

การทำนายรูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม
และการวิเคราะห์ความถดถอยเชิงพหุ

นางสาวภัทรพร เพชรสังคุณ

นายภาคภูมิ รวยลาภ

นางสาวเมษาศิณี กลิ่นมาลัย

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต

สาขาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2550

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

**การทำนายรูปร่างของรอยเชื่อมโดยใช้ โคร่งข่ายประสาทเทียม
และการวิเคราะห์ความถดถอยเชิงพหุ**

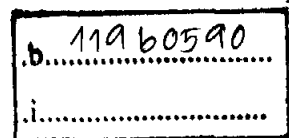
นางสาวภัทรพร เพชรสังคุณ
นายภาคภูมิ รวยลาภ
นางสาวเมษาศิณี กลิ่นมาลัย

ร.พ.

ว 3677

2550

เลขที่.....
เลขทะเบียน..... **83166**
วัน,เดือน,ปี..... - 6 ส.ค. 2551



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต
สาขาวิศวกรรมอุตสาหกรรม คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2550

**PREDICTION OF WELD GEOMETRY USING ARTIFICIAL
NEURAL NETWORKS AND MULTIPLE REGRESSION**

MISS PATTARAPORN PETSANGKON

MR. PHAKPHOOM RUAILARP

MISS MAESASINEE KLINMALAI

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
BACHELOR OF ENGINEERING IN INDUSTRIAL ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2007**

คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองปริญญาโท

หัวข้อปริญญาโท

การทำนายรูปร่างของรอยเชื่อม โดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์
ความถดถอยเชิงพหุ
PREDICTION OF WELD GEOMETRY USING ARTIFICIAL NEURAL
NETWORKS AND MULTIPLE REGRESSION

นักศึกษา

นางสาวภัทรพร เพชรสังคุณ	รหัสประจำตัว	47010550
นายภาคภูมิ รวยลาภ	รหัสประจำตัว	47010561
นางสาวเมษาศิณี กลิ่นมาลัย	รหัสประจำตัว	47010601

หลักสูตร

วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมอุตสาหการ

อาจารย์ผู้ควบคุมปริญญาโท


(ผศ.ดร.กรรณชัย กัลยาศิริ)

หัวข้อปริญญานิพนธ์	การทำนารูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ
นักศึกษา	นางสาวภัทรพร เพชรสังคุณ นายภาคภูมิ รวยลาภ นางสาวเมษาศิณี กลิ่นมาลัย
หลักสูตร	วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมอุตสาหการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา	2550
อาจารย์ผู้ควบคุมปริญญานิพนธ์	ผศ.ดร.กรรณชัย กัลยาศิริ

บทคัดย่อ

ปริญญานิพนธ์นี้ได้นำเสนอการพยากรณ์รูปร่างของรอยเชื่อมที่เกิดจากการเชื่อมแบบ Gas tungsten arc welding โดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ ซึ่งมีปัจจัยในการเชื่อมได้แก่ Arc current, Arc Voltage, Travel speed, Wire speed และรูปร่างของรอยเชื่อมที่ต้องการพยากรณ์คือ Bead width, Penetration, Reinforcement height และ Cross section โดยผลจากการพยากรณ์ของ โครงข่ายประสาทเทียมจะนำไปเปรียบเทียบกับผลจากการพยากรณ์ด้วยวิธีการวิเคราะห์ความถดถอยเชิงพหุที่ใช้ปัจจัย และจำนวนข้อมูลทั้งหมดในการวิเคราะห์โครงสร้างเหมือนกัน จากการเปรียบเทียบความถูกต้องของการพยากรณ์ด้วยวิธี โครงข่ายประสาทเทียมและวิธีการวิเคราะห์ความถดถอยเชิงพหุด้วยค่าความผิดพลาดกำลังสองเฉลี่ย พบว่ารูปร่างของรอยเชื่อมคือ Penetration, Reinforcement height และ Cross section ที่ได้จากการพยากรณ์ด้วยวิธีโครงข่ายประสาทเทียมมีค่าความผิดพลาดกำลังสองเฉลี่ยน้อยกว่าค่าที่ได้จากวิธีการวิเคราะห์ความถดถอยเชิงพหุ ยกเว้น Bead width เพียงค่าเดียวที่วิธีการวิเคราะห์ความถดถอยเชิงพหุมีค่าความผิดพลาดกำลังสองเฉลี่ยน้อยกว่า จึงสรุปได้ว่าวิธี โครงข่ายประสาทเทียมสามารถพยากรณ์รูปร่างรอยเชื่อม ได้ดีกว่าวิธีการวิเคราะห์ความถดถอยเชิงพหุ

Thesis Title Prediction of Weld Geometry Using Artificial Neural Networks and Multiple Regression
Student Miss Pattaraporn Petsangkoon
Mr. Phakphoom Ruailarp
Miss Maesasinee Klinmalai
Degree Bachelor of Engineering in Industrial Engineering
King Mongkut's Institute of Technology Ladkrabang
Academic Year 2007
Thesis Advisor Asst.Prof.Dr. Kannachai Kanlayasiri

ABSTRACT

This thesis presented artificial neural networks and multiple regression in predicting weld geometry of gas tungsten arc welding. The welding variables studied were arc current, arc voltage, travel speed and wire speed. Predicting outputs were bead width, penetration, reinforcement height and cross section. Mean square error was used to determine the prediction ability of both analysis methods. Results showed that the mean square errors of penetration, reinforcement height and cross section from artificial neural networks were lower than those of multiple regression while mean square error of bead width prediction from artificial neural networks model was higher than that of multiple regression model. In conclusion, artificial neural networks generally provided better prediction ability than multiple regression.

กิตติกรรมประกาศ

ปริญญาานิพนธ์เรื่อง การทำนาขรุขระของรอยเขื่อนโดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ สามารถสำเร็จลุล่วงได้ด้วยดี กลุ่มผู้วิจัยขอกราบขอบพระคุณบุคคลทุกท่านที่มีส่วนเกี่ยวข้องที่ช่วยให้ปริญญาานิพนธ์ฉบับนี้เสร็จสมบูรณ์

ผศ.ดร.กรรณชัย กัลยาศิริ อาจารย์ที่ปรึกษาปริญญาานิพนธ์ กลุ่มผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงสำหรับการให้โอกาสในการศึกษาปริญญาานิพนธ์ฉบับนี้ รวมทั้งความรู้ คำแนะนำ ความช่วยเหลือ และความเอาใจใส่ในทุกๆ ด้านตลอดระยะเวลาที่ผ่านมา

ผศ.ดร.สรรพสิทธิ์ ลิ้มบรรณรัตน์ กลุ่มผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงสำหรับ ความรู้ คำแนะนำ การรับฟังปัญหา กำลังใจในการทำงาน ความเอาใจใส่ และความช่วยเหลือในทุกๆด้าน ตลอดการศึกษาระดับปริญญาตรี ในหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมอุตสาหกรรม

