แบ่งได้หลักๆดังนี้

1.น้ำเสียจากชุมชน หมายถึง น้ำที่เกิดจากการใช้ประโยชน์ในกิจกรรมต่างๆ และระบายน้ำทิ้งลงสู่ท่อระบายน้ำ แหล่งรองรับน้ำเสีย หรือแหล่งน้ำธรรมชาติ โดยไม่ได้ผ่านการบำบัด

2.หาปริมาณออกซิเจนที่จุลินทรีย์ใช้ย่อยสลายสารอินทรีย์ต่างๆในน้ำเสีย วิธีนี้ถือว่าเป็นวิธีทางชีววิทยา

3.หาปริมาณจุลินทรีย์ในน้ำ

4.วัดความเข้มข้นของสารต่างๆที่ละลายอยู่ในน้ำ

2.2 การประมาณค่าออกซิเจนที่ใช้ในการย่อยสลายสารอินทรีย์โดยแบคทีเรีย (BOD)

BOD หมายถึงปริมาณออกซิเจนที่แบคทีเรียใช้ในการย่อยสลายสารอินทรีย์ชนิดที่ย่อยสลายได้ภายใต้ออกซิเจน โดยจุลินทรีย์จะใช้ออกซิเจนที่ละลายอยู่น้ำเพื่อการเจริญเติบโต หากมีค่า BOD สูง แสดงว่าปริมาณออกซิเจนจะถูกใช้ไปมาก และแสดงว่ามีปริมาณสารอินทรีย์ในน้ำมากด้วย น้ำจึงมีความสกปรกสูง ดังนั้นการตรวจวัดค่า BOD จึงต้องกระทำภายใต้สภาวะที่เหมือนกับเกิดขึ้นในธรรมชาติมากที่สุด นั่นคือ ต้องทำการอบ(Incubate) ที่อุณหภูมิประมาณ 20 องศาเซลเซียส ซึ่งเป็นอุณหภูมิที่ใกล้เคียงกลับน้ำทั่วไป และใช้เวลาในการอบ 5 วัน เนื่องจากเป็นระยะที่เหมาะสมต่อการย่อยสลายของแบคทีเรีย หากใช้เวลาน้อยกว่านี้จะมีการใช้ออกซิเจนน้อย แต่ถ้าให้ระยะเวลานานเกินไป ปฏิกิริยาย่อยสลายจะเกิดในทิศทางย้อนกลับ ทำให้ไม่ได้ค่าที่แท้จริง ดังนั้นจึงเรียกค่า BOD มาตรฐานนี้เรียกว่า BOD_5

การประมาณการค่าปริมาณออกซิเจนที่ใช้ในการย่อยสลายสารอินทรีย์โดยแบคทีเรียที่มีอยู่ในน้ำสามารถบ่งบอกถึงการเปลี่ยนแปลงของคุณภาพน้ำและช่วยในการประเมินการเฝ้าระวัง

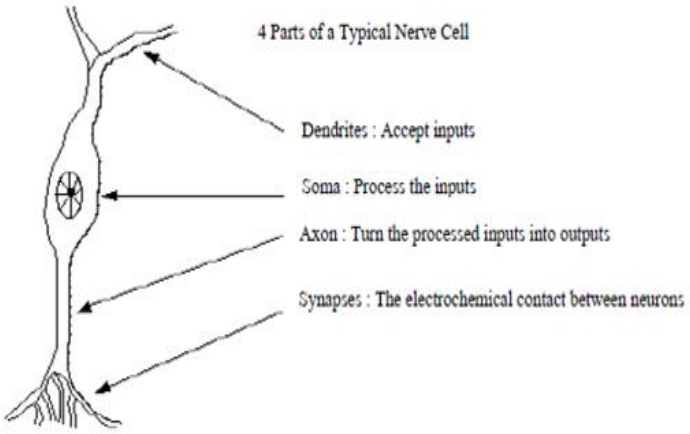
วางแผนจัดการป้องกันและแก้ไขปัญหาคุณภาพน้ำได้อย่างเหมาะสม โดยเฉพาะบริเวณที่มีแนวโน้มระดับมลพิษน้ำสูงหรือเป็นบริเวณที่มีประชากรในกลุ่มที่มีผลกระทบต่อระดับมลพิษทางน้ำ เช่น บริเวณอาคารบ้านเรือนที่อยู่ริมน้ำ การคมนาคมทางน้ำ เป็นต้น ซึ่งจำเป็นต้องทราบระดับคุณภาพทางน้ำ การตรวจวัดจริงเป็นวิธีการที่ใช้ดำเนินการศึกษาที่นิยมใช้ แต่การตรวจวัดจริงนั้นทำได้จำกัด เช่น สามารถตรวจวัดได้เพียงชั่วระยะเวลาหนึ่ง หรือบางครั้งต้องใช้เวลามาก ความไม่เพียงพอของจำนวนเครื่องมือและงบประมาณมีจำนวนจำกัด ดังนั้น การประมาณค่าปริมาณออกซิเจนที่ใช้ในการย่อยสลายสารอินทรีย์โดยแบคทีเรียที่เป็นพารามิเตอร์หลักตัวหนึ่งที่สามารถบ่งบอกถึงระดับคุณภาพน้ำได้โดยการใช้แบบจำลองโครงข่าย ซึ่งเป็นอีกแนวทางของการเฝ้าระวังคุณภาพน้ำและ ลดข้อจำกัดของการตรวจวัดจริง

2.3 โครงข่ายใยประสาทประดิษฐ์ (Artificial Neural Network : ANN)

แบบจำลองโครงข่ายใยประสาทประดิษฐ์ ประกอบด้วย ระบบการคำนวณแบบไม่เป็นเชิงเส้นซึ่งมีวิธีการดำเนินระบบเป็นแบบขนานและมีรูปแบบการเรียนรู้แบบโครงข่ายใยประสาทชีวภาพ(Lippmann, 1987) โดยประกอบไปด้วยนิวรอล (หรือ โหนด หรือ หน่วยประมวลผล) ซึ่งรวมกันอยู่เป็นชั้น ๆ ซึ่งสามารถรับข้อมูลเข้าได้หลายค่า และสามารถคำนวณผล โดยจะให้ผลลัพธ์ค่าเดียวหรือหลายค่าก็ได้ (Klimasauskas, 1993) ซึ่งการคำนวณในระบบประกอบไปด้วยฟังก์ชั่นง่าย ๆ เช่นฟังก์ชั่นการรวม และ ฟังก์ชั่นการคูณ (Arciszewski and Ziarko, 1992) โดยมีความสามารถในการเรียนรู้จากตัวอย่างหลาย ๆ ตัวอย่าง ซึ่งจะหาแนวทางการแก้ปัญหา แม้แต่ข้อมูลที่ถูกป้อนเข้ามาไม่สมบูรณ์หรือผิดพลาด ระบบจะเปรียบเทียบผลลัพธ์ที่คลาดเคลื่อน และปรับเปลี่ยนวิธีการประมวลผลเพื่อให้ผลลัพธ์ถูกต้องที่สุด ระบบจะประมวลผลข้อมูลโดยคอมพิวเตอร์อย่างรวดเร็ว (Flood and Kartam, 1994)

