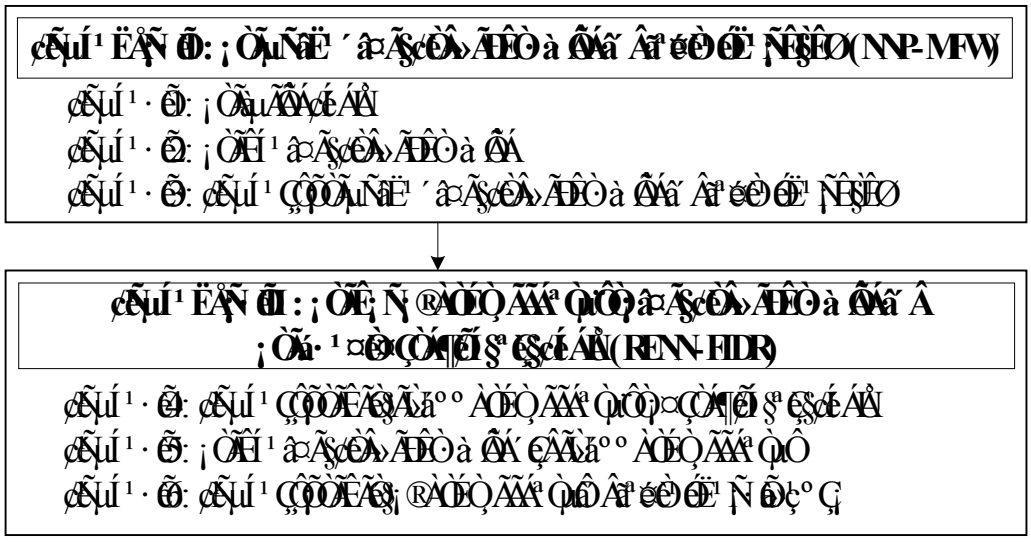
บทที่ 3

แบบจำลองการสกัดกฎภาษาธรรมชาติโดยใช้โครงข่ายประสาทเทียม

วิทยานิพนธ์นี้ใช้เทคนิคโครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้น มาสร้างแบบจำลองการสกัดกฎภาษาธรรมชาติ ที่มีชื่อว่าแบบจำลองการสกัดกฎภาษาธรรมชาติจากโครงข่ายประสาทเทียมที่ผ่านการตัดโหนดโดยใช้การแทนค่าความถี่ของช่วงข้อมูล (Rule Extraction form Neural Networks Pruning using Frequency Interval Data Representation: RENN-P-FIDR) [37] มีการทำงาน 2 ขั้นตอนหลักแสดงได้ดังภาพประกอบ 3.1 ขั้นตอนหลักที่ I การตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด (Neural Network Pruning using Maximum Feature Weighting: NNP-MFW) เป็นการเตรียมข้อมูล (Data Preprocessing) ก่อนที่จะสกัดกฎภาษาธรรมชาติ และขั้นตอนหลักที่ II การสกัดกฎภาษาธรรมชาติจากโครงข่ายประสาทเทียมโดยการแทนค่าความถี่ของช่วงข้อมูล (Rule Extraction form Neural Networks using Frequency Interval Data Representation: RENN-FIDR) [38] รายละเอียดของแต่ละขั้นตอนมีดังนี้



ภาพประกอบ 3.1 แบบจำลองการสกัดกฎภาษาธรรมชาติ RENN-P-FIDR

3.1 การตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด (NNP-MFW)

การตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด (Neural Network Pruning using Maximum Feature Weighting: NNP-MFW) เป็นการเตรียมข้อมูลก่อนที่จะสกัดกฎ

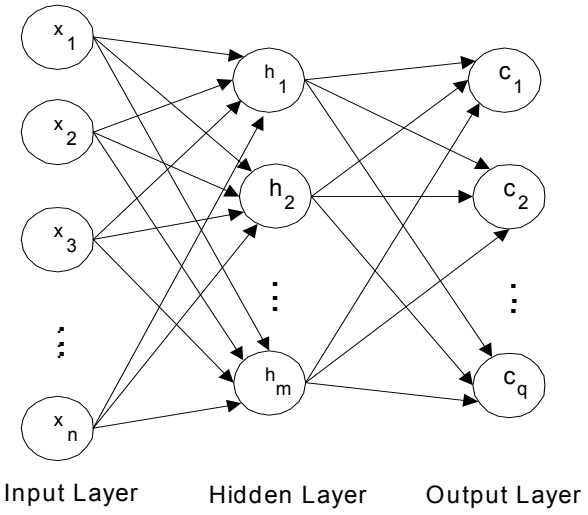
ภาษาธรรมชาติ เนื่องจากต้องการสกัดกฎภาษาธรรมชาติที่มีค่าความถูกต้องสูง มีจำนวนกฎ มีตัวเชื่อมของเงื่อนไข และมีเวลาที่ใช้ในการสกัดกฎน้อย ดังนั้นจึงต้องมีการตัดโหนดโครงข่ายประสาทเทียมก่อนเข้าสู่ขั้นตอนการสกัดกฎภาษาธรรมชาติจากโครงข่ายประสาทเทียมโดย การแทนค่าความถี่ของช่วงข้อมูล การตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุดมีการทำงาน 3 ขั้นตอนคือ ขั้นตอนที่ 1 การเตรียมข้อมูล ขั้นตอนที่ 2 การสอนโครงข่ายประสาทเทียมด้วยการเรียนรู้แบบแพร่ย้อนกลับ และขั้นตอนที่ 3 ขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุดดังภาพประกอบ 3.1 ในแต่ละขั้นตอนมีรายละเอียดดังนี้

ขั้นตอนที่ 1 การเตรียมข้อมูล

เนื่องจากฐานข้อมูลทางการแพทย์ที่นำสกัดกฎภาษาธรรมชาติ อาจมีบางแถว ข้อมูลมีตัวแปรข้อมูลเข้าหายไป (Missing Values) ดังนั้นจึงต้องแทนค่าตัวแปรข้อมูลเข้าที่หายไปด้วยค่าเฉลี่ยของกลุ่ม (Class Mean)

ขั้นตอนที่ 2 การสอนโครงข่ายประสาทเทียมโดยใช้การเรียนรู้แบบแพร่ ย้อนกลับ

กำหนดให้สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบเพอร์เซฟตรอนหลายชั้นมีโครงสร้างดังนี้คือ ชั้นข้อมูลเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นผลลัพธ์ (Output Layer) ดังภาพประกอบ 3.2 โดยจะทำการสอนโครงข่ายประสาทเทียมจนกระทั่งครบจำนวนรอบที่ตั้งไว้หรือได้ค่าความผิดพลาด (MSE) น้อยกว่าค่าที่ยอมรับได้



โดยที่ x คือโหนดข้อมูลเข้า n โหนด

h คือโหนดซ่อน m โหนด

c คือโหนดผลลัพธ์ q โหนด

ภาพประกอบ 3.2 โครงสร้างของโครงข่ายประสาทเทียมสำหรับขั้นตอนหลักที่ I NNP-MFW

ขั้นตอนที่ 3 ขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด

ขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุดจะพิจารณาจากค่าน้ำหนักที่ยอมรับได้ (accept_w) ของแต่ละชั้นของโครงข่ายประสาทเทียมตั้งแต่ระหว่างชั้นผลลัพธ์ (Output Layer) มายังชั้นซ่อน (Hidden Layer) และระหว่างชั้นซ่อนมายังชั้นข้อมูลเข้า (Input Layer) ย้อนกลับมา เนื่องจากเส้นเชื่อมที่มีน้ำหนักมากจะมีความสำคัญในการแบ่งกลุ่มมากกว่าเส้นเชื่อมที่มีน้ำหนักน้อยกว่า [21] กล่าวคือระบบจะทำการหาค่าน้ำหนักสูงสุด (max_w) ของเส้นเชื่อมและทำการคำนวณค่าน้ำหนักที่ยอมรับได้ (accept_w) จากร้อยละน้ำหนัก (□) ของการตัดโหนดที่ผู้ใช้กำหนด โดยที่ □ มากจะตัดโหนดได้มาก □ น้อยจะตัดโหนดได้น้อย ทั้งนี้ระบบจะตัดโหนดโครงข่ายประสาทเทียมก็ต่อเมื่อ □ มีค่าระหว่าง (0,100] นั่นคือถ้า □ = 0 จะหมายถึงไม่มีการตัดโหนดโครงข่ายประสาทเทียม ค่าน้ำหนักที่ยอมรับได้สามารถนิยามได้ดังนี้

นิยามที่ 1 ค่าน้ำหนักที่ยอมรับได้ (accept_w) หมายถึงกับค่าน้ำหนักสูงสุด (max_w) ของเส้นเชื่อมน้ำหนักคูณด้วยร้อยละน้ำหนัก (□) หารด้วย 100 ดังสมการที่ (3.1)

$$accept_w = \frac{max_w \times \square}{10} \qquad (3.1)$$

ในการตัดโหนดโครงข่ายประสาทเทียมจะพิจารณาหาโหนดซ่อนที่มีผลต่อ การทำนายกลุ่มที่ 1 โดยพิจารณาจากค่าน้ำหนักของเส้นเชื่อม (w) ที่มีค่ามากกว่าหรือเท่ากับ ค่าน้ำหนักที่ยอมรับได้ จากนั้นในทำนองเดียวกันจะพิจารณาย้อนกลับระหว่างชั้นซ่อนไปยังชั้น ข้อมูลเข้าว่าโหนดข้อมูลเข้าใดมีผลต่อโหนดซ่อนนั้น (หมายเหตุ โหนดซ่อนและโหนดข้อมูลเข้าที่ไม่ได้รับการยอมรับจะถูกตัดออกจากโครงข่ายประสาทเทียม) ระบบจะทำเช่นนี้ไปเรื่อยๆในทุกกลุ่มข้อมูล จากขั้นตอนวิธีดังกล่าวผลลัพธ์ที่ได้จะทำให้สามารถกำหนดจำนวนโหนดซ่อนที่ เหมาะสมและกำหนดค่าโหนดข้อมูลเข้าที่จำเป็นในการแบ่งกลุ่มข้อมูล เพื่อใช้ในการสกัดกฎภาษาธรรมชาติ

ในลำดับถัดไป รายละเอียดของขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมแสดงได้ดังภาพประกอบ 3.3

ขั้นตอนที่ 3: ขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด

3.1 กำหนดให้ □ คือร้อยละน้ำหนักของการตัดโหนดที่ผู้ใช้กำหนด

3.2 ถ้า □ = 0

แล้ว ให้ยอมรับโหนดข้อมูลเข้า x และโหนดซ่อน h ทุกโหนดและทำขั้นตอนย่อยที่ 3.4

3.3 จากโหนดผลลัพธ์ที่ q ในแต่ละกลุ่ม

3.3.1 พิจารณาเส้นเชื่อมระหว่างชั้นข้อมูลออก (Output Layer) และชั้นซ่อน (Hidden Layer)