อาจารย์เชาวลิต หามนตรี กลุ่มผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง สำหรับความรู้ คำแนะนำ และความช่วยเหลือในทุกๆด้าน

ดร.อนิรุท ไชยจาวณิช กลุ่มผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง สำหรับความรู้ คำแนะนำ และความช่วยเหลือในทุกๆด้าน

อาจารย์กิตติวัฒน์ สิริเกษมสุข กลุ่มผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง สำหรับความรู้ คำแนะนำ และความช่วยเหลือตลอดระยะเวลาที่ผ่านมา

สำหรับคุณงามความดีที่เกิดจากปริญญาานิพนธ์ฉบับนี้ ขอมอบให้กับสมาชิกในกลุ่มผู้วิจัยทุกคนและคณะอาจารย์ทุกท่านที่มีส่วนในการให้ความรู้ และถ่ายทอดประสบการณ์ต่างๆ จนทำให้การวิจัยสำเร็จลุล่วงได้ด้วยดี

นางสาวภัทรพร เพชรสังคุณ

นายภาคภูมิ รวยลาภ

นางสาวเมษาศิณี กลิ่นมาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ

บทที่ 1 บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	1
1.3 ขอบเขตการศึกษา.....	1
1.4 ประโยชน์ที่คาดว่าจะได้รับ	1

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง

2.1 การเชื่อมแบบ Gas tungsten arc welding	2
2.2 โครงข่ายประสาทเทียม	4
2.2.1 ความเป็นมาของโครงข่ายประสาทเทียม.....	4
2.2.1.1 ช่วงทศวรรษที่ 40 ยุคเริ่มต้นของโครงข่ายประสาทเทียม	4
2.2.1.2 ช่วงทศวรรษที่ 50-60 ยุคทองของโครงข่ายประสาทเทียม	5
2.2.1.3 ช่วงทศวรรษที่ 70 ยุคเงียบของโครงข่ายประสาทเทียม	5
2.2.1.4 ช่วงทศวรรษที่ 80 ยุคความกระตือรือร้นกลับมาใหม่ของโครงข่ายประสาทเทียม.....	6
2.2.2 เซลล์ประสาททางชีววิทยา.....	7
2.2.2.1 นิวรอน	7
2.2.2.2 โครงสร้างของนิวรอนและไซแนปส์.....	7
2.2.2.3 การส่งผ่านกระแสประสาท.....	8
2.2.3 เซลล์ประสาทเทียม.....	9
2.2.3.1 เซลล์ประสาทเทียมอินพุตเดียว.....	10
2.2.3.2 เซลล์ประสาทเทียมหลายอินพุต	11
2.2.3.3 คุณสมบัติและความสามารถของโครงข่ายประสาทเทียม.....	12
2.2.3.4 ลักษณะปัญหาที่เหมาะสมกับโครงข่ายประสาทเทียม.....	12

สารบัญ (ต่อ)

	หน้า
2.2.4 ชนิดของฟังก์ชันถ่ายโอน	12
2.2.5 โครงสร้างของโครงข่ายประสาทเทียม	14
2.2.5.1 Feedforward Networks	14
2.2.5.2 Recurrent Networks.....	16
2.2.6 การเรียนรู้ของโครงข่ายประสาทเทียม.....	17
2.2.6.1 การเรียนรู้แบบควบคุม (Supervised Learning).....	17
2.2.6.2 การเรียนรู้แบบไม่มีการควบคุม (Unsupervised Learning).....	17
2.2.7 โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีหลายชั้น	17
2.2.8 กฎการเรียนรู้แบบแพร่ค่าย้อนกลับ (Back-propagation Algorithm).....	19
2.2.9 การ Normalization และ Unnormalization ชุดข้อมูล.....	24
2.3 การวิเคราะห์ความดัดออยเชิงพหุ	25
2.3.1 การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยสุด	25
2.3.2 การวิเคราะห์ความแปรปรวน (Analysis of Variance)	28
2.3.3 การประมาณค่าความคลาดเคลื่อน.....	30
2.3.4 การประมาณค่าเบี่ยงเบนมาตรฐานของสัมประสิทธิ์การดัดออย	30
2.3.5 ความหมายของสัมประสิทธิ์ความดัดออยเชิงส่วน.....	31
2.3.6 สัมประสิทธิ์การตัดสินใจเชิงพหุ (Multiple Coefficient of Determination : R^2 หรือ r^2).....	32
2.3.7 สัมประสิทธิ์สหสัมพันธ์เชิงพหุ (Multiple Coefficient of Correlation).....	33
2.3.8 สัมประสิทธิ์สหสัมพันธ์บางส่วน (Coefficient of Partial Correlation)	33
2.3.9 การตรวจสอบสมมติฐานของการวิเคราะห์การดัดออย.....	34
2.3.9.1 ค่า X ต้องเป็นค่าที่กำหนดไว้ล่วงหน้าหรือทราบค่า.....	34
2.3.9.2 ค่าความคลาดเคลื่อน e มีค่าเฉลี่ย = 0 $E(e) = 0$ หรือ $V(e) = \sigma^2$	34
2.3.9.3 การตรวจสอบ e_i และ e_j เป็นอิสระต่อกัน	35
2.3.9.4 การตรวจสอบว่า e_i มีการแจกแจงแบบปกติ	37
2.3.10 การตรวจสอบรูปแบบความสัมพันธ์ของตัวแปรเชิงปริมาณ 2 ตัว.....	37
บทที่ 3 วิธีการดำเนินการวิจัย	
3.1 วิธีการดำเนินงานของโครงข่ายประสาทเทียม	39
3.1.1 ข้อมูลจริงทั้งหมด	39
3.1.2 การจัดเตรียมข้อมูล.....	41

สารบัญ (ต่อ)

	หน้า
3.1.3 การทำนารูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม.....	42
3.1.3.1 จำนวนเซลล์ประสาทของตัวแปรด้านเข้าและชั้นตัวแปรด้านออก.....	42
3.1.3.2 ชนิดของฟังก์ชันถ่ายโอน.....	43
3.1.3.3 สถาปัตยกรรมของโครงข่ายประสาทเทียม.....	43
3.1.3.4 จำนวนเซลล์ประสาทของชั้นซ่อนที่เหมาะสม.....	45
3.1.3.5 การปรับค่าน้ำหนักและค่าไบอัส.....	45
3.1.3.6 การวัดประสิทธิภาพของโครงข่ายประสาทเทียม.....	46
3.2 การดำเนินงานของการวิเคราะห์ความถดถอยเชิงพหุ.....	46
บทที่ 4 ผลการดำเนินงาน	
4.1 การพยากรณ์รูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม.....	48
4.2 การพยากรณ์รูปร่างของรอยเชื่อมโดยใช้การวิเคราะห์ความถดถอยเชิงพหุ.....	51
4.2.1 Bead width.....	51
4.2.2 Penetration.....	56
4.2.3 Reinforcement height.....	59
4.2.4 Cross section.....	65
4.3 สรุปผลการพยากรณ์รูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียมและ การวิเคราะห์ความถดถอยเชิงพหุ.....	72
บทที่ 5 สรุปผลและข้อเสนอแนะ	
5.1 สรุปผลการวิจัย.....	80
5.1.1 การพยากรณ์รูปร่างรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม.....	80
5.1.2 การพยากรณ์รูปร่างของรอยเชื่อมโดยใช้การวิเคราะห์ความถดถอยเชิงพหุ.....	80
5.2 สิ่งที่ควรพัฒนาต่อไป.....	81
หนังสืออ้างอิง.....	82