2.3.1 แนวคิดพื้นฐานของโครงข่ายใยประสาทประดิษฐ์

Lippmann (1987), Chester (1993) และ Kireetoh (1995) ได้อธิบายถึงนิวรอลว่าคล้ายกับระบบประสาทของมนุษย์ โดยรับสัญญาณข้อมูลที่ส่งเข้ามา และจะถูกกระตุ้น แต่ละเซลล์ประกอบด้วย ปลายในการรับกระแสประสาท เรียกว่า "เดนไดรท์" (Dendrite) ซึ่งเป็นข้อมูลป้อนเข้า และปลายในการส่งกระแสประสาท เรียกว่า "แอคซอน" (Axon) ดังรูปที่ 2.3 ซึ่งเป็นเหมือนข้อมูลที่ส่งออกมาของเซลล์ เซลล์เหล่านี้ทำงานด้วยปฏิกิริยาไฟฟ้าเคมี เมื่อมีการกระตุ้นด้วยสิ่งเร้าภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรท์เข้าสู่นิวเคลียส ซึ่งจะเป็นตัวตัดสินว่าต้องกระตุ้นเซลอื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็จะกระตุ้นเซลอื่น ๆต่อไปผ่านทางแอคซอนของมัน ผลการกระตุ้นด้วยสิ่งเร้าที่เหมือนหรือมีลักษณะพิเศษบางอย่างเหมือนกัน จะให้ผลลัพธ์สุดท้ายเป็นค่าที่ค่อนข้างแน่นอน



รูปที่ 2.1 โครงสร้างระบบประสาทในสมอง

โครงข่ายใยประสาทประดิษฐ์เป็นปัญญาประดิษฐ์ (Artificial Intelligence) ที่สร้างเลียนแบบระบบประสาทของมนุษย์ (Lippmann, 1987; Caudill and Butler, 1990; Klimasauskas,1993; Medsker et al., 1993) โดยระบบการทำงานต่าง ๆ ของโครงข่ายใยประสาทประดิษฐ์มีลักษณะคล้ายกับระบบประสาทของมนุษย์ คือ มีการเรียนรู้จากประสบการณ์ที่ได้รับ และสามารถให้คำตอบได้แม้แต่ข้อมูลที่ป้อนเข้าจะผิดพลาด หรือไม่สมบูรณ์ โดยหาวิธีการแก้ปัญหาจากประสบการณ์การเรียนรู้ที่ผ่านมา และสามารถที่จะพัฒนาให้ไปเป็นโปรแกรมที่สามารถโต้ตอบกับมนุษย์ได้

2.3.2 การสร้างแบบจำลองโครงข่ายใยประสาทประดิษฐ์

โครงข่ายใยประสาทประดิษฐ์ คือ เครื่องจักรการเรียนรู้ ตั้งอยู่บนพื้นฐานความคิดเกี่ยวกับการปรับปรุงตัวแปรควบคุมต่าง ๆ ภายในด้วยตัวมันเอง โดยมีองค์ประกอบภายในระบบซึ่งประกอบไปด้วย 5 องค์ประกอบหลัก คือ หน่วยการเรียนรู้, โครงข่ายใยประสาท, แผนการเรียนรู้, กระบวนการเรียนรู้ และ กระบวนการวิเคราะห์ (Adeli, 1992) ซึ่ง Elazouni et al. (1997) ได้จำแนกส่วนประกอบของโครงข่ายใยประสาทประดิษฐ์ เป็ น 3 ขั้น คือ 1) การออกแบบ 2) การสร้างแบบจำลอง และ 3) การทดสอบและหาผลลัพธ์โดยขั้นการออกแบบจะประกอบไปด้วยกัน 2 ส่วน คือ การวิเคราะห์โครงสร้างของปัญหา และการวิเคราะห์ปัญหา ส่วนขั้นการสร้างแบบจำลอง จะแบ่งย่อยออกเป็น 3 ขั้นตอนคือ1) การเลือกข้อมูล 2) การเลือกรูปแบบโครงข่าย 3) การสอนและการทดสอบโครงข่าย

2.3.3 ข้อมูลป้อนเข้า

โครงข่ายใยประสาทประดิษฐ์ ประกอบไปด้วยตัวแปรอิสระ หรือ ข้อมูลป้อนเข้าและตัวแปรตาม หรือ ผลลัพธ์ โดยหลักการเลือกตัวแปรที่ใช้ในโครงข่ายที่เกี่ยวข้องจะมี 2 แบบ(Smith, 1993) คือ วิธีแรก ข้อมูลจะต้องถูกแปลงรูปให้อยู่ในรูปที่เหมาะสม และวิธีที่สอง คือการเลือกข้อมูลโดยใช้พื้นฐานระหว่าง predictiveness และ covariance โดยปกติแล้ว ตัวแปรอิสระที่ถูกเลือกจะมีความสามารถในการทำนายผลหากตัวแปรที่เลือกมีความสัมพันธ์กันในทางตรงกันข้ามหากตัวแปรอิสระ 2 ตัวมีความสัมพันธ์ต่อกัน จะทำให้แบบจำลองมีความอ่อนไหว (sensitive) และเกิดปัญหาที่

เรียกว่า over fitting และ limit generalization ด้วยเหตุผลนี้ การเลือกข้อมูลจะต้องเลือกเฉพาะตัวแปรอิสระที่มีความสามารถทำนายผลเพื่อให้ได้ ผลลัพธ์หรือตัวแปรตาม โดยตัวแปรอิสระที่เลือกมานั้นจะต้องไม่มีความสัมพันธ์กัน แต่อย่างไรก็ดี ก็ขึ้นอยู่กับรูปแบบของโครงข่ายที่ใช้และเพื่อที่จะลดจำนวนตัวอย่างที่ใช้ในการสอนและเวลาที่ใช้ในการเรียนรู้ ก็ควรจะต้องมีการคัดเลือกข้อมูลป้อนเข้าให้เหมาะสม เพราะการคัดเลือกข้อมูลเป็นปัจจัยที่สำคัญในการสร้างแบบจำลอง (Wu and Lim, 1993)

2.3.4 ชั้นซ่อน (Hidden Layer)