1) หาค่า max_w ค่าน้ำหนักสูงสุดของเส้นเชื่อมระหว่างชั้นผลลัพธ์และชั้นซ่อน

2) คำนวณค่า accept_w = (max_w x □) /100

3) สำหรับแต่ละเส้นเชื่อม

ถ้า $w \geq$ accept_w

แล้ว ให้ยอมรับโหนดซ่อน h ของเส้นเชื่อมนั้น

3.3.2 จากโหนดซ่อนที่ h ในแต่ละโหนดที่ถูกยอมรับ

1) พิจารณาเส้นเชื่อมระหว่างชั้นซ่อน (Hidden Layer) และชั้นข้อมูลเข้า (Input Layer)

1.1) หาค่า max_w′ ค่าน้ำหนักสูงสุดของเส้นเชื่อมระหว่างชั้นซ่อนและชั้นข้อมูลเข้า

1.2) คำนวณค่า accept_w′ = (max_w′ x □) /100

1.3) สำหรับแต่ละเส้นเชื่อม

ถ้า $w' \geq$ accept_w′

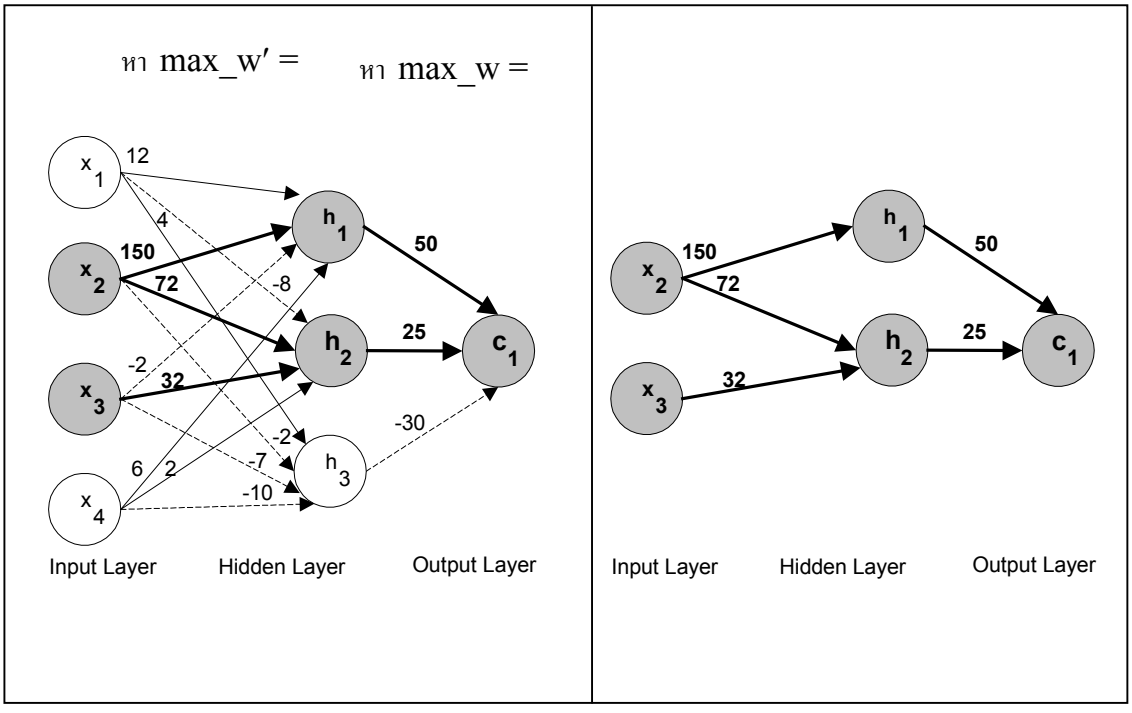
แล้ว ให้ยอมรับโหนดข้อมูลเข้า x ของเส้นเชื่อมนั้น

3.4 จบการทำงาน

ภาพประกอบ 3.3 ขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด

ตัวอย่างในการตัดโหนดโครงข่ายประสาทเทียมเมื่อกำหนดค่า □ = 10 โดยพิจารณาค่าน้ำหนักของเส้นเชื่อมระหว่างชั้นต่างๆ ดังภาพประกอบ 3.4 (a) จะเห็นได้ว่า ค่าน้ำหนักที่สูงสุด (max_w) ระหว่างชั้นผลลัพธ์และชั้นซ่อนพบว่ามีค่าเท่ากับ 50 ดังนั้นสามารถคำนวณค่าน้ำหนักที่ยอมรับได้ (accept_w) เท่ากับ (50 x 10)/100 หรือมีค่าเท่ากับ 5 หมายถึงเราจะยอมรับน้ำหนัก (w) ที่มีค่ามากกว่าหรือเท่ากับ 5 ซึ่งก็คือเส้นเชื่อมน้ำหนักระหว่าง โหนดผลลัพธ์ที่ 1 ไปโหนดซ่อนที่ 1 และโหนดผลลัพธ์ที่ 1 ไปโหนดซ่อนที่ 2 ดังนั้นจึงยอมรับโหนดซ่อนที่ 1 และโหนดซ่อนที่ 2 ส่วนโหนดซ่อนที่ 3 ถูกตัดออกจากโครงข่ายประสาทเทียม

จากนั้นพิจารณาค่าน้ำหนักที่สูงสุด (max_w′) ระหว่างชั้นซ่อนและชั้นเข้ามูลเข้าของโหนดซ่อนที่ 1 พบว่ามีค่าเท่ากับ 150 ดังนั้นค่าน้ำหนักที่ยอมรับได้ (accept_w′) มีค่าเท่ากับ (150 x 10)/100 หรือ15 โหนดข้อมูลเข้าที่มีน้ำหนัก (w′) ของเส้นเชื่อมมากกว่า 15 คือโหนดข้อมูลเข้าที่ 2 ดังนั้นจึงยอมรับโหนดข้อมูลเข้าที่ 2 ในทำนองเดียวกันพิจารณาน้ำหนักของโหนดซ่อนที่ 2 จะได้ว่ายอมรับโหนดข้อมูลเข้าที่ 3 เพราะฉะนั้นจึงยอมรับโหนดข้อมูลเข้าที่ 2 และ 3 ส่วนโหนดข้อมูลเข้าที่ 1 และ 4 ถูกตัดออกจากโครงข่ายประสาทเทียมดังภาพประกอบ 3.4(b)



โดยที่ accept_w′ ของ h_1 = 15, accept_w ของ c_1 = 5

accept_w′ ของ h_2 = 7

(a) ก่อนการตัดโหนด (b) หลังการตัดโหนด

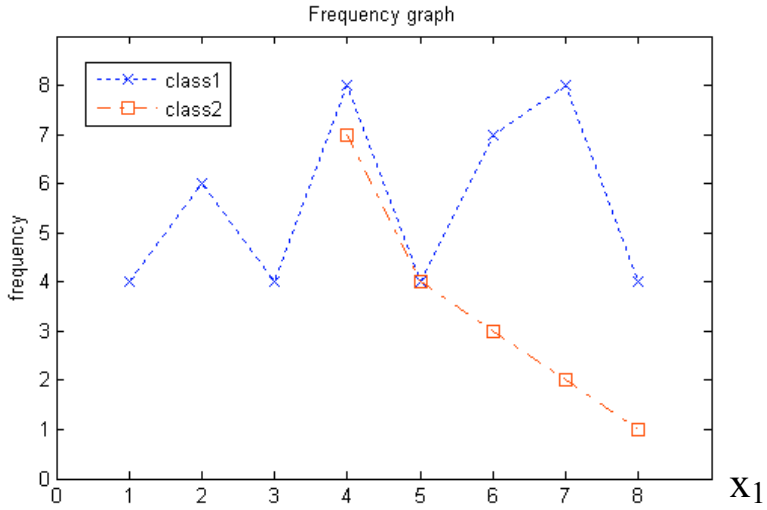
ภาพประกอบ 3.4 ตัวอย่างการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด

3.2 การสกัดกฎภาษาธรรมชาติจากโครงข่ายประสาทเทียมโดยการแทนค่าความถี่ของช่วงข้อมูล (RENN-FIDR)

การสกัดกฎภาษาธรรมชาติจากโครงข่ายประสาทเทียมโดยการแทนค่าความถี่ของช่วงข้อมูล: RENN-FIDR ซึ่งเป็นขั้นตอนหลักที่ II จะพิจารณาเฉพาะโหนดซ่อน และโหนดข้อมูลเข้าที่เสนอแนะจากขั้นตอนหลักที่ I (การตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด: NNP-MFW) ในขั้นตอนหลักที่ II นี้มีอีก 3 ขั้นตอนคือ ขั้นตอนที่ 4 ขั้นตอนวิธีการสร้างรูปแบบภาษาธรรมชาติจากความถี่ของช่วงข้อมูล ขั้นตอนที่ 5 การสอนโครงข่ายประสาทเทียมด้วยด้วยรูปแบบภาษาธรรมชาติ และขั้นตอนที่ 6 ขั้นตอนวิธีการสร้างกฎภาษาธรรมชาติโดยใช้ ค่าน้ำหนักที่เป็นบวก ดังภาพประกอบ 3.1 ซึ่งรายละเอียดในแต่ละขั้นตอนมีดังนี้

ขั้นตอนที่ 4 ขั้นตอนวิธีการสร้างรูปแบบภาษาธรรมชาติจากความถี่ของช่วงข้อมูล

เนื่องจากข้อมูลมีลักษณะเป็นตัวเลขจำนวนจริง และเราต้องการที่จะแทนด้วยภาษาธรรมชาติ เช่น เล็ก กลาง และใหญ่ ดังนั้นจึงต้องมีการเตรียมข้อมูลก่อนการสอนโครงข่ายประสาทเทียม โดยการสร้างรูปแบบภาษาธรรมชาติจากความถี่ของช่วงข้อมูล เราจะพิจารณา ค่าความถี่ของข้อมูลในแต่ละตัวแปรข้อมูลเข้า (Input Attribute) และพิจารณาค่าความกำกวมของข้อมูลนั้นว่าเป็นสมาชิกของกลุ่มใดมากที่สุดเพื่อใช้แทนค่าข้อมูลตัวอย่างเช่น ภาพประกอบ 3.5 แสดงกราฟความถี่ของข้อมูล ที่มีตัวแปรข้อมูลเข้าคือ x_1 ที่มีค่าอยู่ระหว่าง 1 ถึง 8 และมีข้อมูลอยู่ ทั้งหมด 2 กลุ่ม โดยที่เครื่องหมาย x แทนกลุ่มที่ 1 และเครื่องหมาย □ แทนกลุ่มที่ 2