สารบัญตาราง

	หน้า
ตารางที่ 2.1 การวิเคราะห์ความแปรปรวนของการวิเคราะห์เชิงพหุ	30
ตารางที่ 2.2 การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอย β_1 ในขั้นที่ 2.....	32
ตารางที่ 3.1 ข้อมูลจริงทั้งหมด 42 ข้อมูล	40
ตารางที่ 4.1 แสดงผลจากการฝึกสอนและทดสอบโครงข่ายประสาทเทียมที่มี 1 ชั้นซ่อน	49
ตารางที่ 4.2 แสดงผลจากการฝึกสอนและทดสอบโครงข่ายประสาทเทียมที่มี 2 ชั้นซ่อน	50
ตารางที่ 4.3 Variables Entered/Removed of Bead width	51
ตารางที่ 4.4 Model Summary(d) of Bead width.....	52
ตารางที่ 4.5 ANOVA(d) of Bead width.....	53
ตารางที่ 4.6 Coefficient(a) of Bead width	54
ตารางที่ 4.7 Variables Entered/Removed(b) of Penetration.....	56
ตารางที่ 4.8 Model Summary(b) of Penetration	56
ตารางที่ 4.9 ANOVA(b) of Penetration.....	57
ตารางที่ 4.10 Coefficient(a) of Penetration	58
ตารางที่ 4.11 Variables Entered/Removed(a) of Reinforcement height	59
ตารางที่ 4.12 Model Summary(e) of Reinforcement height	60
ตารางที่ 4.13 ANOVA(e) of Reinforcement height.....	62
ตารางที่ 4.14 Coefficient(a) of Reinforcement height.....	63
ตารางที่ 4.15 Variables Entered/Removed(a) of Cross section.....	65
ตารางที่ 4.16 Model Summary(e) of Cross section.....	66
ตารางที่ 4.17 ANOVA(e) of Cross section.....	68
ตารางที่ 4.18 Coefficient(a) of Cross section	70
ตารางที่ 4.19 แสดงค่าความผิดพลาดกำลังสองเฉลี่ยของรูปร่างรอยเชื่อมที่ได้จากการพยากรณ์ โดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ.....	72
ตารางที่ 4.20 แสดงค่าที่ได้จากการพยากรณ์โดยใช้การวิเคราะห์ความถดถอยเชิงพหุ	73
ตารางที่ 4.21 แสดงค่าความผิดพลาดกำลังสองเฉลี่ยของรูปร่างรอยเชื่อมที่ได้จากการพยากรณ์ โดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ.....	73

สารบัญรูป

	หน้า
รูปที่ 2.1 การเชื่อมแบบ Gas tungsten arc welding	3
รูปที่ 2.2 รูปร่างรอยเชื่อมแบบ Gas tungsten arc welding	3
รูปที่ 2.3 ลักษณะรอยเชื่อมที่เกิดจากการเชื่อมแบบ Gas tungsten arc welding	3
รูปที่ 2.4 ตัวอย่างของนิวรอนที่มีสองขั้ว	8
รูปที่ 2.5 โครงสร้างของประสาท	8
รูปที่ 2.6 เซลล์ประสาทเทียมที่มีอินพุตเดียว	10
รูปที่ 2.7 เซลล์ประสาทเทียมที่มีหลายอินพุต	11
รูปที่ 2.8 ฟังก์ชันถ่ายโอนชนิด Hard limit	13
รูปที่ 2.9 ฟังก์ชันถ่ายโอนชนิด Linear	13
รูปที่ 2.10 ฟังก์ชันถ่ายโอนชนิด Log-sigmoid	13
รูปที่ 2.11 ฟังก์ชันถ่ายโอนชนิด Tan-sigmoid	14
รูปที่ 2.12 โครงข่ายประสาทเทียมแบบชั้นเดียว	15
รูปที่ 2.13 โครงข่ายประสาทเทียมแบบหลายชั้น	15
รูปที่ 2.14 Recurrent Networks	16
รูปที่ 2.15 โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีสองชั้น	18
รูปที่ 2.16 ระบายถดถอยเชิงพหุที่มีตัวแปรตาม Y และตัวแปรอิสระ X_1 และ X_2	26
รูปที่ 2.17 กราฟแสดงความสัมพันธ์ระหว่างค่า e กับ \hat{Y}	35
รูปที่ 2.18 กราฟแสดงความสัมพันธ์ระหว่าง e_t กับ t	36
รูปที่ 2.19 การแจกแจงของ Durbin-Watson	37
รูปที่ 2.20 กราฟแสดงความสัมพันธ์ระหว่าง X และ Y	38
รูปที่ 3.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมที่ใช้ในการพยากรณ์	44
รูปที่ 3.2 ขั้นตอนการพยากรณ์รูปร่างของรอยเชื่อม โดยใช้การวิเคราะห์ความถดถอยเชิงพหุ	47
รูปที่ 4.1 Bead width ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง	74
รูปที่ 4.2 Penetration ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง	74
รูปที่ 4.3 Reinforcement height ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง	75
รูปที่ 4.4 Cross section ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง	75
รูปที่ 4.5 การกระจายของ Bead width ที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียม	76
รูปที่ 4.6 การกระจายของ Bead width ที่ได้จากการพยากรณ์โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ	76
รูปที่ 4.7 การกระจายของ Penetration ที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียม	77
รูปที่ 4.8 การกระจายของ Penetration ที่ได้จากการพยากรณ์โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ	77

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 4.9 การกระจายของ Reinforcement height ที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียม	78
รูปที่ 4.10 การกระจายของ Reinforcement height ที่ได้จากการพยากรณ์ โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ	78
รูปที่ 4.11 การกระจายของ Cross section ที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียม	79
รูปที่ 4.12 การกระจายของ Cross section ที่ได้จากการพยากรณ์โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ	79