ชั้นซ่อนเป็นชั้นประมวลผลที่อยู่ระหว่างชั้นข้อมูลป้อนเข้า และชั้นแสดงผลลัพธ์โดยปกติแล้ว ชั้นซ่อนอาจมีมากกว่า 1 ชั้น โดยโครงข่ายจะสามารถประมวลหาฟังก์ชั่นที่เหมาะสมจากปัญหาที่ซับซ้อนได้หากมีชั้นซ่อนที่มากพอ (Lippmann, 1987) ข้อมูลที่ได้จากชั้นซ่อนจะได้เป็นตัวแปรใหม่ที่จะถูกส่งต่อให้กับชั้นแสดงผลลัพธ์ หรือชั้นตัวแปรตาม ถ้าโครงข่ายแบบแพร่กลับ(Back-propagation) มีชั้นซ่อนที่น้อยเกินไปแล้วจะทำให้โครงข่ายไม่สามารถที่จะหาทางแก้ปัญหาได้ (Karunasekera, 1992) แต่ถึงอย่างไรก็ตามถ้าเกิดโครงข่ายมีชั้นซ่อนที่มากจนเกินไป จะทำให้โครงข่ายมีระยะเวลาในการเรียนรู้นาน William (1993) ได้ให้ข้อคิดเห็นว่า หากมีชั้นซ่อนที่มากก็จะไม่ช่วยให้โครงข่ายมีประสิทธิภาพมากขึ้น หรืออีกนัยหนึ่ง Rumelhart (1988) ได้กล่าวว่าการที่มีโหนดในแต่ละชั้นที่มากเกิน จะทำให้โครงข่ายไม่สามารถที่จะหาจุดสิ้นสุดได้ การที่มีโหนดในชั้นซ่อนมากเกินจะทำให้เกิดปัญหาที่เรียกว่า over fitting โดยโครงข่ายจะจำลองโครงสร้างใหม่เกินความเป็นจริงจาก noise ของข้อมูล แทนที่จะหาฟังก์ชั่นที่เหมาะสมในการวิเคราะห์ปัญหาให้ถูกต้องตามที่ควรจะเป็น (Smith, 1993)ดังนั้น การที่จะทำให้โครงข่ายเกิดประสิทธิภาพสูงสุด ต้องกำหนดให้มีโหนดในชั้นซ่อนอยู่ให้น้อยเท่าที่จะเป็นไปได้ (Khan, Topping and Bahreininejad, 1993) Berke และ Hajela(1991) ได้ให้ความเห็นว่า จำนวนของโหนดในชั้นซ่อนควรอยู่ระหว่างค่าเฉลี่ยของผลรวมของโหนดในชั้นข้อมูลป้อนเข้า และชั้นแสดงผลลัพธ์ Rogers และ Ramarsh (1992) ได้ให้ความคิดเห็นว่าในการกำหนดโหนดในชั้นซ่อนควรจะดูจากผลรวมของโหนดในชั้นข้อมูลป้อนเข้าและชั้นแสดงผลลัพธ์ Soemardi (1996) ได้แสดงความเห็นว่า จำนวนโหนดในชั้นซ่อนควรจะมีค่าเท่ากับร้อยละ75 ของโหนดในชั้นข้อมูลป้อนเข้า ดังนั้นจึงสรุปจากข้อคิดเห็นได้ว่าจำนวนโหนดในชั้นซ่อนที่มากที่สุดควรจะเท่ากับผลรวมของโหนดในชั้นข้อมูลป้อนเข้า และชั้นแสดงผลลัพธ์ และจำนวนโหนดที่น้อยที่สุดควรจะเท่ากับร้อยละ 75 ของโหนดในชั้นข้อมูลป้อนเข้า หรือ เท่ากับค่าเฉลี่ยของผลรวมของโหนดในชั้นข้อมูลป้อนเข้าและชั้นแสดงผลลัพธ์

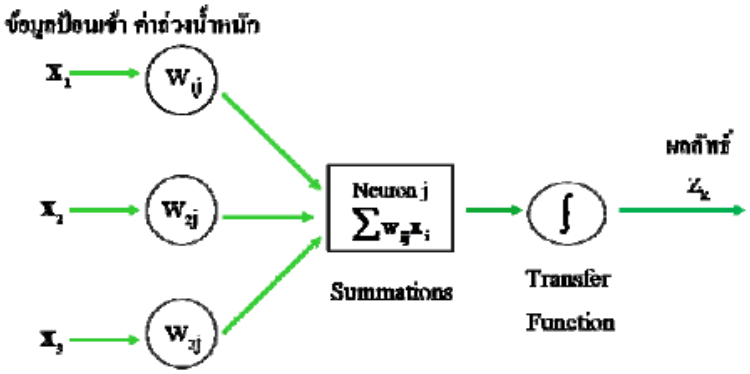
2.3.5 ค่าถ่วงน้ำหนักและไบแอส (Weights and biases)

ค่าถ่วงน้ำหนักถูกแทนด้วยตัวเลขเพื่อแสดงถึงความแรงในการเชื่อมต่อของโหนดแต่ละโหนดที่ถูกเชื่อมต่อเข้าด้วยกัน ซึ่งผลรวมของค่าถ่วงน้ำหนักที่ป้อนเข้าจะไปปรับปรุงการประมวลผลในแต่ละโหนด ค่าถ่วงน้ำหนักคือค่าความแรงสัมพัทธ์ (ในทางคณิตศาสตร์) ของการเชื่อมต่อซึ่งส่งผลต่อการส่งผ่านข้อมูลจากชั้นหนึ่งไปยังชั้นต่อไป (Medsker et al., 1993) โดยปกติค่าถ่วงน้ำหนักจะถูกกำหนด และเริ่มป้อนเข้าสู่โครงข่ายในขั้นตอนการเรียนรู้ ซึ่งต้องมีหลักการในการกำหนดค่าเพื่อที่จะให้โครงข่ายสามารถแก้โจทย์ปัญหา และลดเวลาการเรียนรู้ได้ สำหรับโครงข่ายใด ๆ ค่าถ่วง

น้ำหนักจะมีค่าเท่ากับผลคูณของจำนวนโหนดของทุก ๆ การเชื่อมต่อ และค่าของไบแอสจะเท่ากับผลรวมของจำนวนโหนดของทุก ๆ การเชื่อมต่อ

2.3.6 ฟังก์ชั่นการรวมและฟังก์ชั่นการแปลงค่า (Summation and transfer function)