ภาพประกอบ 3.5 กราฟแสดงความถี่ของข้อมูลกลุ่ม 1 และกลุ่ม 2

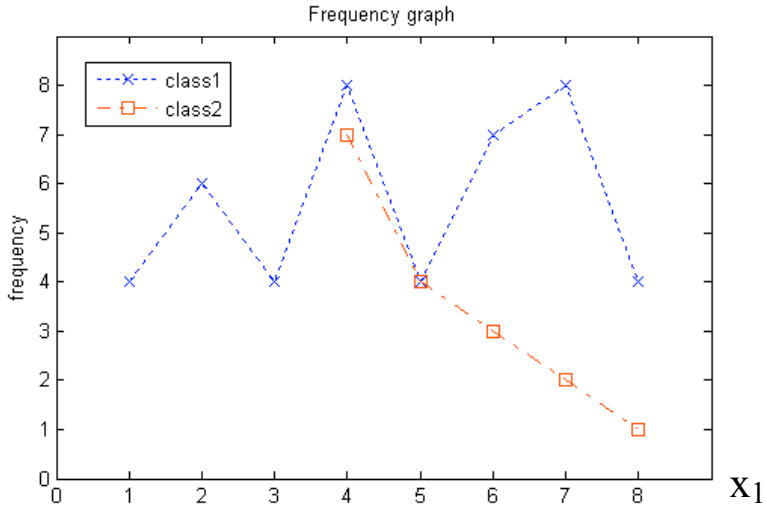
ในการพิจารณาค่าความกำกวมในแต่ละจุดของตัวแปรข้อมูลเข้า (ค่าความกำกวมของจุด) สามารถนิยามได้ดังต่อไปนี้คือ

นิยามที่ 2 ค่าความกำกวมของจุด (ambiguity) หมายถึงค่าความถี่ข้อมูลสูงสุดรองลงมา ($freq_{2nd_max}$) ต่อค่าความถี่สูงสุด ($freq_{max}$) คูณด้วย 100 ดังสมการที่ (3.2)

$$\text{ambiguity} = \frac{freq_{2nd}}{freq_{max}} \times 100 \qquad (3.2)$$

ค่าความกำกวมของจุดมีไว้เพื่อพิจารณาว่าตัวแปรที่มีความถี่สูงสุด ($freq_{max}$) และตัวแปรที่มีความถี่สูงสุดรองลงมา ($freq_{2nd_max}$) มีสัดส่วนที่สัมพันธ์กันอย่างไร สาเหตุที่ต้องพิจารณาค่าความถี่สูงสุดรองลงมา ($freq_{2nd_max}$) ของกลุ่มหารความถี่สูงสุด ($freq_{max}$) ของอีกกลุ่มหนึ่งคูณด้วย 100 เนื่องจากต้องการค่าความกำกวมของจุดที่มีช่วงอยู่ระหว่างค่า 0 ถึง 100

ภาพประกอบ 3.6 แสดงค่าความกำกวมของจุดจากกราฟความถี่ของข้อมูลจากภาพประกอบ 3.5 ตัวอย่างในการคำนวณหาค่าความกำกวมเช่นจากภาพประกอบ 3.6 ที่จุด x_1 มีค่าเท่ากับ 6 จะได้ค่าความถี่สูงสุดคือกลุ่มที่ 1 มีค่าเท่ากับ 7 และค่าความถี่สูงสุดรองลงมาคือกลุ่มที่ 2 มีค่าเท่ากับ 3 ดังนั้นค่าความความกำกวมของจุด x_1 เท่ากับ 6 มีค่าเท่ากับ (3/7)×100 หรือ 43 เป็นต้น



ค่าความกำกวมของจุด 0 0 0 88 100 45 25 32

ภาพประกอบ 3.6 ค่าความกำกวมของจุดในแต่ละจุดของตัวแปรข้อมูลเข้า

ถ้ากำหนดให้ ϕ คือค่าความกำกวมของจุดที่ยอมรับได้ (Ambiguity Threshold Value) สามารถแบ่งค่าความกำกวมของจุดได้เป็น 3 กรณี ซึ่งจาก 3 กรณีนี้จะได้ช่วงต่างๆ และจะต้องคำนวณหาค่าความกำกวมของช่วงที่ได้ โดยค่าความกำกวมของช่วงสามารถนิยามได้ดังต่อไปนี้ คือ

นิยามที่ 3 ค่าความกำกวมของช่วง (ambiguity_interval) หมายถึงค่าผลรวมความถี่ข้อมูลสูงสุดรองลงมาทุกจุดของตัวแปรข้อมูลเข้าที่อยู่ในช่วง ($sum_freq_{2nd_max}$) ต่อค่าผลรวมความถี่สูงสุดทุกจุดของตัวแปรข้อมูลเข้าที่อยู่ในช่วง (sum_freq_{max}) คูณด้วย 100 ดังสมการที่ (3.3)

$$ambiguity_interval = \frac{sum_freq_{2nd}}{sum_fre} \times 100 \qquad (3.3)$$

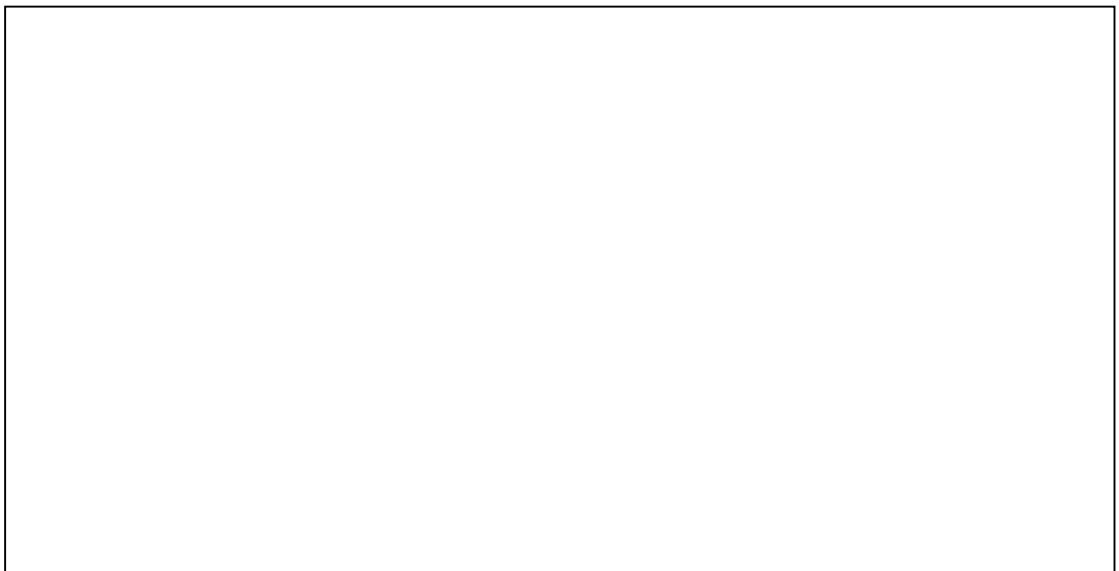
ในการคำนวณหาค่าความกำกวมของช่วงจะต้องคำนวณหาผลรวมของความถี่ใหม่ เพราะช่วงที่ได้อาจเกิดจากการรวมช่วงซึ่งจะกล่าวในขั้นตอนวิธีในการหาช่วงเพื่อใช้ใน การแทนข้อมูลจากค่าความกำกวมของความถี่ของช่วงข้อมูล โดยช่วงที่นำมารวมกันเป็นช่วงใหม่อาจมีกลุ่มข้อมูลที่มีความถี่สูงสุดต่างกลุ่มข้อมูลกัน ดังนั้นในการคำนวณหาค่าความกำกวมของช่วงใหม่จะไม่สามารถนำค่าความกำกวมของช่วงในแต่ละช่วงมารวมกันแล้วหารจำนวนช่วงได้

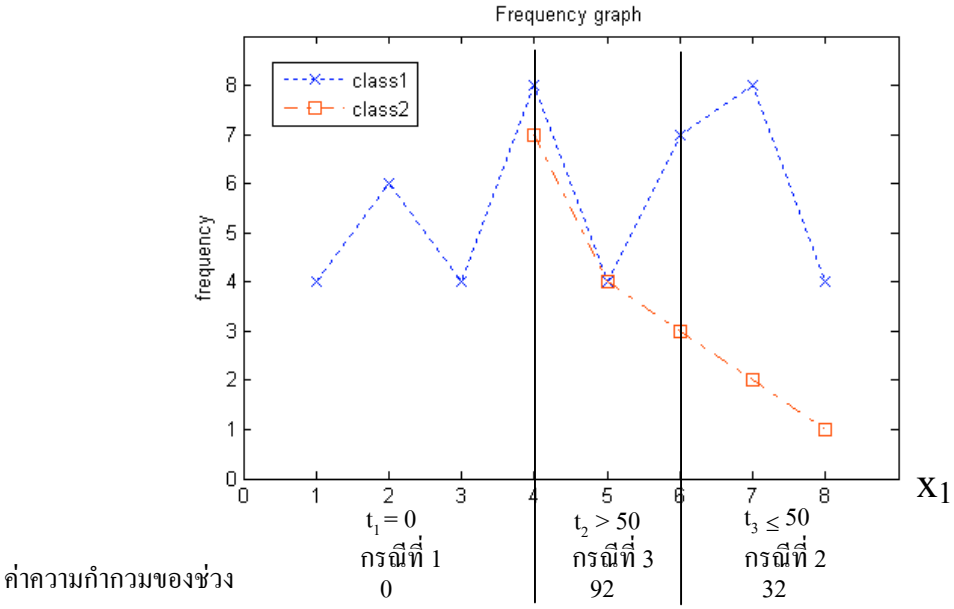
จากภาพประกอบ 3.6 ถ้ากำหนดให้ ϕ คือค่าความกำกวมของจุดที่ยอมรับได้ (Ambiguity Threshold Value) มีค่าเท่ากับ 50 สามารถแบ่งค่าความกำกวมของจุดได้เป็น 3 กรณี ดังนี้

กรณีที่ 1 ค่าความกำกวมของจุดมีค่าเป็นศูนย์ (ambiguity = 0) หมายถึงไม่มีความกำกวมเลยระหว่างกลุ่มที่ 1 และกลุ่มที่ 2 ตัวอย่างเช่นจากภาพประกอบ 3.6 ที่จุด x_1 มีค่าเท่ากับ 1 จะได้ค่าความกำกวมของจุดเท่ากับ 0 ที่จุด x_1 มีค่าเท่ากับ 2 จะได้ค่าความกำกวมของจุดเท่ากับ 0 และที่จุด x_1 มีค่าเท่ากับ 3 จะได้ค่าความกำกวมของจุดเท่ากับ 0 ทั้ง 3 จุดนี้มีกลุ่มที่ 1 เป็นกลุ่มที่มีความถี่สูงสุด ดังนั้นทั้ง 3 จุดนี้จึงอยู่ในช่วงเดียวกันคือ [1,4) และคำนวณ ค่าความกำกวมของช่วงโดยใช้สมการที่ (3.3) ได้เท่ากับ (0+0+0)/(4+6+4) x 100 หรือ 0 ดังภาพประกอบ 3.7