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เนื่องจากปัจจุบันในงานติดตั้งโครงสร้างเหล็กและอุตสาหกรรมการผลิตต่างๆ มักจะมีการนำการเชื่อมมาเป็นส่วนหนึ่งในการสร้างหรือผลิต แต่การเชื่อมเป็นกระบวนการผลิตที่ซับซ้อนจึงทำให้ยากต่อการทำนายรูปร่างของรอยเชื่อม และการลงมือทดลองหรือเชื่อมจริงนั้นเป็นสิ่งที่ยุ่งยากและทำให้เกิดความสิ้นเปลือง จึงมีความพยายามในการสร้างแบบจำลองเพื่อใช้ในการทำนายรูปร่างของรอยเชื่อม การสร้างแบบจำลองโดยใช้ โครงข่ายประสาทเทียมมาช่วยในการทำนายค่าต่างๆ เป็นวิธีที่นิยมใช้กับระบบที่ซับซ้อน ดังนั้น ปริญญาพนธ์นี้จึงได้นำวิธี โครงข่ายประสาทเทียมมาทำนายรูปร่างของรอยเชื่อมที่เกิดขึ้นจากการเชื่อมที่เงื่อนไขต่างๆ ซึ่งผลที่ได้จากการทำนายจะช่วยให้สามารถเลือกใช้ปัจจัยในการเชื่อมให้เหมาะสม และเกิดประสิทธิภาพสูงสุดในการผลิตที่ต้องการได้ นอกจากนี้ผลที่ได้จากการทำนายด้วยวิธี โครงข่ายประสาทเทียมจะนำมาเปรียบเทียบกับผลการทำนายโดยใช้วิธี การวิเคราะห์ความถดถอยเชิงพหุ ซึ่งเป็นเทคนิคการทำนายที่นิยมใช้กัน โดยทั่วไปในการทำนายค่าต่างๆ

1.2 วัตถุประสงค์

1. เพื่อทำนายรูปร่างของรอยเชื่อมแบบไฟฟ้าโดยใช้ โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ
2. เพื่อเปรียบเทียบประสิทธิภาพของการทำนายด้วยวิธี โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ

1.3 ขอบเขตการศึกษา

1. ทำนายรูปร่างของรอยเชื่อมในรูปของ Bead width, Penetration, Reinforcement height และ Cross section
2. ทำนายรูปร่างของรอยเชื่อมที่เกิดขึ้นจากการเชื่อมแบบ Gas tungsten arc welding

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำเทคนิค โครงข่ายประสาทเทียมและ การวิเคราะห์ความถดถอยเชิงพหุ ไปทำนายลักษณะรอยเชื่อม
2. ได้รับความรู้ในการใช้ โครงข่ายประสาทเทียมและ การวิเคราะห์ความถดถอยเชิงพหุ ที่จะนำไปใช้ในงานลักษณะอื่นๆ ได้

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

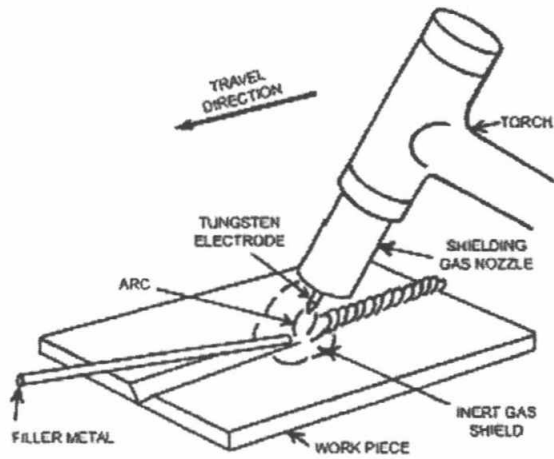
2.1 การเชื่อมแบบ Gas tungsten arc welding

การเชื่อมโลหะ (Welding) คือการทำให้โลหะหลอมละลายเข้าด้วยกันโดยใช้ความร้อน การเชื่อมมีหลายวิธีซึ่งแตกต่างกันตรงที่การเกิดความร้อนและวิธีการให้ความร้อนแก่ชิ้นงาน โลหะหลายชนิดสามารถเชื่อมประสานกันได้โดยใช้ความร้อนเพียงอย่างเดียว เรียกว่า การเชื่อมหลอมละลาย โดยที่ไม่ต้องเติมวัสดุประสาน หรือลวดเชื่อมก็ได้

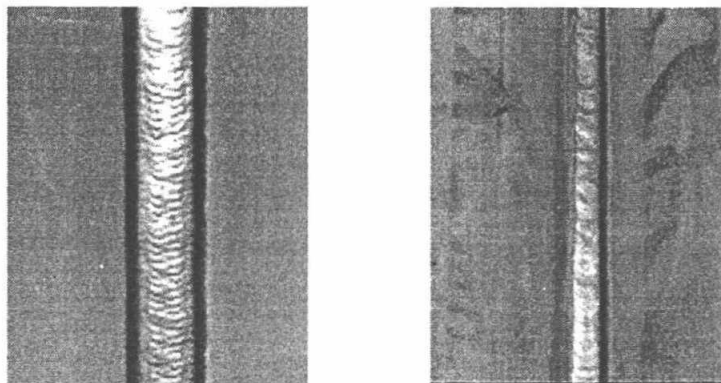
ในสมัยโบราณการหล่อโลหะจะเป็นการหลอมละลายโลหะแล้วเทลงบนแม่พิมพ์ ใช้ทำมีด คานโดยใช้วิธีการตีเหล็กที่ร้อนให้เชื่อมติดกัน งานเชื่อมโลหะได้ถูกค้นพบโดยชาวตะวันตก โดยการนำไปใช้ในการเชื่อมทองคำซึ่งในปัจจุบันเราเรียกวิธีการนี้ว่า การบัดกรีแข็ง จนกระทั่งพุทธศตวรรษที่ 25 (พ.ศ. 2401-2500) ได้มีการค้นพบการเชื่อมไฟฟ้าขึ้น โดยใช้กระแสไฟฟ้าแรงสูงผ่านโลหะ 2 ชิ้นแล้วนำมาแตะกันทำให้โลหะติดกันได้ ต่อมาในปีพ.ศ. 2433 นักเคมีชาวฝรั่งเศสค้นพบว่าสามารถนำแก๊สออกซิเจนผสมกับแก๊สอะเซทิลีน เมื่อนำมาจุดไฟได้อุณหภูมิสูงมากซึ่งเป็นกระบวนการเชื่อมที่นิยมกันกว้างขวางมาก

ปัจจุบันงานเชื่อมมีบทบาทอย่างมากในวงการอุตสาหกรรมการก่อสร้างและอุตสาหกรรมเหล็ก อุตสาหกรรมการผลิตและการซ่อมแซมทั่วไป โดยในปริญญาโทศึกษานี้ศึกษาเกี่ยวกับการทำนารูปปร่างรอยเชื่อมที่เกิดจากการเชื่อมแบบ Gas tungsten arc welding