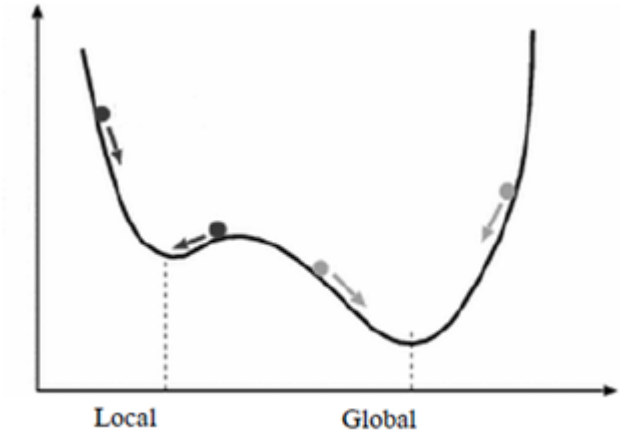
ฟังก์ชั่นการรวม คือ ฟังก์ชั่นการหาผลเฉลี่ยค่าถ่วงน้ำหนักของทุก ๆ โหนดที่เชื่อมต่อกัน โดยมีขั้นตอนคือ นำค่าของข้อมูลป้ อนเข้า (Input) ในแต่ละโหนดคูณกับค่าถ่วงน้ำหนักของแต่ละโหนดและรวมผลลัพธ์ของทุก ๆ โหนดเข้าด้วยกัน ดังรูปที่ 2.5 ส่วนฟังก์ชั่นการแปลงค่าคือความสัมพันธ์ระหว่างระดับการกระตุ้นภายในโหนด (N) กับผลลัพธ์ที่ได้ (Output) โดยจะอยู่ในรูปของฟังก์ชั่นซิกมอยด์ f(N) โดยมีข้อแม้ว่า 1) มีความต่อเนื่อง 2) ค่าของฟังก์ชั่นซิกมอยด์จะต้องเพิ่มขึ้นเมื่อ N เพิ่มขึ้น (Smith, 1993)



รูปที่ 2.2 กระบวนการทำงานของโครงข่ายใยประสาทประดิษฐ์ในโหนดย่อย

2.3.7 อัตราการเรียนรู้และโมเมนตัม (Learning rate and momentum)

ถ้าโครงข่ายที่มีอัลกอลิทึมแบบแพร่กลับ (Back-propagation) มีขนาดที่ใหญ่และมีชุดการสอนโครงข่ายที่ใหญ่จะทำให้โครงข่ายมีการเรียนรู้นาน (Khan et al., 1993) โครงข่ายแบแพร่กลับ จะมีข้อจำกัด คือ ไม่สามารถกำหนดเวลาในการเรียนรู้ที่แน่นอนได้ โดยโครงข่ายมีโอกาสที่จะหลงทาง ไม่สามารถค้นหาคำตอบที่มีความผิดพลาดน้อยที่สุด ทำให้พบกับคำตอบที่ local minimumก่อนที่จะพบ global minimum ได้ ดังรูปที่ 2.6 ดังนั้นจึงมีความสำคัญอย่างยิ่งในการเลือกอัตราการเรียนรู้ และค่าโมเมนตัมที่เหมาะสม สำหรับการใช้อัลกอลิทึมแบบแพร่กลับ แต่ถึงอย่างไรก็ดีหลักการในการหาอัตราการเรียนรู้ และค่าโมเมนตัมที่เหมาะสม คือ ต้องใช้วิธีลองผิดลองถูก (trial-and-error) (Anderson et al., 1993)



รูปที่ 2.3 แสดงจุด local minimum และ global minimum

เมื่ออัตราการเรียนรู้สูงจะส่งผลให้การเรียนรู้เป็นไปอย่างรวดเร็วซึ่งอาจจะทำให้การเรียนรู้จบลงที่ local minimum แต่ในทางตรงกันข้าม หากให้โครงข่ายมีอัตราการเรียนรู้ที่ต่ำจะทำให้เวลาที่ใช้ในการเรียนรู้เพื่อให้เข้าใกล้ global minimum นานขึ้น (Khan et al., 1993) ซึ่งสำหรับในแต่ละชั้นของโครงข่ายเดียวกันอาจมีค่าอัตราการเรียนรู้แตกต่างกันได้ (Bhokha, 1998)สำหรับการแก้ปัญหาที่เกิดขึ้นนี้ทำได้ด้วยการป้ อนค่าโมเมนตัม ซึ่งจะนำไปคูณกับค่าถ่วงน้ำหนักที่ถูกปรับแก้ในรอบที่ผ่านมา ทำให้อัตราการเรียนรู้เร็วขึ้น (Khan et al., 1993)

2.3.8 กระบวนการสอนหรือการเรียนรู้ (Training or Learning)

กระบวนการเรียนรู้เป็นกระบวนการหนึ่งในโครงข่ายที่เรียนรู้จากความผิดพลาดโดยมีหลักการ 3 ข้อ คือ 1) คำนวณหาคำตอบ 2) ตรวจสอบคำตอบว่าถูกต้องหรือไม่ 3) ปรับแก้ค่าถ่วงน้ำหนักแล้วคำนวณใหม่อีกครั้ง (Medsker et al., 1993) กระบวนการสอน คือ การนำข้อผิดพลาดจากการคำนวณครั้งก่อนมาปรับแก้ค่าถ่วงน้ำหนักให้กับการสอนในรอบต่อไป ทำให้คำตอบที่ได้มีความถูกต้องสูงขึ้น (Klimasauskas, 1993)

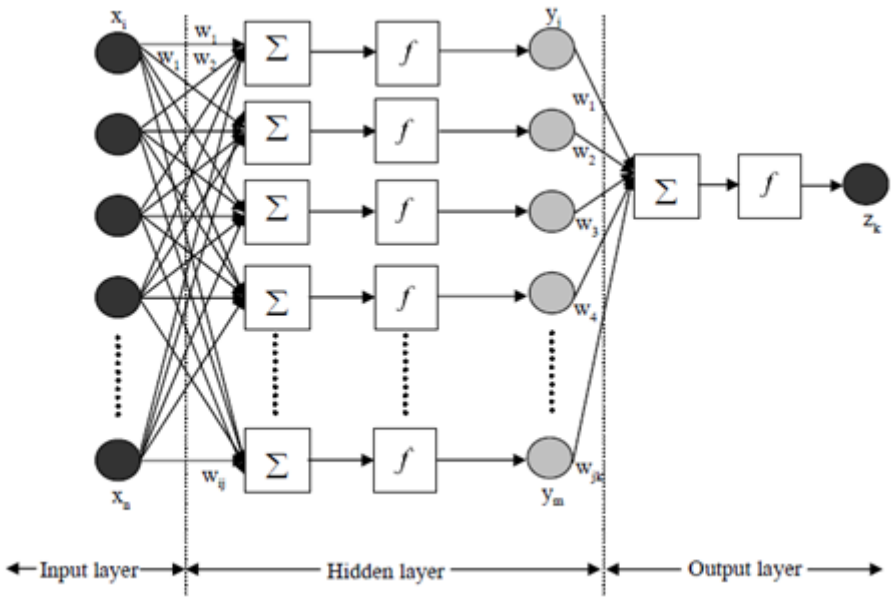
ในขั้นตอนการเรียนรู้จะเป็นกระบวนการที่เกี่ยวข้องกับ การปรับค่าน้ำหนักของแต่ละโหนด โดยอาศัยจากประสบการณ์การเรียนรู้ของโครงข่าย การเรียนรู้ของโครงข่ายในแต่ละรอบจะนำผลลัพธ์ที่ประมวลได้ มาเปรียบเทียบกับค่าจริงที่ได้จากการทดลอง ซึ่งอาจจะมีความคลาดเคลื่อน หลังจากนั้นโครงข่ายจะย้อนกลับไปเรียนรู้ในรอบต่อไปพร้อมกับการปรับแก้ค่าถ่วงน้ำหนักเพื่อให้การประมวลผลรอบต่อไปมีความแม่นยำมากขึ้น Bhokha (1998) ได้กล่าวว่าการปรับแก้อาจจะเป็นการปรับขึ้นหรือปรับลงก็ได้ Klimasauskas (1993) กล่าวว่า การวัดผลว่าโครงข่ายสามารถเรียนรู้ได้ดีเพียงใดจะดูจากตัวชี้วัดต่าง ๆ เช่น ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(mean square error) ในชั้นแสดงผลลัพธ์Lippmann (1987) และ Smith (1993) ได้กล่าวไว้ว่า กระบวนการเรียนรู้สามารถแบ่งออกเป็น 2 ลักษณะ คือ