กรณีที่ 2 ค่าความกำกวมของจุดมีค่าน้อยกว่าหรือเท่ากับค่าความกำกวมของจุดที่ยอมรับได้ (ambiguity $\leq \phi$) ตัวอย่างเช่น กำหนดให้ค่า ϕ มีค่าเท่ากับ 50 หมายถึง เราจะยอมรับค่าความกำกวมระหว่างข้อมูลที่แทนด้วยกลุ่มที่ 1 และกลุ่มที่ 2 ได้ไม่เกิน 50 จาก ภาพประกอบ 3.6 ที่จุด x_1 มีค่าเท่ากับ 6 จะได้ค่าความกำกวมของจุดเท่ากับ 43 ที่จุด x_1 มีค่า เท่ากับ 7 จะได้ค่าความกำกวมของจุดเท่ากับ 25 และที่จุด x_1 มีค่าเท่ากับ 8 จะได้ค่าความกำกวมของจุดเท่ากับ 25 ซึ่งมีค่าน้อยกว่าค่าความกำกวมของจุดที่ยอมรับได้ และมีกลุ่มที่ 1 เป็นกลุ่มที่มีความถี่สูงสุด ดังนั้นทั้ง 3 จุดนี้จึงอยู่ในช่วงเดียวกันคือ [6,8] และคำนวณค่าความกำกวมของช่วงได้เท่ากับ (3+2+1)/(7+8+4) x 100 หรือ 32 ดังภาพประกอบ 3.7

กรณีที่ 3 ค่าความกำกวมของจุดมีค่ามากกว่าค่าความกำกวมของจุดที่ยอมรับได้ (ambiguity > ϕ) ตัวอย่างเช่น กำหนดให้ค่า ϕ มีค่าเท่ากับ 50 จากภาพประกอบ 3.6 ที่จุด x_1 มีค่าเท่ากับ 4 จะได้ค่าความกำกวมของจุดเท่ากับ 88 และที่จุด x_1 มีค่าเท่ากับ 5 จะได้ ค่าความกำกวมของจุดเท่ากับ 100 ซึ่งมีค่ามากกว่าค่าความกำกวมของจุดที่ยอมรับได้ และมีกลุ่มที่ 1 เป็นกลุ่มที่มีความถี่สูงสุด ดังนั้นทั้ง 2 จุดนี้จึงอยู่ในช่วงเดียวกันคือ [4,6) และคำนวณ ค่าความกำกวมใหม่ได้เท่ากับ (7+4)/(8+4) x 100 หรือ 92 ดังภาพประกอบ 3.7





ภาพประกอบ 3.7 การแบ่งค่าความกำกวมของจุด

ขั้นตอนที่ 4 ขั้นตอนวิธีในการหาช่วงเพื่อใช้ในการแทนข้อมูลจากค่าความกำกวมของความถี่ของช่วงข้อมูลแสดงได้ดังภาพประกอบ 3.8 มีดังนี้ ขั้นตอนย่อยที่ 4.1 กำหนดให้ L คือรูปแบบของภาษาธรรมชาติที่เราต้องการ เช่นถ้า L มีค่าเท่ากับ 3 หมายถึงกำหนดให้ l_1= เล็ก l_2 = กลาง และ l_3 = ใหญ่เป็นต้น และ ϕ คือค่าความกำกวมของจุดที่ยอมรับได้ ขั้นตอนย่อยที่ 4.2 จะคำนวณหาค่าความกำกวมของจุดต่างๆของตัวแปรข้อมูลเข้าแต่ละตัว ขั้นตอนย่อยที่ 4.3 ให้ t คือช่วงที่ได้จากกรณีต่างๆที่มีค่าตัวแปรข้อมูลเข้าที่อยู่ในกรณีเดียวกันจาก 3 กรณีข้างต้นที่ติดกัน และกลุ่มที่มีความถี่สูงสุดเป็นกลุ่มเดียวกัน โดยจะต้องคำนวณหาค่าความกำกวมของช่วงในแต่ละช่วงที่แบ่งได้ ถ้าค่า L เท่ากับค่า t เราจะได้ช่วงภาษาธรรมชาติที่ต้องการตามค่า t กรณีที่จำนวนภาษาธรรมชาติ L ไม่เท่ากับค่า t จะต้องทำขั้นตอนย่อยที่ 4.4 คือมีการรวมช่วง ข้อมูล หรือแบ่งช่วงข้อมูล

กรณีที่ต้องแบ่งช่วงข้อมูลเพิ่มเติมคือ $t < L$ จะเลือกแบ่งช่วงข้อมูลที่มีค่าความกำกวมของช่วงสูงสุดมาแบ่งครึ่งช่วง เพื่อต้องการไม่ให้มีผลกระทบต่อช่วงที่มีค่าความกำกวมของช่วงน้อยในการแบ่งช่วงข้อมูล จากนั้นคำนวณหาค่าความกำกวมของช่วงใหม่ และกรณีที่ต้องรวมช่วงข้อมูลเข้าด้วยกันคือค่า $t > L$ จะเลือกรวมช่วงที่มีความกำกวมของช่วงมากที่สุดอีกเช่นกัน โดยจะคำนวณค่าความกำกวมของช่วงใหม่ระหว่างการรวมช่วงฝั่งซ้ายที่ติดกัน กับการรวมช่วงฝั่งขวาที่ติดกัน และเลือกเอาช่วงข้อมูลที่ได้ค่าความกำกวมของช่วงน้อยที่สุดมาเป็นช่วงใหม่ ในขั้นตอนย่อย

ที่ 4.4 ลำดับการทำงานจะทำซ้ำจนกระทั่งได้จำนวนภาษาธรรมชาติ L เท่ากับจำนวน ช่วงข้อมูล t

ขั้นตอนที่ 4: ขั้นตอนวิธีการสร้างรูปแบบภาษาธรรมชาติจากความถี่ของช่วงข้อมูล

4.1 กำหนดให้ L คือรูปแบบภาษาธรรมชาติที่ต้องการ

ϕ คือค่าความกำกวมที่ระบบยอมรับได้

4.2 คำนวณหาค่าความกำกวม (ambiguity) ของแต่ละจุดของตัวแปรข้อมูลเข้าทุกตัวโดยใช้สมการ (3.2)

4.3 t คือช่วงที่ได้จากกรณีต่างๆที่มีค่าตัวแปรข้อมูลเข้าอยู่ในกรณีเดียวกันที่ติดกันและกลุ่มที่มีความถี่สูงสุดเป็นกลุ่มเดียวกัน โดยแยกเป็น 3 กรณีคือ

4.3.1 ambiguity = 0

4.3.2 ambiguity $\leq \phi$

4.3.3 ambiguity > ϕ

คำนวณหาค่าความกำกวมของช่วง (ambiguity_interval) ในแต่ละช่วงโดยใช้สมการ (3.3)

4.4 กรณี $t \neq L$

4.4.1 ถ้า $t < L$ แล้ว ให้เลือกช่วงที่มีความกำกวมของช่วงสูงสุดมาแบ่งครึ่งช่วง คำนวณหาค่าความกำกวมของช่วงใหม่ (ambiguity_interval) โดยใช้สมการ (3.3)

4.4.2 ถ้า $t > L$ แล้ว ให้เลือกช่วงที่มีความกำกวมของช่วงสูงสุดมารวมกับช่วงที่ใกล้เคียงที่รวมแล้วได้ช่วงที่มีค่าความกำกวมของช่วงน้อยกว่า โดยคำนวณหาค่าความกำกวมของช่วงใหม่ (ambiguity_interval) โดยใช้สมการ (3.3)

ทำจนกระทั่ง L = t

ภาพประกอบ 3.8 ขั้นตอนวิธีการสร้างรูปแบบภาษาธรรมชาติจากความถี่ของช่วงข้อมูล

จากภาพประกอบ 3.7 กำหนดให้มีช่วงข้อมูล 3 ช่วงคือ $t_1 = [1,4)$ $t_2 = [4,6)$ และ $t_3 = [6,8]$ ตามลำดับ และค่าความกำกวมของช่วง t_1 คือ 0 ค่าความกำกวมของช่วง t_2 คือ 92 และค่า

ความกำกวมของช่วง t_3 คือ 32 ถ้าผู้ใช้ต้องการค่า L = 3 เราจะสามารถแทนค่า l_1 = [1,5) คือเล็ก l_2 = [4,6) คือกลาง และ l_3 = [6,8] คือใหญ่ได้โดยไม่ต้องมีการแบ่งช่วงหรือรวมช่วงข้อมูลอีก ในกรณีที่ผู้ใช้ต้องการค่า L = 4 คือ เล็ก กลาง ใหญ่ และใหญ่มาก เราจะต้องแบ่งช่วงข้อมูลเพิ่ม โดยพิจารณาค่าความกำกวมของช่วงข้อมูลที่สูงที่สุดคือ t_2 = [4,6) ที่มี ค่าความกำกวมของช่วงเท่ากับ 92 ทำการแบ่งครึ่งช่วง t_2 ออกเป็นสองส่วนเท่ากันคือ [4,5) และ [5,6) โดยช่วงอื่นยังคงเหมือนเดิม ทำให้ได้ช่วงใหม่คือ l_1 = [1,4) l_2 = [4,5) l_3 = [5,6) และ l_4 = [6,8]

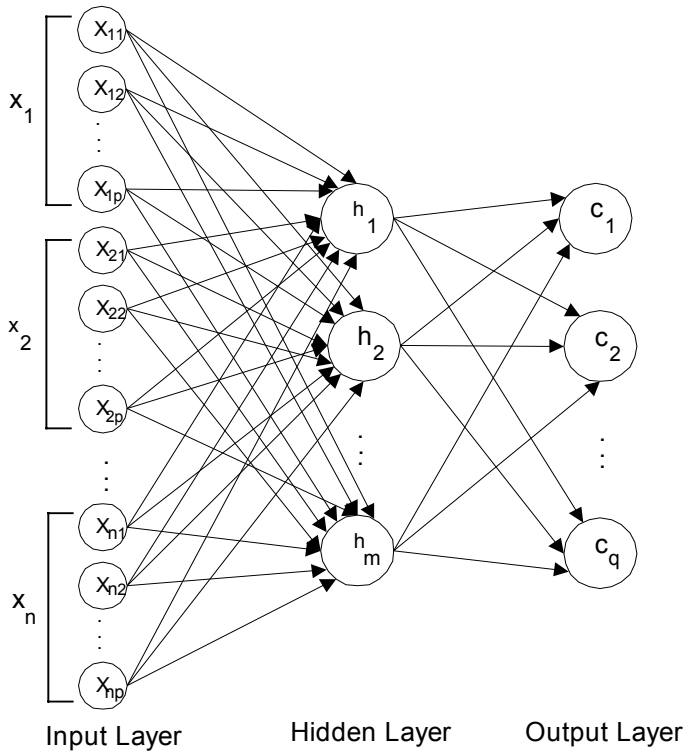
ในทางกลับกันจากภาพประกอบ 3.7 ถ้าผู้ใช้ต้องการค่า L = 2 คือ เล็ก และใหญ่ เราจะเลือกช่วงข้อมูลที่ใช้ในการรวมช่วงที่มีค่าความกำกวมของช่วงมากที่สุดเช่นกัน นั่นคือจะต้องทำการรวมช่วงข้อมูล t_2 = [4,6) เข้ากับทางซ้ายคือ t_1 หรือรวมช่วงข้อมูล t_2 เข้ากับทางขวาคือ t_3 จากการคำนวณค่าความกำกวมของช่วงจากการรวมช่วงทั้ง 2 แบบ ถ้า t_2 รวมกับ t_1 (ฝั่งซ้าย) สามารถคำนวณหาค่าความกำกวมของช่วงโดยใช้สมการที่ (3.3) ซึ่งมีค่าเท่ากับ (0+0+0+7+4)/(4+6+4+8+4) x 100 หรือ 42 และ ถ้ารวม t_2 เข้ากับ t_3 (ฝั่งขวา)ได้ค่าความกำกวมของช่วงเท่ากับ (7+4+3+2+1)/(8+4+7+8+4) x 100 หรือ 55 เราจะเลือกช่วงที่รวมแล้วได้ค่าความกำกวมของช่วงต่ำกว่าซึ่งก็คือช่วงทางฝั่งซ้าย ดังนั้นช่วงใหม่ที่ได้คือ l_1 = [1,6) และ l_2 = [6,8]