Gas Tungsten Arc Welding (GTAW) หรือที่เรียกกันโดยทั่วไปว่าการเชื่อม TIG (Tungsten Inert Gas) เป็นกระบวนการเชื่อมที่ใช้แท่งอิเล็กโทรดแบบไม่สิ้นเปลือง (Non-consumable Electrode) ทำให้เกิดการอาร์คกระแสไฟฟ้าที่ ใช้ประมาณ 200-500 แอมป์ ขึ้นอยู่กับชนิดของวัสดุที่นำมาทำการเชื่อม เพื่อหลอมละลายชิ้นงานให้ติดกันแท่งอิเล็กโทรดที่ใช้ในการอาร์ค ทำจากทังสเตน และใช้แก๊สเฉื่อย ปกคลุมบ่อหลอมละลาย ในกรณีที่ต้องการเติมลวดเชื่อมต้องนำลวดเชื่อม (Filler Wire) จากภายนอกมาเติมในบ่อหลอมละลาย กระบวนการนี้ถูกนำมาประยุกต์ใช้งานโดยนิยมเรียกว่าการเชื่อมทิก หรือบางที่เรียกว่า การเชื่อมอาร์กอน เพราะในขณะที่ทำการเชื่อมส่วนใหญ่จะใช้แก๊สอาร์กอนเป็นแก๊สปกคลุมแนวเชื่อม คือ อาร์กอน และฮีเลียม หรือ แก๊สผสม ถ้าใช้ไฟกระแส AC จะใช้กับการเชื่อมงานที่เป็นอะลูมิเนียม และแมกนีเซียม เพราะว่า กระแสไฟ AC สามารถช่วยขจัดผิว Oxides ที่ผิวงานได้ และให้คุณภาพงานดี สามารถเชื่อมงานหนา ๆ ได้ เหมาะกับวัสดุงานที่เป็นอะลูมิเนียม และสแตนเลส ส่วนเหล็กหล่อ หรือ เหล็กอ่อน จะเชื่อมได้ยาก การเชื่อมด้วยกระบวนการนี้แนวเชื่อมที่ได้จะมีคุณภาพดี ไม่มีการกระเด็นของน้ำโลหะในระหว่างการอาร์ค จึงทำให้บริเวณผิวหน้างานและแนวเชื่อมสะอาด



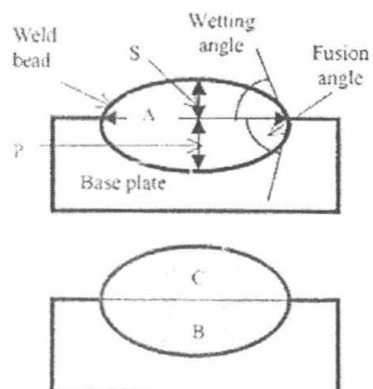
รูปที่ 2.1 การเชื่อมแบบ Gas tungsten arc welding



(a) Front view

(b) Back view

รูปที่ 2.2 รูปร่างรอยเชื่อมแบบ Gas tungsten arc welding



รูปที่ 2.3 ลักษณะรอยเชื่อมที่เกิดจากการเชื่อมแบบ Gas tungsten arc welding

จากรูปที่ 2.3 แสดงลักษณะรอยเชื่อมที่เกิดจากการเชื่อมแบบ Gas tungsten arc welding โดยที่ ระยะ S คือ Reinforcement ระยะ P คือ Penetration ระยะ A คือ Bead Width บริเวณ B คือ Penetration area และบริเวณ C คือ Reinforcement area

2.2 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Artificial Neural Networks, ANNs) เป็นแขนงหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence, AI) ซึ่งมีโครงสร้างเชิงสถาปัตยกรรมและการทำงานคล้ายคลึงกับการทำงานของเซลล์สมอง หรือระบบประสาทมนุษย์ โดยสามารถเรียนรู้และปรับตัวตามข้อมูลที่ได้รับจากสภาพแวดล้อมภายนอก และให้คำตอบในปัญหาที่ซับซ้อนได้นำเชื่อถือ

เนื่องจากคุณสมบัติที่สำคัญคือ ความยืดหยุ่นในการจำลองความสัมพันธ์ระหว่างปัจจัยนำเข้า (Input) และผลลัพธ์ (Output) ที่มีความซับซ้อนและสามารถตอบสนองกับข้อมูลที่ไม่เคยเห็นได้ดี ทำให้มีการนำทฤษฎีโครงข่ายประสาทเทียมมาประยุกต์ใช้แก้ปัญหาจริงหลายด้านอย่างกว้างขวาง เช่น การเงินการธนาคาร อวกาศ ระบบป้องกันประเทศ ระบบรักษาความปลอดภัย การแพทย์ ระบบสื่อสาร ระบบขนส่ง การบันเทิง รวมทั้งงานทางด้านวิศวกรรม โดยนำโครงข่ายประสาทเทียมมาใช้ในการจดจำรูปแบบ (Recognition) การจัดกลุ่ม (Clustering) การจัดหมวดหมู่ (Classification) และการพยากรณ์ (Forecasting) เป็นต้น (อนันต์ วิริยะกิจชัย, 2549) ในบทนี้จะอธิบายเกี่ยวกับหลักการของโครงข่ายประสาทเทียมที่ใช้ในการพยากรณ์รูปร่างรอยเชื่อม

2.2.1 ความเป็นมาของโครงข่ายประสาทเทียม

ความพยายามที่จะทำความเข้าใจการทำงานของสมองมนุษย์ได้มีมานานนับศตวรรษแล้ว เริ่มต้นในปี ค.ศ. 1890 William James นักจิตวิทยาชาวอเมริกัน ได้แต่งหนังสือชื่อ Psychology (Brief Course) ซึ่งอธิบายถึงการทำงานของสมอง และยังกล่าวถึงทฤษฎีที่มีอยู่ในปัจจุบันไว้ล่วงหน้าอีกด้วย (กูริสตันท์ ลักษณะ, 2548)

2.2.1.1 ช่วงทศวรรษที่ 40 ยุคเริ่มต้นของโครงข่ายประสาทเทียม

ในปี ค.ศ.1936 Alan Turing ได้นำสมองไปใช้เป็นแบบอย่างของการคำนวณใน A way of looking at the world of computing ในปี ค.ศ.1943 Warren McCulloch นักประสาทวิทยา และ Walter Pitts นักคณิตศาสตร์ ซึ่งมีอายุเพียง 18 ปี ในขณะนั้น ได้ออกแบบและเป็นที่ยอมรับโดยทั่วไปว่าเป็นโครงข่ายประสาทเทียมชนิดแรก นักวิจัยเหล่านั้นยอมรับว่า การรวมหลายๆนิวรอนไม่ซับซ้อนเข้าสู่ระบบโครงข่ายให้เป็นการเพิ่มกำลังในการคำนวณค่าต่าง ๆ นักของนิวรอน McCulloch – Pitts ได้กำหนดให้มีค่าคงที่ ดังนั้นนิวรอน McCulloch – Pitts ประกอบด้วยชั้นของโครงข่ายจำนวน 2 ชั้น คือชั้นอินพุตและเอาต์พุต โครงข่ายจะมีการเชื่อมโยงจากชั้นอินพุตไปยังชั้นเอาต์พุตโดยผ่านน้ำหนัก (Weight) ความคิดของค่าขีดจำกัด (threshold) กล่าวคือ ถ้าโครงข่ายอินพุตถึงนิวรอนมีค่าใหญ่กว่าค่าขีดจำกัดแล้ว โหนดกระตุ้น เป็นลักษณะอย่างหนึ่งในนิวรอน McCulloch – Pitts ต่อมาในปี ค.ศ.1949 Donald Habb นักจิตวิทยาของมหาวิทยาลัย McGill ได้ออกแบบกฎการเรียนรู้ชนิดแรกสำหรับโครงข่ายประสาทเทียม กฎการเรียนรู้ก็คือ ถ้าสองนิวรอนถูกกระตุ้นพร้อมกันแล้ว ความแข็งแรงของการเชื่อมโยงระหว่างพวกมันควรจะเพิ่มขึ้น และได้แต่งหนังสือชื่อ The organization of behavior ซึ่งกล่าวถึงการเชื่อมโยงกันของจิตวิทยาและสรีรศาสตร์