1) การเรียนรู้แบบมีครูสอน (supervised training) ซึ่งจะประกอบไปด้วย คู่อันดับของข้อมูลป้อนเข้า และผลลัพธ์จริง ซึ่งเมื่อโครงข่ายเริ่มการเรียนรู้จากข้อมูลป้อนเข้า และคำนวณหาผลลัพธ์ได้แล้ว จึงจะนำไปเปรียบเทียบกับผลลัพธ์จริง เพื่อหาความคลาดเคลื่อนซึ่งความคลาดเคลื่อนดังกล่าวจะถูกส่งกลับเข้าไปยังโครงข่ายพร้อมกับการปรับแก้ค่าถ่วงน้ำหนักเพื่อให้โครงข่ายคำนวณผลลัพธ์ใหม่ให้มีความคลาดเคลื่อนน้อยที่สุด

2) การเรียนรู้แบบไม่มีครูสอน (unsupervised training) ได้คิดค้นโดย Kohonen (1984) ซึ่งแตกต่างจากแบบจำลองที่เลียนแบบระบบสมองของมนุษย์ โดยไม่ต้องใช้ผลลัพธ์จริงมาทำการเปรียบเทียบ แต่จะใช้คุณสมบัติทางสถิติของข้อมูลชุดทดสอบมาจัดกลุ่มเป็นหมวดหมู่หลังจากที่ป้อนข้อมูลเข้าไปแล้ว แบบจำลองจะทำการประมวลผลลัพธ์ที่เป็นไปได้ออกมาเป็นชุด ๆ (Heaton, 2004)

2.3.8.1 การเรียนรู้แบบแพร่กลับ (Back-propagation)

ในช่วงปี 1950 Rosenblatt ได้สร้างโครงข่ายชั้นเดียวแบบง่าย ๆ ขึ้นมาโดยมีชื่อเรียกว่า perceptron หลังจากนั้นต่อมา Widrow และ Hoff ได้สร้างอัลกอลิทึมขึ้นมาใหม่ที่สามารถอธิบายได้ด้วยกฎของเดลตา (Delta rule) โดยใช้การเรียนรู้ข้อมูลที่มีการปรับค่าถ่วงน้ำหนักซึ่งมีชื่อเรียกว่า Widrow-Hoff rule (Lippmann, 1987) ต่อมาในปี 1986 Rumelhart, Hilton และ Williams ได้ศึกษาค้นคว้าเพิ่มเติมจนเกิดกระบวนการเรียนรู้แบบแพร่กลับ (Back-propagation) หรือเรียกว่า Generalized Delta Rule (GDR) การเรียนรู้แบบแพร่กลับนี้ เป็นที่ยอมรับกันอย่างแพร่หลายโดยเป็นการเรียนรู้แบบมีครูสอน และมีระบบการเชื่อมโยงแบบเคลื่อนไปข้างหน้าหลายชั้น (multilayer feed forward) (Bhokha, 1998) ดังรูปที่ 2.4 สำหรับวิธีการเรียนรู้แบบแพร่กลับ จะมีการปรับค่าถ่วงน้ำหนัก (w) ในทุก ๆ รอบการเรียนรู้เพื่อให้เกิดค่าความผิดพลาดที่น้อยที่สุด โดยจะเริ่มต้นปรับแก้ค่าถ่วงน้ำหนักตั้งแต่ชั้น output หลังจากนั้นทำการปรับย้อนกลับมาที่ชั้นซ่อนชั้นสุดท้าย และชั้นซ่อนชั้นต่อ ๆ มา จนกระทั่งถึงชั้นซ่อนชั้นแรก



รูปที่ 2.4 โครงข่ายใยประสาทประดิษฐ์แบบแพร่กลับ

2.3.9 การหยุดการสอน (Stop training)

การหยุดการสอนโครงข่ายสามารถกระทำได้ 2 วิธี คือ 1) การกำหนดรอบการสอน (Epochs) 2) การกำหนดค่า error ที่ยอมรับได้ (Bhokha, 1998) Carpenter (1993) ได้แนะนำให้กำหนดจำนวนรอบการสอนอยู่ที่ 20,000 ถึง 100,000 รอบ และอีกวิธี คือ การกำหนดค่าความคลาดเคลื่อนระหว่างข้อมูลจริง และผลลัพธ์ที่โครงข่ายสามารถคำนวณได้ (khan et al., 1993) แต่ข้อควรระวังก็คือ การสอนโครงข่ายที่นานเกินไปจะทำให้เกิดปัญหาที่เรียกว่า over fitting ได้ ดังรูปที่ 2.5 ซึ่งคือปัญหาที่โครงข่ายสามารถที่จะเรียนรู้จนได้ผลลัพธ์ที่มีค่า error ที่น้อยที่สุด แต่เมื่อนำมาตรวจสอบความถูกต้องด้วยชุดทดสอบแล้ว ปรากฏว่าไม่สามารถที่จะให้ผลลัพธ์ได้ดีจริง (Bhokha, 1998)

รูปที่ 2.5 กรณี Over fitting

2.3.10 ชุดข้อมูล (Samples)