ขั้นตอนที่ 5 การสอนโครงข่ายประสาทเทียมด้วยรูปแบบภาษาธรรมชาติ

กำหนดให้รูปแบบโครงสร้างสถาปัตยกรรมของโครงข่ายประสาทเทียมมี 3 ชั้นคือ ชั้นข้อมูลเข้า ชั้นซ่อน และชั้นผลลัพธ์แสดงได้ดังภาพประกอบ 3.9 โดยจะทำการสอนโครงข่ายประสาทเทียมจนกระทั่งครบจำนวนรอบที่ตั้งไว้หรือได้ค่าความผิดพลาด (MSE) น้อยกว่าค่าที่ยอมรับได้

วิธีการแทนค่าข้อมูลเข้ามีดังนี้คือ ถ้าตำแหน่งของโหนดข้อมูลเข้าตรงกับรูปแบบภาษาธรรมชาติ เราจะแทนค่าเป็น 1 และโหนดที่ไม่ตรงกับรูปแบบภาษาธรรมชาติ เราจะแทนค่าเป็น 0 ตัวอย่างเช่น ถ้าที่โหนด x_1 มีการแบ่งช่วงภาษาธรรมชาติออกเป็น 3 ช่วงคือ เล็ก (S) กลาง (M) และใหญ่ (L) เราจะแทนค่าข้อมูลเข้าของ x_1 ด้วยโหนด 3 โหนดคือ [x_{11} x_{12} x_{13}] และแทนตำแหน่งด้วย [เล็ก กลาง ใหญ่] ดังนั้นการแทนข้อมูลเล็กจะแทนด้วย [1 0 0] การแทนข้อมูลกลางจะแทนด้วย [0 1 0] และการแทนข้อมูลใหญ่จะแทนด้วย [0 0 1] เป็นต้น ในการแทนค่าผลลัพธ์ ก็ใช้หลักการทำนองเดียวกันคือ ถ้าเรามีข้อมูล 2 กลุ่ม เช่น กลุ่มที่เป็นโรคหัวใจ และกลุ่มไม่เป็นโรคหัวใจ จะแทนค่าโหนดผลลัพธ์ด้วยโหนดเดียวเท่านั้น นั่นคือ [c_1] ของกลุ่ม1 จะแทนด้วย [1] กลุ่ม 2 จะแทนด้วย [0] เป็นต้น แต่ถ้าเรามีข้อมูลสามกลุ่มคือ กลุ่มที่ 1 กลุ่มที่ 2 และกลุ่มที่ 3 การแทนค่า

โหนดผลลัพธ์ [c_1 c_2 c_3] ของกลุ่ม1 จะแทนด้วย [1 0 0] กลุ่ม 2 จะแทนด้วย [0 1 0] และกลุ่ม 3 จะแทนด้วย [0 0 1] เป็นต้น



โดย x คือตัวแปรข้อมูลเข้า n ตัวแปร

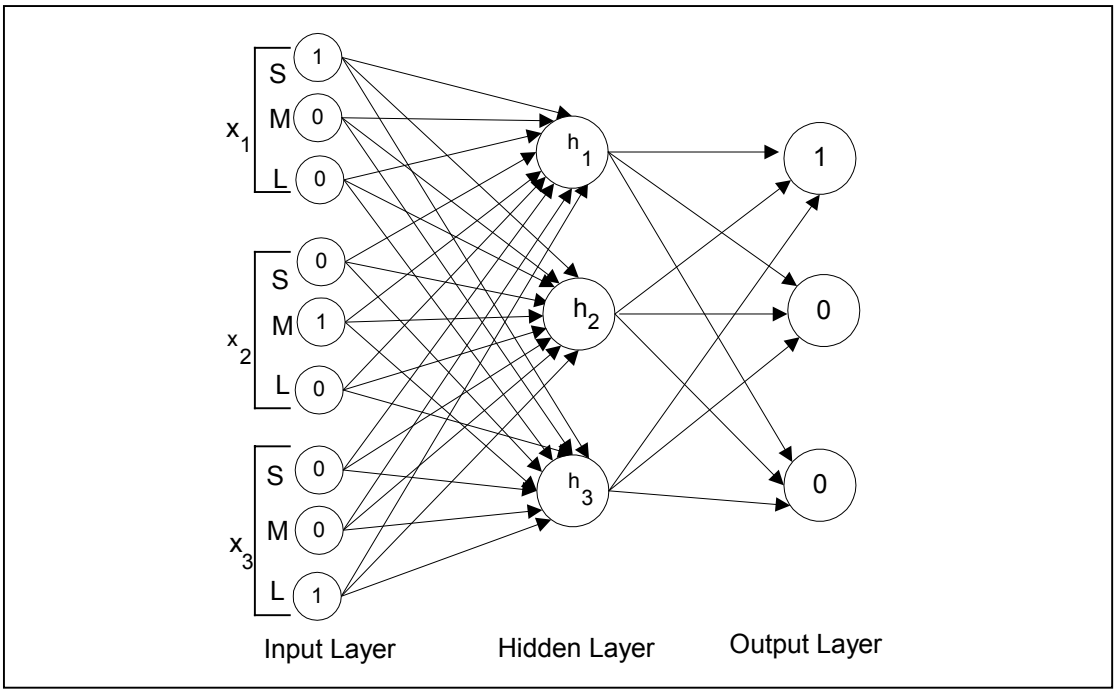
p คือจำนวนรูปแบบภาษาธรรมชาติของข้อมูลเข้าแต่ละตัวปี

h คือโหนดซ่อน m โหนด

c คือโหนดผลลัพธ์เข้า q โหนด

ภาพประกอบ 3.9 โครงสร้างของโครงข่ายประสาทเทียมสำหรับขั้นตอนหลักที่ II RENN-FIDR

ตัวอย่างการแทนค่าข้อมูลเข้าและค่าผลลัพธ์ในโครงข่ายประสาทเทียม กำหนดให้มีตัวแปรข้อมูลเข้า 3 ตัวแปร และผลลัพธ์มี 3 กลุ่ม แต่ละตัวแปรข้อมูลเข้ามีช่วงภาษาธรรม-ชาติได้ 3 ช่วงคือ เล็ก กลาง และใหญ่ มีโหนดซ่อน 3 โหนด และมีผลลัพธ์ได้ 3 กลุ่ม ถ้าตัวแปรข้อมูลเข้าที่ 1 (x_1) = เล็ก ตัวแปรข้อมูลเข้าที่ 2 (x_2) = กลาง ตัวแปรข้อมูลเข้าที่ 3 (x_3) = ใหญ่ และผลลัพธ์ = กลุ่มที่ 1 จะแทนได้ดังภาพประกอบ 3.10



ภาพประกอบ 3.10 ตัวอย่างการแทนค่าข้อมูลเข้าและค่าผลลัพธ์ในโครงข่ายประสาทเทียมสำหรับขั้นตอนหลักที่ II RENN-FIDR

ขั้นตอนที่ 6 ขั้นตอนวิธีการสร้างกฎโดยพิจารณาค่าน้ำหนักที่เป็นบวก

ขั้นตอนวิธีการสร้างกฎโดยพิจารณาค่าน้ำหนักที่เป็นบวกประกอบด้วย 2 ขั้นตอนย่อยดังนี้ ขั้นตอนย่อยที่ 6.1 คือการสร้างกฎจากค่าน้ำหนักที่มีค่าบวกที่ได้จากการสอนโครงข่ายประสาทเทียม ประกอบด้วย 2 ลำดับ คือการพิจารณาเฉพาะค่าน้ำหนักที่มีค่าเป็นบวก และสร้างเมตริกซ์เส้นเชื่อมน้ำหนักที่มีค่าเป็นบวก ขั้นตอนย่อยที่ 6.2 คือการเลือกกฎที่เหมาะสม ประกอบด้วย 3 ลำดับคือ การหาค่าความถูกต้องของข้อมูลในช่วงของกฎ (Inbound Correction: IC) การหาค่าความถูกต้องของข้อมูลนอกช่วงของกฎ (Outbound Correction: OC) และการหาค่าความถูกต้องของข้อมูลช่วงที่เหลือ ดังภาพประกอบ 3.11

ขั้นตอนที่ 6: ขั้นตอนวิธีการสร้างกฎโดยพิจารณาค่าน้ำหนักที่เป็นบวก

6.1 การสร้างกฎจากค่าน้ำหนักที่มีค่าบวกที่ได้จากการสอนโครงข่ายประสาทเทียม
  6.1.1 การพิจารณาเฉพาะค่าน้ำหนักที่มีค่าเป็นบวก
  6.1.2 สร้างเมตริกซ์เส้นเชื่อมน้ำหนักที่มีค่าเป็นบวก
6.2 การเลือกกฎที่เหมาะสม
  6.2.1 การหาค่าความถูกต้องของข้อมูลในช่วงของกฎ (Inbound Correction: IC)
  6.2.2 การหาค่าความถูกต้องของข้อมูลนอกช่วงของกฎ (Outbound Correction: OC)
  6.2.3 การหาค่าความถูกต้องของข้อมูลช่วงที่เหลือ