2.2.1.2 ช่วงทศวรรษที่ 50-60 ยุคทองของโครงข่ายประสาทเทียม

การพัฒนาด้านฮาร์ดแวร์และซอฟต์แวร์ ในช่วงทศวรรษที่ 1950 ทำให้การวิจัยด้านต่างๆเข้าสู่ยุคของการสร้างแบบจำลองด้วยคอมพิวเตอร์ เป็นผลให้การทดลองทฤษฎีที่เกี่ยวข้องกับโครงข่ายประสาทมีความเป็นไปได้มากยิ่งขึ้น ในปี ค.ศ.1956 Dartmouth Summer ได้เริ่มทำงานวิจัยด้านปัญญาประดิษฐ์ (Artificial intelligence : AI) ซึ่งเป็นแรงผลักดันให้เกิดการศึกษาทั้งในด้านปัญญาประดิษฐ์และโครงข่ายประสาทมากขึ้น ในปีค.ศ.1958 John Von Neumann เขียนหนังสือเรื่อง The Computer and the Brain ซึ่งได้นำเสนอฟังก์ชันการทำงานของเซลล์ประสาทโดยใช้รีเลย์ที่ใช้สำหรับส่งโทรเลขและหลอดสุญญากาศ (Vacuum Tube) และ ในปีค.ศ.1957 Frank Rosenblatt ได้เริ่มศึกษาและพัฒนาประเภทใหญ่ของโครงข่ายประสาทเทียมที่เรียกว่า Perceptron ซึ่งเป็นรูปแบบที่สำคัญรูปแบบหนึ่งของโครงข่ายประสาทเทียม โดยส่วนมากสถาปัตยกรรมของโครงข่ายประสาทเทียมชนิดนี้จะประกอบด้วยชั้นอินพุท เชื่อมโยงผ่านด้วยค่าถ่วงน้ำหนัก เพื่อเป็นตัวเชื่อมกับนิวรอนที่ชั้นเอาต์พุท ค่าถ่วงน้ำหนักบนทางผ่านเชื่อมโยงถูกปรับเปลี่ยนกฎการเรียนรู้ Perceptron ปรับเปลี่ยนค่าถ่วงน้ำหนักซ้ำๆกัน ซึ่งมีประสิทธิภาพดีกว่ากฎของ Habb การเรียนรู้ของ Perceptron สามารถแสดงให้เห็นว่า การรู้เข้าสู่ค่าถ่วงน้ำหนักที่ถูกต้อง ค่าถ่วงน้ำหนักเหล่านั้นจะแก้ปัญหาคือเตรียมพร้อม และต่อมางานของ Frank Rosenblatt ได้อธิบาย Perceptron หลายๆชนิดว่า เหมือนนิวรอนได้พัฒนาจาก McCulloch และ Pitts และจาก Habb โครงข่ายประสาทเทียมชนิด Perceptron ใช้ค่าเริ่มต้นฟังก์ชันเอาต์พุท

ในปีค.ศ.1959 Bernard Widrow และนักศึกษาของเขา Marcian Hoff ได้พัฒนาการเรียนรู้ซึ่งใกล้เคียงกับกฎการเรียนรู้ของ Perceptron โดยการเรียนรู้ของ Perceptron ปรับค่าถ่วงน้ำหนักที่เชื่อมโยงที่โหนด เมื่อผลตอบสนองของโหนดไม่ถูกต้อง กฎ Della ปรับค่าถ่วงน้ำหนักเพื่อลดความแตกต่างระหว่างโครงข่ายอินพุทถึงโหนดเอาต์พุทและค่าเอาต์พุทที่ต้องการ สิ่งเหล่านี้จะสรุปในรูปของค่าผิดพลาดยกกำลังสองเฉลี่ยที่น้อยที่สุด (Smallest Mean Squared Error : SMS) และนำสิ่งต่างๆเหล่านี้ไปใช้สร้างแบบจำลองของ MADALINE (Multiple Adaptive Linear Element) ซึ่งพัฒนามาจาก ADALINE และเป็นกรนำเอาโครงข่ายประสาทมาใช้งานในทางปฏิบัติครั้งแรก โดยนำมาใช้ในการลดเสียงก้องในสายโทรศัพท์

2.2.1.3 ช่วงทศวรรษที่ 70 ยุคเงิบของโครงข่ายประสาทเทียม

ในปีค.ศ.1969 Marvin Minsky และ Seymour Papert ได้ตีพิมพ์หนังสือ Perceptron ซึ่งกล่าวถึงข้อจำกัดของ Perceptron ว่าไม่สามารถนำไปใช้ในการแก้ปัญหาหลายๆปัญหาที่มีความน่าสนใจได้ ซึ่งจุดนี้เองที่ทำให้นักวิจัยหลายคนกลับไปให้ความสนใจด้านปัญญาประดิษฐ์(AI) มากขึ้น แต่ยังมีนักวิจัยบางกลุ่มที่ไม่ย่อท้อและยังคงทำงานวิจัยด้านโครงข่ายประสาทต่อไป John Anderson ของมหาวิทยาลัย Brown เริ่มงานวิจัยของเขาในปี ค.ศ.1968 โดยทำการวิจัยโครงข่ายประสาทเทียมด้วยโครงข่ายรวมกลุ่มความสามารถในการจดจำ (Associative Memory Nets) เขาได้พัฒนาแบบจำลองของโครงข่ายประสาทขึ้นมา และให้ชื่อว่า Brian-State-in-A-Box (BSB) งานประยุกต์ที่ใช้โครงข่ายประสาทเทียมชนิดนี้คือ การวินิจฉัยทางการแพทย์ (Medical Diagnosis)