ชุดข้อมูล คือ ข้อมูลที่ทราบตัวแปรต้นและตัวแปรตามเพื่อนำมาใช้สอนโครงข่าย Yeh et al. (1993) ได้กล่าวไว้ว่า แหล่งที่มาของชุดข้อมูลแบ่งออกได้เป็น 3 ลักษณะ คือ 1) แบบสอบถาม 2) ข้อมูลทางสถิติ และ 3) จากการทดลอง โดยชุดข้อมูลดังกล่าวจะนำมาแบ่งออกเป็น 2 ชุด คือ ชุดการสอน (Training set) และ ชุดทดสอบ (Test set) Klimasauskas (1993) ได้ให้ความเห็นว่า ควรให้มีจำนวนชุดการสอนอย่างน้อย 5 ชุด เพื่อใช้ในการสอนโครงข่าย

2.3.11 การทดสอบโครงข่าย (Testing)

Smith (1993) กล่าวว่า การทดสอบโครงข่ายเป็นการทดสอบว่าโครงข่ายสามารถที่จะเรียนรู้จากชุดการสอน (Training set) ได้ดีเพียงไร โดยใช้ชุดข้อมูลที่ไม่เคยใช้สำหรับการสอนมาทดสอบ เรียกว่า ชุดทดสอบ (Test set) ซึ่งโครงข่ายที่สามารถให้ผลลัพธ์ที่แม่นยำได้เมื่อใช้ชุดทดสอบมาทดสอบ จะเป็นโครงข่ายที่น่าเชื่อถือในการทดสอบโครงข่ายสามารถแบ่งออกได้เป็น 2 ลักษณะ คือ 1) แบ่งข้อมูลออกเป็น 2 ชุด (McKim et al., 1996) โดยชุดแรกไว้สำหรับสอนโครงข่ายให้จดจำรูปแบบของข้อมูล และข้อมูลชุดที่สองไว้สำหรับทดสอบโครงข่าย โดยผลลัพธ์ที่แตกต่างระหว่างค่าจริง และค่าที่ได้จากชุดทดสอบจะถูกคำนวณออกมาเป็นค่าผิดพลาดของระบบ (system error) ซึ่งค่าผิดพลาดของระบบที่น้อย จะแสดงถึงความสามารถในการทำนายที่สูง2) ใช้ชุดข้อมูลทั้งหมดเป็นทั้งชุดการสอน และชุดทดสอบ โดยนำชุดข้อมูลทั้งหมดมาสอนโครงข่ายก่อน และหลังจากนั้นนำชุดข้อมูลชุดเดิมมาทดสอบโครงข่าย

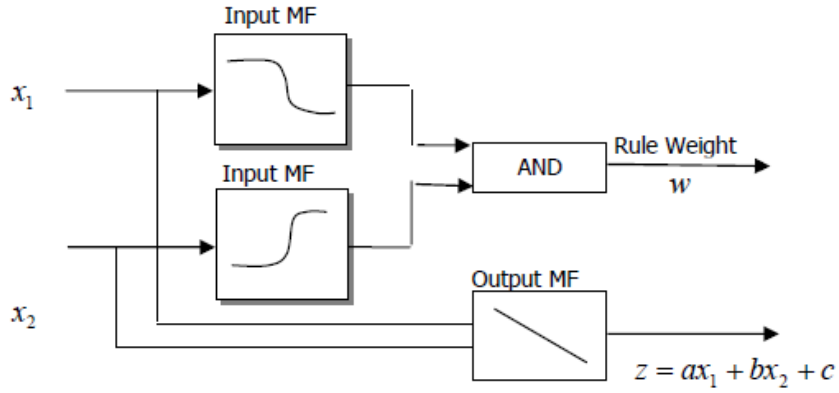
2.4 โครงข่ายประสาทเทียมฟัซซี่ (Adaptive Neuro-Fuzzy Inference System : ANFIS)

ANFIS เป็นโมเดลหนึ่งของฟัซซีที่ได้รับการอ้างถึงและนำไปประยุกต์ใช้งานหลากหลาย การทำงานภายในของ ANFIS เป็นการประยุกต์ผสมผสานทั้งวิธีการโครงข่ายประสาทเทียมและฟัซซีเข้าด้วยกัน ทำให้สามารถนำข้อดีในแต่ละวิธีการมาสนับสนุนกันและช่วยลดข้อจำกัดของแต่ละวิธีการ การปรับตัวเรียนรู้รูปแบบของข้อมูลนับว่าเป็นข้อเด่นของโครงข่ายประสาทเทียม แต่ด้วยข้อจำกัดในเรื่องการอธิบายการปรับเรียนรู้ภายในที่ยากต่อการสื่อเพื่อทำความเข้าใจเมื่อเทียบกับในมุมมองของมนุษย์ทั่วไปหรือในภาษาสื่อทั่วไป ขณะที่พื้นฐานของโมเดลฟัซซีซึ่งมีการพัฒนาจากการตัดสินใจแบบตรรกะเชิงทวินัย (Crisp Logic) ด้วยการเปรียบเทียบกฎแบบแบบถ้า-แล้ว(if-then) มาเป็นการตัดสินใจแบบคลุมเครือหรือตรรกะแบบฟัซซี (Fuzzy Logic)

โมเดลฟัซซีพื้นฐานที่ได้รับการอ้างถึง คือ โมเดลฟัซซีแมนดานิ (Mandani Fuzzy) และโมเดลฟัซซีซูเกโน (Sugeno Fuzzy) ทั้งสองวิธีการมีข้อแตกต่างหลักที่ผลค่าเอาต์พุตฟังก์ชัน ความเป็นสมาชิก (Output Membership Function) ซึ่งวิธีหลังสามารถเลือกได้ทั้งฟังก์ชันเชิงเส้นหรือเลือกเป็นค่าคงที่ได้ กรณีโมเดล ANFIS ซึ่งได้รับการพัฒนามาจากโมเดลฟัซซีซูเกโน ดังนั้นในบทนี้จึงเน้นอธิบายเฉพาะวิธีที่สอง มีรายละเอียดในหัวข้อที่ 2.4.1 และอธิบายการทำงานภายในโมเดล ANFIS ในหัวข้อที่ 2.4.2 ส่วนหัวข้อที่ 2.4.3 ขยายผลการนำ ANFIS ไปใช้งาน โดยเริ่มต้นที่การฝึกเรียนรู้ก่อนใช้งาน และการนำไปทดสอบและใช้งาน ตามลำดับ

2.4.1 ฟัซซีโมเดลฟัซซีซูเกโน (Sugeno Fuzzy)