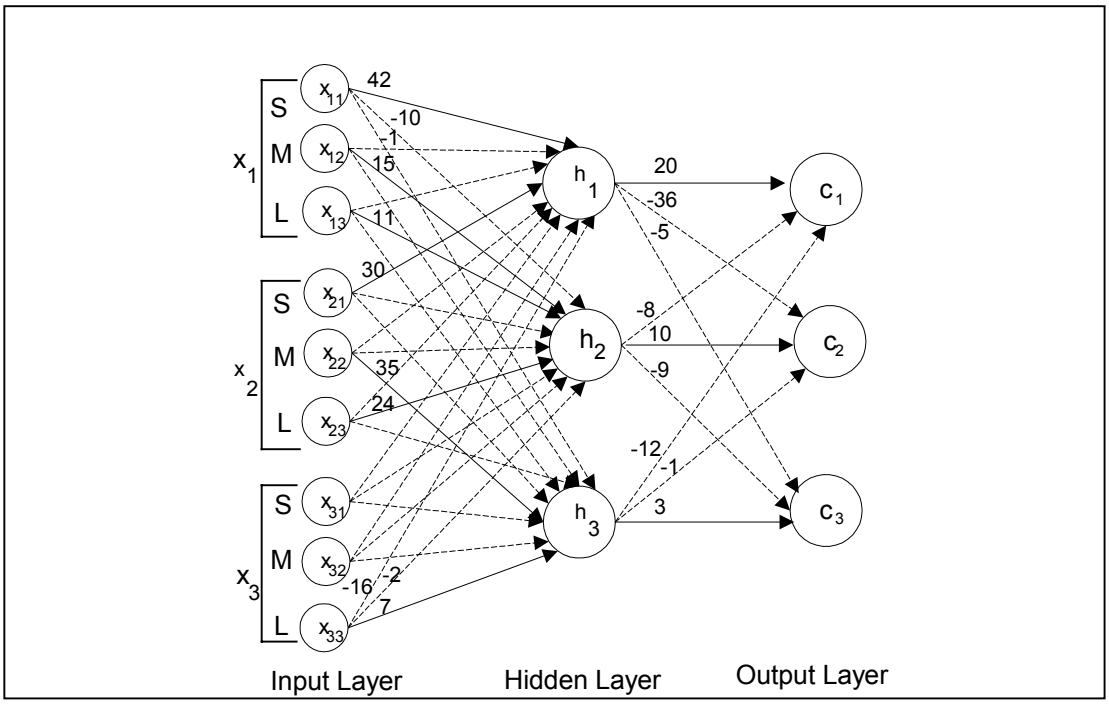
ภาพประกอบ 3.11 ขั้นตอนวิธีการสร้างกฎโดยพิจารณาค่าน้ำหนักที่เป็นบวก

ขั้นตอนย่อยที่ 6.1 การสร้างกฎจากค่าน้ำหนักที่มีค่าบวกที่ได้จากการสอนโครงข่ายประสาทเทียม จะพิจารณาเฉพาะค่าน้ำหนักที่มีค่าเป็นบวก ระหว่างชั้นข้อมูลเข้าและ ชั้นซ่อน รวมทั้งระหว่างชั้นซ่อนและชั้นผลลัพธ์ โดยไม่พิจารณาค่าน้ำหนักที่เป็นลบ เนื่องจาก น้ำหนักที่เป็นบวกจะมีผลต่อการทำนายกลุ่มที่เราสนใจมากกว่าน้ำหนักที่เป็นลบ แล้วนำมาหากฎที่เป็นไปได้ทั้งหมดในการแทนค่าข้อมูลของแต่ละกลุ่ม ตัวอย่างเช่น ในภาพประกอบ 3.12 กำหนดให้มีข้อมูลเข้า 3 ตัวแปรคือ x_1 x_2 และ x_3 โดยทั้ง 3 ตัวแปรมีจำนวนภาษาธรรมชาติ 3 ค่าคือ เล็ก (S) กลาง (M) และใหญ่ (L) ภาพประกอบ 3.12(a) แสดงค่าเส้นเชื่อมน้ำหนักและ ค่าน้ำหนักที่มีค่าเป็นบวกจากการสอนโครงข่ายประสาทเทียม และภาพประกอบ 3.12(b) แสดงเฉพาะเส้นเชื่อมน้ำหนักและค่าน้ำหนักที่เป็นบวกเท่านั้น

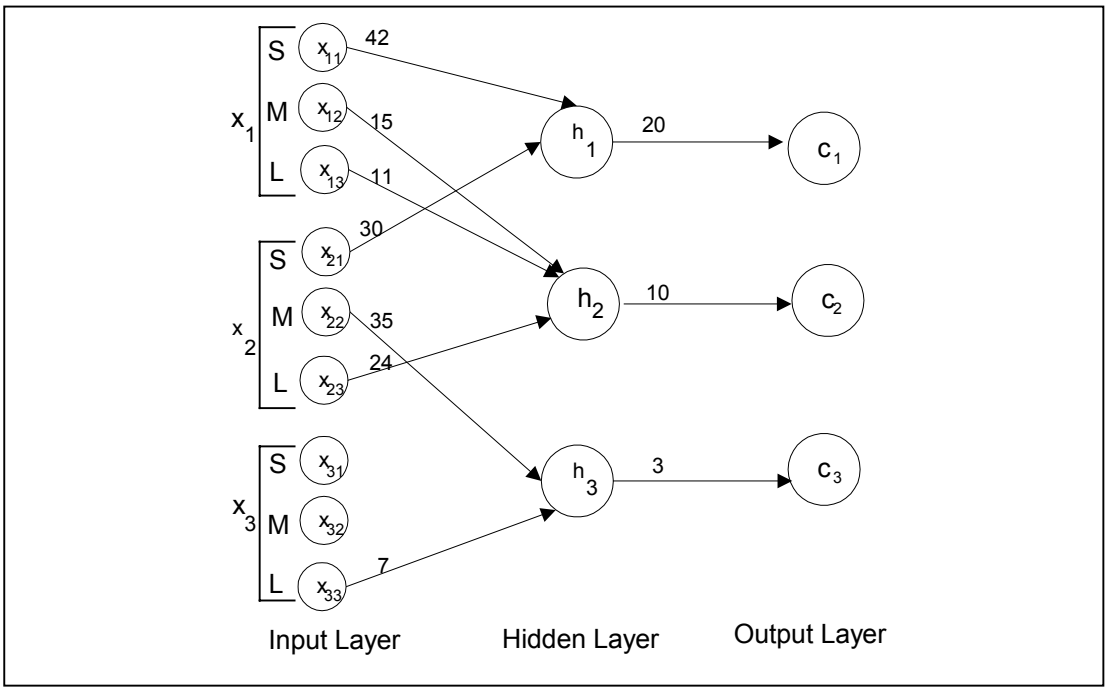
จากภาพประกอบ 3.13 กำหนดให้ R_i แทนค่าเส้นเชื่อมน้ำหนักของกลุ่มข้อมูลกลุ่มที่ q ที่มีค่าน้ำหนักเป็นบวกมีจำนวนทั้งหมด r เส้น แสดงได้ดังภาพประกอบ 3.13(a) จากนั้นเรานำเส้นเชื่อมดังกล่าวมาเรียงสับเปลี่ยนโดยใช้ตัวเชื่อมเงื่อนไข “และ” เพื่อหาค่าความเป็นไปได้ของกฎทั้งหมดของกลุ่มที่ q ซึ่งแทนด้วยเมตริกซ์ดังภาพประกอบ 3.13 (b) โดยกำหนดให้น้ำหนักของเส้นเชื่อมที่มีค่ามากที่สุดแทนด้วยกฎ R_{11} น้ำหนักของเส้นเชื่อมรองลงมาแทนด้วยกฎ R_{22} และค่าน้ำหนักของกฎที่ R_{rr} มีค่าน้อยที่สุด โดยค่าน้ำหนักของกฎอยู่ในลักษณะ $[R_{11} > R_{22} > R_{33} > \ldots > R_{rr}]$ ตัวอย่างกฎที่สร้างได้เช่น กฎ R_{13} หมายถึง กฎ “R_1 and R_2 and R_3” ภาพประกอบ 3.13(c) แสดง

กฎที่ตัดค่าซ้ำซ้อนระหว่างกฎ R_{ij} และกฎ R_{ji} ซึ่งหมายถึงกฎเดียวกัน ดังนั้นจำนวนกฎที่เหลือของแต่ละกลุ่มข้อมูลมีค่าเท่ากับผลรวมของจำนวนเส้นเชื่อมน้ำหนักที่มีค่าเป็นบวก เมื่อ r คือจำนวนเส้นเชื่อมน้ำหนักที่มีค่าเป็นบวกดังสมการที่ (3.4)

$$\text{จำนวนกฎที่เหลือของกลุ่มข้อมูล} = \frac{r \times (r+1)}{10} \qquad (3.4)$$

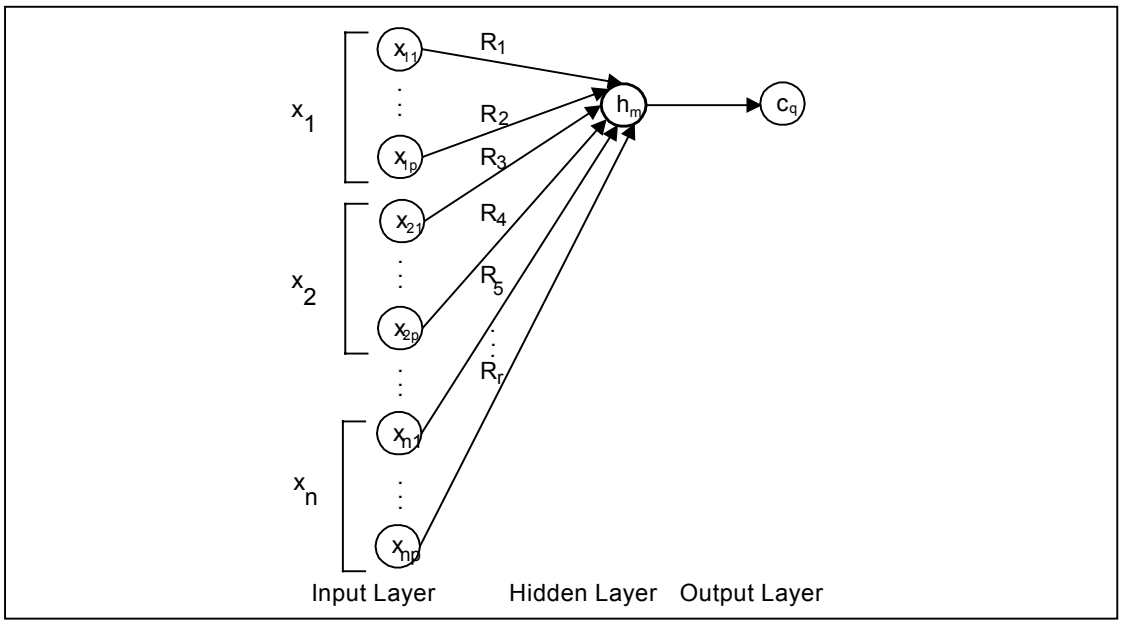


(a) เส้นเชื่อมน้ำหนักและค่าน้ำหนักทั้งหมด



(b) เส้นเชื่อมน้ำหนักที่มีค่าน้ำหนักเป็นบวก

ภาพประกอบ 3.12 ค่าน้ำหนักของโครงสร้างโครงข่ายประสาทเทียม RENN-FIDR



(a) เส้นเชื่อมน้ำหนัก r เส้นในกลุ่มที่ q

$$\text{Matrix } c_q = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1r} \\ R_{21} & R_{22} & \dots & R_{2r} \\ M & M & M & M \\ R_{r1} & R_{r2} & \dots & R_{rr} \end{bmatrix}$$