ในปี ค.ศ.1972 งานเริ่มต้นของ Teuvo Kohonen ของมหาวิทยาลัยเทคโนโลยี Helsinki วิศวกรไฟฟ้าชาวฟินแลนด์ ได้จัดการกับโครงข่ายประสาทรวมกลุ่มความสามารถในการจดจำ และต่อมาได้พัฒนาหลาย เป็น Self Organizing Map ขึ้นมา สถาปัตยกรรมของโครงข่ายประสาทเทียมชนิด Self Organizing Map นี้ ประกอบด้วยจำนวนชั้น 2 ชั้น คือ ชั้นอินพุทและเอาต์พุท การเชื่อมโยงที่ชั้นอินพุทไปยังนิวรอนที่ชั้นเอาต์พุทนั้นจะต่อผ่านน้ำหนัก ข้อแตกต่างจากโครงข่ายอื่นๆคือกระบวนการเรียนรู้ที่สามารถจัดกลุ่มข้อมูลอินพุทได้ด้วยตัวเองและเรียกการเรียนรู้นี้ว่า

การเรียนรู้แบบคอมเพทิทีฟ (Competitive Learning) โครงข่ายประสาทเทียมชนิดนี้ถูกนำไปใช้เกี่ยวกับการจดจำคำพูด (Speech Recognition) , การแก้ปัญหของ “Traveling Salesman Problem” เป็นต้น

2.2.1.4 ช่วงทศวรรษที่ 80 ยุคความกระตือรือร้นกลับมาใหม่ของโครงข่ายประสาทเทียม

ในปีค.ศ.1982 สหรัฐอเมริกาและญี่ปุ่นได้ร่วมมือกันจัด Conference on Cooperative / Competitive Neural Network ขึ้นที่ญี่ปุ่น ซึ่งเป็นจุดเริ่มต้นของการประชุมสัมมนาทางด้านโครงข่ายประสาท ในปี ค.ศ.1985 American Institute of Physics ได้จัดงาน Neural Network for Computing Meeting ในปี ค.ศ.1987 Institute of Electrical and Electronic Engineers (IEEE) ได้จัดประชุมนานาชาติเกี่ยวกับโครงข่ายประสาทเทียมขึ้นเป็นครั้งแรก โดยมีผู้ร่วมงานมากกว่า 1,800 คน ในปีเดียวกัน International Neural Network Society (INNS) ถูกก่อตั้งขึ้นโดยมี Grossberg จากสหรัฐอเมริกา Kohonen จากฟินแลนด์และ Amari จากญี่ปุ่นเป็นผู้นำในการก่อตั้ง โดยหลังจากที่ INNS ก่อตั้งได้ 2 ปี ก็มีสมาชิกเข้าร่วมมากกว่า 3,000 คนจากทั่วโลก

1. โครงข่ายประสาทเทียมชนิด Back-Propagation

สองเหตุผลหลักที่ทำให้เกิดยุคเรียบของโครงข่ายประสาทเทียมคือ ความผิดพลาดของ Perceptron ชั้นเดียวที่ไม่สามารถแก้ปัญหาที่ไม่ซับซ้อนได้เช่น ฟังก์ชัน XOR และอีกเหตุผลก็คือ ไม่มีวิธีฝึกสอนโครงข่ายประสาทเทียมหลายชั้น วิธีสำหรับการแพร่ข้อมูลค่าผิดพลาดที่โหนดเอาต์พุตย้อนกลับไปที่ชั้นซ่อน ได้ถูกค้นพบโดย Paul Werbos ในปี ค.ศ.1974 แต่ไม่ได้กระจายใช้อย่างกว้างขวาง วิธีนี้ถูกค้นพบอีกครั้งในปี ค.ศ.1985 โดย David Parker และ Lecun ก่อนที่จะกลายเป็นที่รู้จักกันอย่างกว้างขวาง เปรียบเสมือนเป็นวิธีคิดแรกเริ่มในทฤษฎีควบคุมอย่างเหมาะสม (Optional Control Theory) ในปี ค.ศ.1986 งานของ Parker กลายเป็นที่สนใจของกลุ่ม Parallel Distributed Processing ซึ่งนำทีมโดยนักวิจัย David Rumelhart ของมหาวิทยาลัย California ที่ Sandiego และ James McClelland ของมหาวิทยาลัย Carnegie-Mellon ผู้ซึ่งนำวิธีดังกล่าวมาปรับปรุงให้ดีขึ้นและถูกใช้อย่างกว้างขวางในเวลาต่อมา หลักการของโครงข่ายประสาทเทียมชนิดนี้สามารถนำมาใช้แก้ปัญหาในลักษณะเชิงเส้น และปัญหาที่ไม่เชิงเส้นได้ ทำให้ประยุกต์เข้ากับโครงการต่างๆ ได้อย่างหลากหลาย

2. โครงข่ายประสาทเทียมชนิด Hopfield

ในปี ค.ศ.1982 John Hopfield ทำงานที่สถาบันเทคโนโลยี California พร้อมกับ David Tank นักวิจัยที่ AT&T Hopfield ได้พัฒนาโครงข่ายประสาทเทียมที่ประสบความสำเร็จเป็นอย่างดีในนามโครงข่ายประสาทเทียมชนิด Hopfield ส่งผลให้ John Hopfield ได้รับรางวัลโนเบล สาขาฟิสิกส์และได้พัฒนาจำนวนโครงข่ายประสาทเทียมพื้นฐานค่าถ่วงน้ำหนักคงที่ และ Adaptive Activation โครงข่ายประสาทเทียมเหล่านั้นสามารถมีประโยชน์เหมือนกับโครงข่ายประสาทเทียมรวมกลุ่มความสามารถในการจดจำและสามารถนำไปใช้เพื่อแก้ปัญหการบีบบังคับความพอใจ เช่น “Traveling Salesman Problem” และเขาได้นำเสนอบทความต่อ National Academy of Sciences ซึ่งมีเนื้อหาเกี่ยวกับการใช้โครงข่ายประสาทเทียมในการทำนายส่วนที่ขาดหายไปของข้อมูลหรือรูปภาพ

3. โครงข่ายประสาทเทียมชนิด Neocognitron

ในปีค.ศ.1975-1988 Kunihiko Fukushima และเพื่อนร่วมงานที่ห้องทดลอง NHK ในกรุงโตเกียว ได้พัฒนาตระกูลโครงข่ายประสาทเทียมที่ใช้งานเฉพาะสำหรับการจดจำตัวอักษรหนึ่งในตัวอย่างของโครงข่ายประสาทเทียมกลุ่มนี้คือ Neocognitron โครงข่ายประสาทเทียมชนิด Self-Organizing ในช่วงแรกถูกเรียกว่า Cognitron แต่เนื่องจากมีความผิดพลาดเกี่ยวกับสถานะการจดจำ ข้อบกพร่องเหล่านี้ได้ถูกแก้ไขด้วยโครงข่ายประสาทเทียมชนิด Neocognitron