โครงสร้างของฟัซซีซูเกโนหรือฟัซซีทีเอส (TS: Takagi-Sugeno Fuzzy) แสดงไว้ดังรูปที่ 2.6 ตัวอย่างจำนวนข้อมูลอินพุตขนาด 2 มิติ (Dimension) นำมาประมวลผลในส่วนของฟังก์ชันความเป็นสมาชิกด้านอินพุต (Input MF) ผลที่ได้นำมาเพื่อใช้ประกอบการกำหนดค่าน้ำหนักของกฎ (Rule Weight หรือ Firing Strength) สำหรับใช้เป็นค่าตัวแปรของการประมวลผลที่เอาต์พุต และสามารถแสดงความสัมพันธ์ในสมการ (2.1)



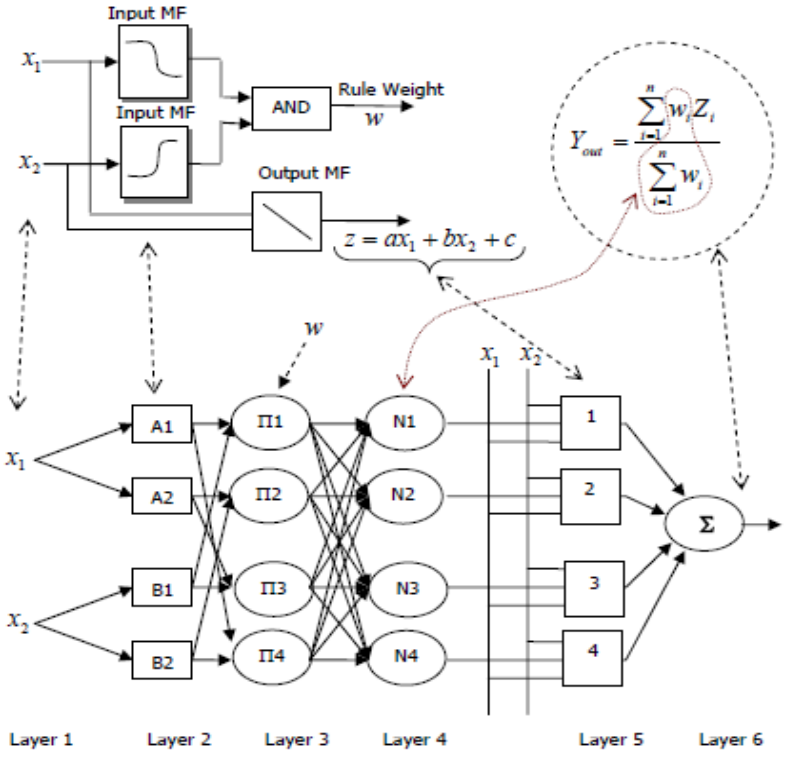
รูปที่ 2.6 โครงสร้างของฟัซซีโมเดลฟัซซีซูเกโน

$$Y_{out} = \frac{\sum_{i=1}^{n} w_i Z_i}{\sum_{i=1}^{n} w_i} \qquad (2.1)$$

โดยที่ n แทนจำนวนกฎของฟัซซี ขณะที่ w ได้จากการจับเทียบค่าของกฎฟัซซีที่ได้จากฟังก์ชันความเป็นสมาชิกด้านอินพุตตามที่ผู้เชี่ยวชาญกำหนดขึ้น ซึ่งสามารถเลือกฟังก์ชันใช้ได้หลายรูปแบบ สำหรับงานวิจัยนี้เลือกใช้แบบฟังก์ชันเกาส์เซียน (Gaussian Function) เนื่องจากลักษณะของข้อมูลมีลักษณะกระจายตัวที่สอดคล้องกับฟังก์ชันเกาส์เซียนฟังก์ชันความเป็นสมาชิกด้านเอาต์พุต (Output MF) สามารถเลือกใช้ได้ 2 รูปแบบ คือ ลำดับที่ศูนย์ (Zero Order) หรือการกำหนดความเป็นสมาชิกด้านเอาต์พุตด้วยค่าคงที่กำหนดให้ค่า a = b = 0 และลำดับที่หนึ่ง (First Order) หรือกำหนดแบบสมการเชิงเส้น $z = ax_1 + bx_2 + c$ สำหรับกรณีนำไปประยุกต์ใช้ใน ANFIS เลือกใช้แบบหลัง โดยแสดงความสัมพันธ์ระหว่างโมเดลฟัซซีซูเกโนและ ANFIS ในหัวข้อถัดไป

2.4.2 โครงสร้างของ ANFIS

เพื่อให้สามารถอธิบายขั้นตอนการฝึกเรียนรู้ก่อนใช้งานได้ พิจารณาที่ฟัซซีซูเกโนสามารถกระจายรูปแบบของโครงสร้างออกเป็น 5 ชั้น ดังความสัมพันธ์แสดงในรูปที่ 2.7 โดยสมการ (2.1) สามารถนำมาแยกเขียนใหม่ตามแต่ละชั้นของ ANFIS ผลตาม (2.2) ถึง (2.7)



รูปที่ 2.7 ความสัมพันธ์ระหว่างโมเดลฟัซซีทีเอส และ ANFIS

ชั้นที่ 1 หรือชั้นอินพุต กำหนดให้ข้อมูลอินพุตแทนด้วย x มีขนาดมิติเท่ากับ n สามารถแทนด้วยสมการ (2.2) ตัวอย่างกรณีภาพที่ 2.7 ค่า n มีค่าเท่ากับ 2

$$\mathbf{x}_{\theta} = \left[x_{1,\theta}, x_{2,\theta}, \cdots, x_{i,\theta}, \cdots, x_{n,\theta}\right]' \qquad (2.2)$$

โดยที่ $1 \le i \le n$ และ θ แทนหมายเลขระเบียนของข้อมูลชุดทดสอบ

ชั้นที่ 2 หรือชั้นการทำค่าฟัซซี (Fuzzification Layer) กรณีงานวิจัยเลือกใช้ฟังก์ชันความเป็นสมาชิกแบบเกาส์เซียน ดังนั้นผลที่ได้สามารถแสดงในสมการ (2.3)

$$y_{i,j}^{(\mu)} = e^{-\frac{\left(x_{i,\theta} - c_{i,j}\right)}{2\sigma_{i,j}^2}} \qquad (2.3)$$

โดยที่ (μ) แทนสัญลักษณ์ของเอาต์พุตในชั้นที่ 2 $1 \leqslant i \leqslant n$ $1 \leqslant j \leqslant R$ n แทนขนาดมิติของอินพุต R แทนจำนวนกฎของฟัซซี c และ σ แทนค่าจุดกึ่งกลางและความกว้างของฟังก์ชันความเป็นสมาชิกเกาส์เซียนที่ได้จากการกำหนดค่าเริ่มต้นด้วยฟังก์ชันการจัดกลุ่มแบบลบออก (Subtractive Clustering Function)