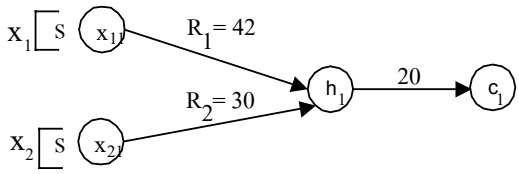
(b) เมตริกซ์ของเส้นเชื่อมกลุ่มที่ q

$$\text{Matrix } c_q = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1r} \\ & R_{22} & \dots & R_{2r} \\ & & M & M \\ & & & R_{rr} \end{bmatrix}$$

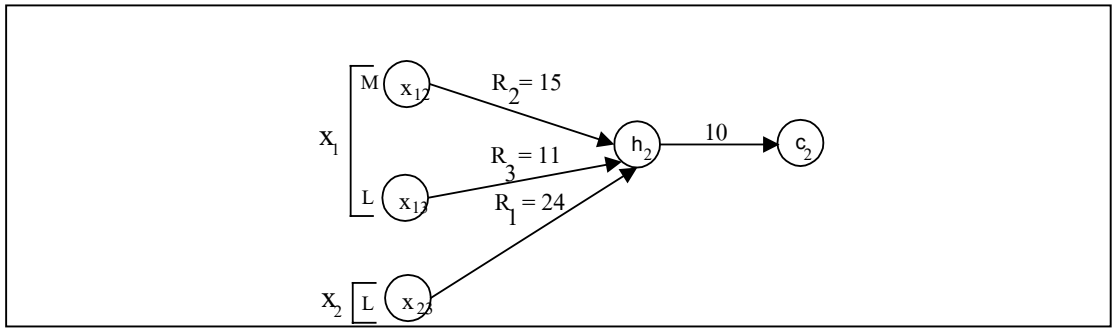
(c) เมตริกซ์ที่ตัดค่าซ้ำซ้อนระหว่างกฎ R_{ij} และกฎ R_{ji} ของกลุ่มที่ q

ภาพประกอบ 3.13 การสร้างกฎที่เป็นไปได้ทั้งหมดจากค่าน้ำหนักที่มีค่าเป็นบวก

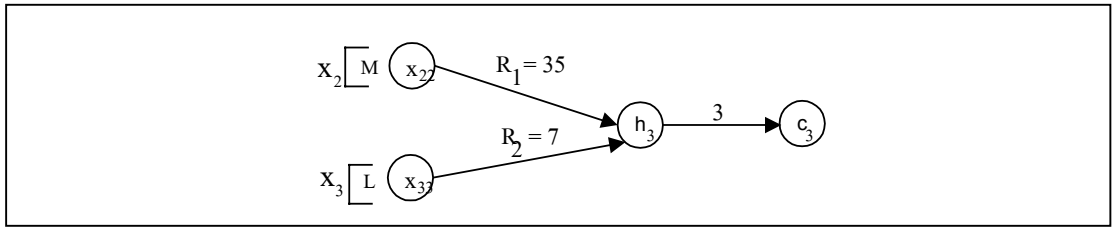
จากภาพประกอบ 3.12(b) สามารถแสดงตัวอย่างในการแทนกฎต่างๆจาก ความสัมพันธ์ระหว่างเส้นเชื่อมให้อยู่ในรูปแบบของกฎ “ถ้า-แล้ว” ในลักษณะรูปแบบภาษา ธรรมชาติได้ดังต่อไปนี้ ภาพประกอบ 3.14 แสดงค่าน้ำหนักของเส้นเชื่อมของข้อมูลในแต่ละกลุ่ม ข้อมูลกลุ่มที่ 1 มีค่าเส้นเชื่อมที่เป็นบวก 2 เส้น ดังนั้นจะได้ค่า r = 2 [R_1 R_2] แสดงได้ดังภาพประกอบ 3.14(a) ข้อมูลกลุ่มที่ 2 มีค่าเส้นเชื่อม 3 เส้นหรือ r = 3 [R_1 R_2 R_3] แสดงได้ดังภาพประกอบ 3.14(b) และข้อมูลกลุ่มที่ 3 มีค่าเส้นเชื่อม 2 เส้นหรือ r = 2 [R_1 R_2] แสดงได้ดังภาพประกอบ 3.14(c) จำนวนกฎที่ได้จากเมตริกซ์ของกฎของข้อมูลกลุ่มที่ 1 แสดงได้ดังภาพประกอบ 3.15(a) มีจำนวนกฎ 3 กฎคือ R_{11} R_{12} และ R_{22} ข้อมูลกลุ่มที่ 2 แสดงได้ดังภาพประกอบ 3.15(b) มีจำนวนกฎ 6 กฎคือ R_{11} R_{12} R_{13} R_{22} R_{23} และ R_{33} และข้อมูลกลุ่มที่ 3 แสดงได้ดังภาพประกอบ 3.15(b) มีจำนวนกฎ 3 กฎคือ R_{11} R_{12} และ R_{22} จากตัวอย่างในภาพประกอบ 3.14 และภาพประกอบ 3.15 ของแต่ละกลุ่มข้อมูลสามารถสร้างกฎภาษาธรรมชาติได้ดังภาพประกอบ 3.16 ข้อมูลกลุ่มที่ 1 กฎ R_{11} คือกฎ R_1 and R_1 ซึ่งกฎ R_1 เป็นกฎที่เกิดจาก เส้นเชื่อมของโหนดข้อมูลเข้า x_1 มีค่าเท่ากับ small ดังภาพประกอบ 3.14(a) ดังนั้น R_{11} จึงมีค่าเท่ากับ If x_1 = small then class 1 กฎ R_{12} คือกฎ R_1 and R_2 ซึ่งกฎ R_2 เป็นกฎที่เกิดจากเส้นเชื่อมของโหนดข้อมูลเข้า x_2 มีค่าเท่ากับ small ดังภาพประกอบ 3.14(a) ดังนั้น R_{12} มีค่า เท่ากับ If x_1 = small and x_2 = small then class 1 และกฎ R_{22} มีค่าเท่ากับ If x_2 = small then class 1 แสดงได้ดังภาพประกอบ 3.16 (a) กฎภาษาธรรมชาติของข้อมูลกลุ่มที่ 2 และข้อมูล กลุ่มที่ 3 แสดงได้ดังภาพประกอบ 3.16(b) และ 3.16(c) ตามลำดับ



(a) ข้อมูลกลุ่มที่ 1

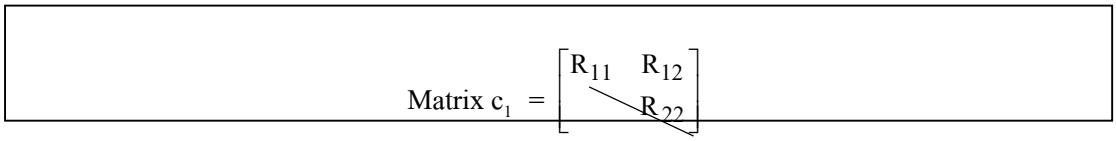


(b) ข้อมูลกลุ่มที่ 2



(c) ข้อมูลกลุ่มที่ 3

ภาพประกอบ 3.14 ค่าน้ำหนักเส้นเชื่อมของข้อมูลในแต่ละกลุ่มจากตัวอย่างภาพประกอบ 3.12



(a) ข้อมูลกลุ่มที่ 1

$$\text{Matrix } c_2 = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ & R_{22} & R_{23} \\ & & R_{33} \end{bmatrix}$$

(b) ข้อมูลกลุ่มที่ 2

$$\text{Matrix } c_3 = \begin{bmatrix} R_{11} & R_{12} \\ & R_{22} \end{bmatrix}$$

(c) ข้อมูลกลุ่มที่ 1

ภาพประกอบ 3.15 เมตริกซ์เส้นเชื่อมของข้อมูลจากตัวอย่างภาพประกอบ 3.14

R_{11} : If x_1 = small then class 1

R_{22}: If x_2 = small then class 1

R_{12} : If x_1 = small and x_2 = small then class 1

(a) ข้อมูลกลุ่มที่ 1

R_{11} : If x_2 = large then class 2

R_{22}: If x_1 = medium then class 2

R_{33}: If x_1 = large then class 2

R_{12} : If x_2 = large and x_1 = medium then class 2

R_{23} : If x_1 = medium and x_1 = large then class 2

R_{13} : If x_2 = large and x_1 = medium and x_1 = large then class 2

(b) ข้อมูลกลุ่มที่ 2

R_{11} : If x_2 = medium then class 3

R_{22}: If x_3 = large then class 3

R_{12} : If x_2 = medium and x_3 = large then class 3

(c) ข้อมูลกลุ่มที่ 3

ภาพประกอบ 3.16 กฎที่สร้างได้ของข้อมูลจากตัวอย่างภาพประกอบ 3.14 และ 3.15

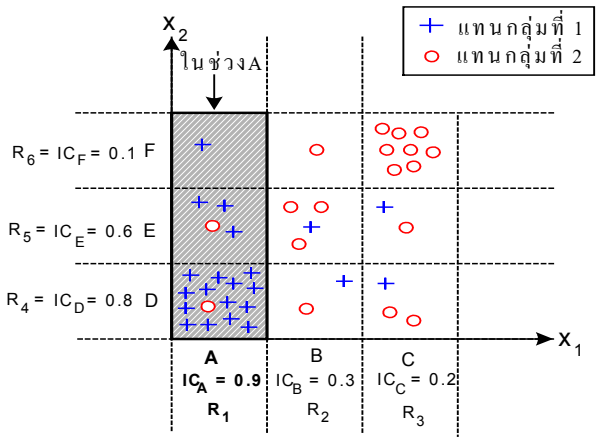
ขั้นตอนย่อยที่ 6.2 การเลือกกฎที่เหมาะสม หรืออีกนัยหนึ่งคือการลดจำนวนกฎที่ได้จากขั้นตอนย่อยที่ 6.1 วิธีการลดจำนวนกฎมี 3 ลำดับคือ 6.2.1 การหาค่าความถูกต้องของข้อมูลในช่วงของกฎ (Inbound Correction: IC) 6.2.2 การหาค่าความถูกต้องของข้อมูลนอกช่วงของกฎ (Outbound Correction: OC) และ 6.2.3 การหาค่าความถูกต้องของข้อมูลช่วงที่ไม่ได้เลือกในลำดับที่ 6.2.1 และ 6.2.2

ลำดับที่ 6.2.1 การหาค่าความถูกต้องของข้อมูลในช่วง (IC) จะพิจารณาค่าความถูกต้องของกฎที่ได้มาจากขั้นตอนย่อยที่ 6.1

นิยามที่ 4 ค่าความถูกต้องของข้อมูลในช่วงของกฎ (Inbound Correction: IC) มีค่าเท่ากับจำนวนข้อมูลที่ทำนายถูกต้องในช่วงของกฎหารด้วยจำนวนข้อมูลทั้งหมดในช่วงนั้น มีค่าอยู่ระหว่าง [0,1] แสดงได้ดังสมการที่ (3.5)

$$\text{ค่าความถูกต้องของข้อมูลในช่วงของกฎ} = \frac{\text{จำนวนข้อมูลที่ทำนายถูกต้องในช่วงของกฎ}}{\text{จำนวนข้อมูลทั้งหมดในช่วงของกฎ}} \qquad (3.5)$$

ตัวอย่างการคำนวณค่าความถูกต้องของข้อมูลในช่วงของกฎแสดงได้ดังภาพประกอบ 3.17 กำหนดให้เครื่องหมายบวก (+) แทนข้อมูลกลุ่มที่ 1 และเครื่องหมาย (o) แทนข้อมูลกลุ่มที่ 2 ถ้าพิจารณา ช่วง A ของตัวแปรข้อมูลเข้า x_1 จะได้ว่าข้อมูลกลุ่มที่ 1 มี 18 จำนวน และข้อมูลกลุ่มที่ 2 มี 2 จำนวน ทำให้ได้ค่าความถูกต้องของ ข้อมูลกลุ่มที่ 1 ในช่วง A มีค่า เท่ากับ 18/(18+2) หรือ 0.9 เป็นต้น



ภาพประกอบ 3.17 ค่าความถูกต้องในช่วง (IC) ของกฎจากข้อมูลกลุ่มที่ 1

หลังจากคำนวณค่าความถูกต้องของข้อมูลในช่วงของทุกช่วงแล้ว กำหนดให้ α คือค่าที่ยอมรับได้ (Inbound Threshold Value) ของกลุ่มข้อมูลที่เราสนใจเท่ากับ 0.5 เราจะพิจารณาคัดเลือกช่วงที่แทนสำหรับกฎของกลุ่มข้อมูลนั้นๆ โดยจะทิ้งกฎที่มีค่าน้อยกว่าค่าที่ยอมรับได้ จากตัวอย่างในภาพประกอบ 3.15 จะได้กฎที่ยอมรับได้ คือ [R_1 R_4 R_5] เพราะมีค่า IC คือ [0.9 0.8 0.6] สำหรับกฎที่ถูกตัดทิ้งไปคือ [R_2 R_3 R_6] ที่มีค่า IC คือ [0.3 0.2 0.1]

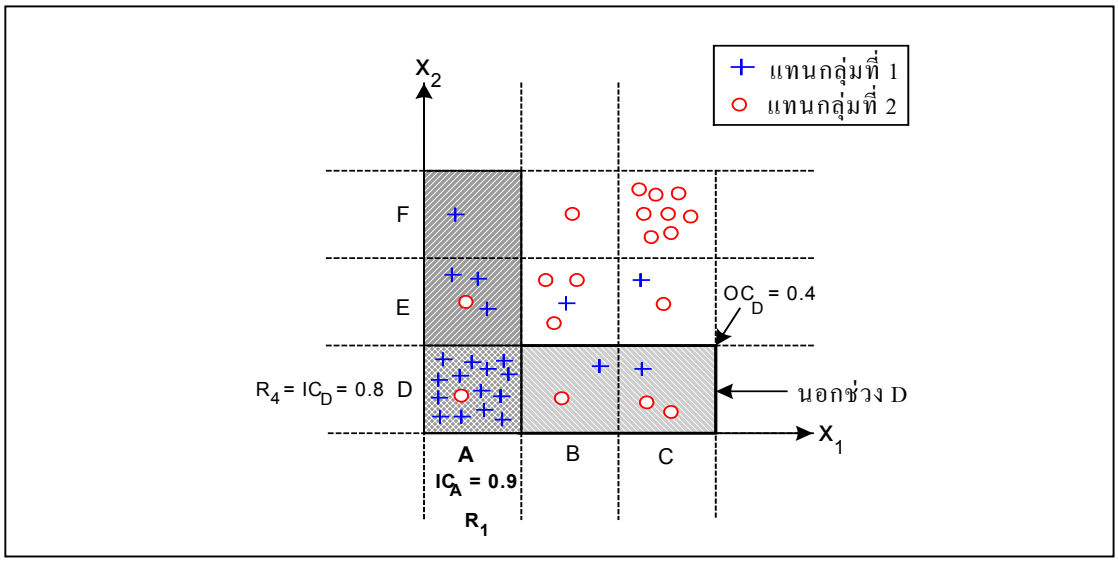
ลำดับที่ 6.2.2 การหาค่าความถูกต้องของข้อมูลนอกช่วงของกฎ (OC) เนื่องจากกฎที่เลือกได้จากขั้นตอนย่อยที่ 2.1 อาจมีช่วงของข้อมูลที่ซ้ำซ้อนกัน ดังนั้นจึงต้องมี การคำนวณค่าความถูกต้องของข้อมูลนอกช่วง

นิยามที่ 5 ค่าความถูกต้องของข้อมูลนอกช่วง (Outbound Correction: OC) มีค่าเท่ากับจำนวนข้อมูลที่ทำนายถูกต้องนอกช่วงของกฎหารด้วยจำนวนข้อมูลทั้งหมดนอกช่วงกฎ มีค่าอยู่ระหว่าง [0,1] แสดงได้ดังสมการที่ (3.6)

$$\text{ค่าความถูกต้องของข้อมูลนอกช่วง} = \frac{\text{จำนวนข้อมูลที่ทำนายถูกต้องนอกช่วงของกฎ}}{\text{จำนวนข้อมูลทั้งหมดนอกช่วงของกฎ}} \quad (3.6)$$

ตัวอย่างการคำนวณค่าความถูกต้องของข้อมูลนอกช่วงแสดงได้ดังภาพประกอบ 3.18 สมมติว่าช่วงที่เลือกไปแล้วคือช่วง A ของตัวแปรข้อมูลเข้า x_1 และกำลังพิจารณา ข้อมูลในช่วง D ของตัวแปรข้อมูลเข้า x_2 จะเห็นได้ว่าบริเวณ AD จะมีความซ้ำซ้อนของข้อมูล ดังนั้น

จึงคำนวณค่าความถูกต้องของข้อมูลของช่วง D ที่อยู่นอกช่วง A เพราะข้อมูลในช่วง A ได้พิจารณาไปก่อนหน้านี้แล้วจากกฎ R_1 ซึ่งคือช่วง B ถึง C นั่นเอง จากภาพประกอบ 3.18 มีจำนวนข้อมูลที่เป็นกลุ่มที่ 1 (+) อยู่ 2 จำนวน และมีจำนวนข้อมูลที่เป็นกลุ่มที่ 2 (o) อยู่ 3 จำนวน ทำให้ได้ค่าความถูกต้องของข้อมูลนอกช่วงของช่วง D เท่ากับ 2/(2+3) หรือ 0.4 เป็นต้น ในการคัดเลือกกฎจะใช้ค่าที่ยอมรับได้ (Outbound Threshold Value) α' เท่ากับ 0.5 ในการเลือกกฎเช่นกัน จากตัวอย่างในภาพประกอบ 3.18 กฎ R_4 จะถูกตัดทิ้งเนื่องจากมีค่าความ-ถูกต้องของข้อมูลนอกช่วงเท่ากับ 0.4 ซึ่งน้อยกว่าค่าที่ยอมรับได้



ภาพประกอบ 3.18 ค่าความถูกต้องนอกช่วง (OC) ของกฎ R_4 จากข้อมูลกลุ่มที่ 1

ลำดับที่ 6.2.3 การหาค่าความถูกต้องของข้อมูลช่วงที่ไม่ได้ถูกเลือกในลำดับที่ 6.2.1 และลำดับที่ 6.2.2 มาเลือกเป็นส่วนหนึ่งของกฎที่เราพิจารณาเนื่องจากอาจมีบางบริเวณของช่วงข้อมูลที่มีค่าความถูกต้องของข้อมูลในช่วง และค่าความถูกต้องของข้อมูลนอกช่วงน้อยกว่า 0.5

ดังนั้นเราจึงต้องหาค่าความถูกต้องของข้อมูลในช่วงที่เหลือดังกล่าวว่าอยู่ในกลุ่มใด และแทนค่าช่วงที่เหลือนี้ด้วยกลุ่มข้อมูลที่มากที่สุดหลังคำสั่ง else ในกฎสุดท้าย

วิทยานิพนธ์ได้นำเสนอแบบจำลองการสกัดกฎภาษาธรรมชาติจากโครงข่ายประสาทเทียมที่ผ่านการตัดโหนดโดยใช้การแทนค่าความถี่ของช่วงข้อมูล (Rule Extraction form Neural Networks Pruning using Frequency Interval Data Representation: RENN-P-FIDR) มีการทำงาน 2 ขั้นตอนหลักคือ ขั้นตอนหลักที่ I การตัดโหนดโครงข่ายประสาทเทียมโดยใช้ ค่าน้ำหนักสูงสุด (Neural Network Pruning using Maximum Feature Weighting: NNP-MFW) มีการทำงาน 3 ขั้นตอนคือ ขั้นตอนที่ 1 การเตรียมข้อมูล ขั้นตอนที่ 2 การสอนโครงข่ายประสาทเทียมด้วยการเรียนรู้แบบแพร่ย้อนกลับ ขั้นตอนที่ 3 ขั้นตอนวิธีการตัดโหนดโครงข่ายประสาทเทียมโดยใช้ค่าน้ำหนักสูงสุด และขั้นตอนหลักที่ II การสกัดกฎภาษาธรรมชาติจาก โครงข่ายประสาทเทียมโดยการแทนค่าความถี่ของช่วงข้อมูล (Rule Extraction form Neural Networks using Frequency Interval Data Representation: RENN-FIDR) มีการทำงาน 3 ขั้นตอนคือ ขั้นตอนที่ 4 ขั้นตอนวิธีการสร้างรูปแบบภาษาธรรมชาติจากความถี่ของช่วงข้อมูล ขั้นตอนที่ 5 การสอนโครงข่ายประสาทเทียมด้วยด้วยรูปแบบภาษาธรรมชาติ และขั้นตอนที่ 6 ขั้นตอนวิธีการสร้างกฎภาษาธรรมชาติโดยใช้ค่าน้ำหนักที่เป็นบวก ซึ่งได้กฎภาษาธรรมชาติที่ได้ ผู้ใช้สามารถเข้าใจได้ง่าย