ชั้นที่ 3 หรือชั้นกฎของฟัซซี (Fuzzy Rule Layer) ในชั้นนี้นำผลที่ได้จากชั้นที่ 2 มารวมกันด้วยกฎของฟัซซีที่กำหนดขึ้นตามสมการ (2.4) และสามารถเทียบได้กับค่าน้ำหนักของกฎ w ในโมเดลฟัซซีทีเอส โดยที่ (Ru) แทนสัญลักษณ์ของเอาต์พุตในชั้นที่ 3

$$y_j^{(\mathrm{Ru})} = \prod_{i=1}^{n} \left(y_{i,j}^{(\mu)} \right) \qquad (2.4)$$

ชั้นที่ 4 หรือชั้นนอลมัลไลเซชัน (Normalization Layer) แสดงผลในสมการ (2.5) โดยที่ (μ)แทนสัญลักษณ์ของเอาต์พุตในชั้นที่ 4

$$y_j^{(\bar{\mu})} = \frac{y_j^{(\mathrm{Ru})}}{\sum_{j=1}^{R} y_j^{(\mathrm{Ru})}} \qquad (2.5)$$

ชั้นที่ 5 หรือชั้นการทำค่าฟัซซีให้เป็นค่าปกติ (Defuzzification Layer) แสดงไว้ใน (2.6) ชั้นนี้เทียบได้กับฟังก์ชันความเป็นสมาชิกด้านเอาต์พุตของฟัซซีซูเกโน

$$y_j^{(\mathrm{Df})} = y_j^{(\bar{\mu})} \cdot \left(\sum_{i=1}^{n} \left(k_{j,i} x_{i,\theta} \right) + k_{j,(n+1)} \right) \qquad (2.6)$$

โดยที่ (Df) แทนสัญลักษณ์ของเอาต์พุตในชั้นที่ 5 k แทนค่าพารามิเตอร์ที่ได้จากการแก้สมการในส่วนของการเรียนรู้ไปข้างหน้าด้วยวิธีการแก้ปัญหาเมตริกซ์แบบซูโดอินเวิร์ส (Moore-Penrose Pseudo Inverse of a Matrix) ผลที่ได้ค่า k มีขนาดเมตริกซ์เท่ากับ [R × (n +1)]

ชั้นที่ 6 ชั้นผลรวมนิวรอน (Summation Neuron) หรือ เอาต์พุตของ ANFIS ตามสมการ (2.7) โดยผลที่ได้เท่ากับสมการ (2.1)

$$\mathrm{TS}\left(x_\theta, \{k, c, \sigma\}\right) = y_\theta^{(\mathrm{TS})} = \sum_{j}^{R} y_j^{(\mathrm{Df})} \qquad (2.7)$$

โดยที่ TS (x_θ $\{k,c,\sigma\}$) แทนฟังก์ชันฟัซซีซูเกโน สามารถประมวลผลได้เมื่อมีการกำหนดค่าอินพุต x_θ และค่าพารามิเตอร์ $\{k,c,\sigma\}$

2.5 งานวิจัยที่เกี่ยวข้อง

จากผลการศึกษา Sundarambal.P et al. (2008) ได้พัฒนาการสร้างแบบจำลองโดยใช้ระบบโครงข่ายใยประสาทประดิษฐ์สำหรับทำนายปริมาณออกซิเจนในสัปดาห์ถัดไปในน้ำทะเลประเทศสิงค์โปร์ และ Prybutok et al. (2000) ได้พยายามที่จะทำนายค่าระดับความเข้มข้นสูงสุดของโอโซนของแต่ละวัน บริเวณเขตอุตสาหกรรมทางตอนเหนือของสหรัฐอเมริกา (Houston site) และได้กล่าวไว้ว่าเป็นการยากที่จะสร้างแบบจำลองให้แม่นยำได้ เนื่องมาจากตัวแปรทางสิ่งแวดล้อมมีความซับซ้อนมาก ซึ่งในการศึกษาครั้งนี้ได้ทำกาสร้างแบบจำลอง Neural Network เพื่อใช้ในการทำนายค่าระดับความเข้มข้นสูงสุดของโอโซนแต่ละวันโดยทำการศึกษาเปรียบเทียบกับแบบจำลองทางสถิติทั่วไป คือ regression และ Box – Jenkins ARIMA ซึ่งข้อมูลที่ใช้ได้จาก Texas Natural Resource Conservation Commission ประกอบไปด้วยค่าความเข้มข้นเฉลี่ยรายชั่วโมงของโอโซน และ ค่าเฉลี่ยรายชั่วโมงของข้อมูลทางอุตุอนิยมวิทยา เช่น อุณหภูมิ ความเร็วลม และทิศทางลม รวมทั้งข้อมูลมลพิษทางอากาศอื่น ๆ เช่น NO NO2 CO2 และ NOx จากสถานีตรวจวัดอากาศ โดยข้อมูลนี้จะเลือกใช้ในช่วงเดือนที่มีความเข้มข้นของโอโซนสูงสุด คือ ตั้งแต่วันที่ 1 มิถุนายน ถึง 31 ตุลาคม เนื่องจากต้องการลดความผิดพลาดในการทำนายจากปัจจัยทางด้านฤดูกาล โดยที่ข้อมูลในช่วงวันที่ 1 มิถุนายน ถึง 30 กันยายน จะใช้ในการสร้างแบบจำลอง และได้ให้ความเห็นไว้ว่าค่าเฉลี่ยของข้อมูลตัวแปรอิสระที่จะทำให้แบบจำลองสามารถทำนายได้ดีควรจะอยู่ในช่วงเวลา 6.00 – 9.00 น. ส่วนข้อมูลในช่วงวันที่ 1 – 10 ตุลาคม จะใช้เป็นตัวทดสอบแบบจำลอง โดยใช้โปรแกรม SAS ในการสร้างแบบจำลอง Regression และ ARIMA และใช้โปรแกรม NeuralWare III ในการสร้างแบบจำลอง Neural Networks หลังจากที่ทำการเลือกตัวแปรอิสระแล้ว ต้องทำการพิจารณาปัจจัยที่ส่งผลกระทบต่อการเกิดโอโซนด้วย คือ photochemical production และ atmospheric accumulation (Robeson and Steyn, 1990) รวมทั้งปัจจัยที่สำคัญอีกประการหนึ่ง คือ มลพิษที่ปลดปล่อยออกมาจากยวดยานพาหนะบนท้องถนน คือ NO NO2 CO2และ NOx หลังจากทำการสร้างแบบจำลองทั้งสามแล้วพบว่าค่า Mean absolute deviations (MAD)ของแบบจำลอง Regression ARIMA และ ANNs มีค่า 0.025741, 0.02879 และ 0.012945 ตามลำดับและค่า RMSE มีค่า 0.031239, 0.033023 และ 0.016418 ตามลำดับ