4. โครงข่ายประสาทเทียมชนิด Boltzmann

ในช่วงปีค.ศ.1983-1987 นักวิจัยหลายท่านได้พัฒนาการหาค่าเชื่อมโยงและฟังก์ชันกระตุ้น โดยใช้ฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function) ต่างจากเดิมที่มักจะใช้แบบ Nondeterministic โดยพัฒนาโครงข่ายประสาทเทียมในนาม Boltzmann ซึ่งอาศัยแนวคิดการจำลองการอบอุ่นตัวลงสู่สถานะเสถียรภาพ (Simulated Annealing) และทฤษฎีการตัดสินใจของเบย์เซียน (Bayesian Decision Theory)

2.2.2 เซลล์ประสาททางชีววิทยา

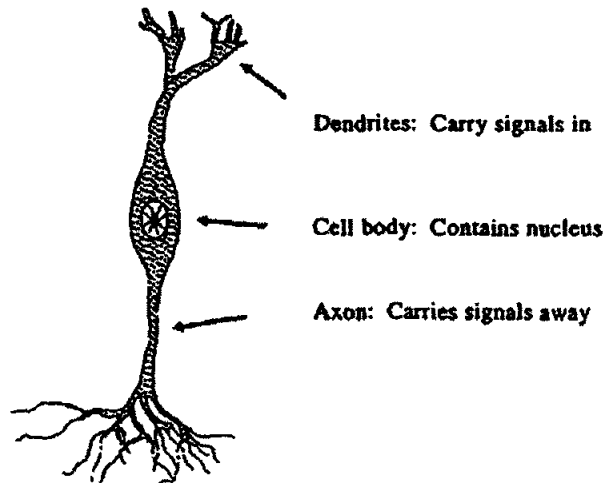
สมองมนุษย์ประกอบด้วยเซลล์ประสาท (Neuron) จำนวนมากประมาณ 10^{11} ยูนิต ที่มีการเชื่อมต่อกันอย่างแน่นหนา (การเชื่อมโยงของแต่ละเซลล์ประสาทประมาณ 10^4 ยูนิต) โดยแต่ละเซลล์ประสาทจะประกอบกันด้วยเดนไดรต์ (Dendrites) ซึ่งทำหน้าที่รับความรู้สึกจากเส้นใยประสาท โดยส่งข้อมูลเป็นสัญญาณไฟฟ้า (Electrical Signal) มาที่เดนไดรต์ และที่เดนไดรต์แต่ละกิ่งก็จะรับรู้ด้วยน้ำหนัก (Weight) ที่ต่างกันและการแทนค่าด้วยน้ำหนักด้วยความแข็งแรง (Strength) ของแต่ละไซแนปส์ (Synapse) จากนั้นจึงส่งข้อมูลตัวเซลล์ (Cell Body or Soma) ซึ่งทำหน้าที่รวบรวมสิ่งที่ได้รับรู้แล้วส่งให้แอกซอน (Axon) แอกซอนจะส่งสัญญาณออกไป โดยสัญญาณที่ส่งออกมาจะเป็นฟังก์ชันของผลรวมของสิ่งที่ได้รับรู้จากตัวเซลล์

2.2.2.1 นิวรอน

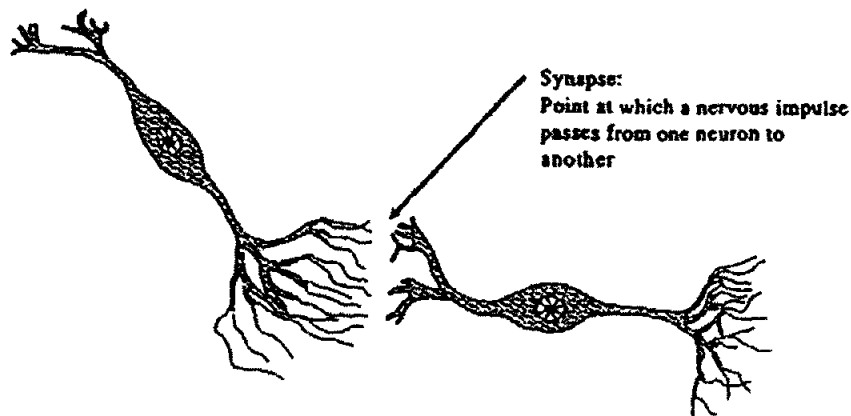
นิวรอน หมายความว่า เซลล์ประสาทและกระบวนการทั้งหมดของเซลล์ประสาท รูปที่ 2.4 แสดงตัวอย่างของนิวรอนที่สามารถพบได้ในเรตินา เป็นนิวรอนที่มีสองขั้ว (Bipolar Neural) ในตัวเซลล์ (Cell Body) ของนิวรอนจะมีนิวเคลียส (Nucleus) อยู่ตรงกลางเซลล์และมีแขนงอย่างน้อย 1 เส้นเชื่อมต่อกับเซลล์ แขนงที่ทำหน้าที่นำกระแสประสาท ซึ่งมีลักษณะเป็นสัญญาณอิมพัลส์ (Impulse Signal) เข้ามายังตัวเซลล์นั้นเรียกว่า เดนไดรต์ (Dendrite) ส่วนแขนงที่ทำหน้าที่นำกระแสประสาทออกจากตัวเซลล์เรียกว่า แอกซอน (Axon)

2.2.2.2 โครงสร้างของนิวรอนและไซแนปส์

เมื่อนิวรอนมากกว่า 1 นิวรอน อยู่รวมด้วยกันมีการจัดเรียงเป็นโครงสร้างของประสาทดังแสดงในรูปที่ 2.5 จะมีการส่งกระแสประสาทจากนิวรอนหนึ่งไปยังนิวรอนอื่น โดยผ่านจุดเชื่อมต่อระหว่างปลายของแอกซอนของนิวรอนหนึ่งกับตัวเซลล์หรือเดนไดรต์ของอีกนิวรอนหนึ่ง หรือเรียกว่า ไซแนปส์ (Synapse) ซึ่งเป็นช่องว่างที่เล็กมากระหว่างนิวรอนสองนิวรอนที่อยู่ติดกันแต่ไม่ได้สัมผัสกัน โดยตรง ไซแนปส์จะเป็นบริเวณที่ควบคุมให้มีการส่งผ่านกระแสประสาทต่อไปหรือยับยั้งการส่งกระแสประสาท รวมถึงเปลี่ยนแปลงลักษณะของกระแสประสาทอีกด้วย



รูปที่ 2.4 ตัวอย่างของนิวรอนที่มีสองขั้ว



รูปที่ 2.5 โครงสร้างของประสาท

2.2.2.3 การส่งผ่านกระแสประสาท

ในสภาวะปกติ เมื่อนิวรอนกระแสประสาทที่ได้รับจากอวัยวะที่ทำหน้าที่รับรู้ (อาทิเช่น ตา หรือ หู) ไปยังอวัยวะที่ทำหน้าที่ตอบสนอง (อาทิเช่น กล้ามเนื้อ หรือ ต่อมต่างๆภายในร่างกาย) กระแสประสาทจะถูกส่งผ่านจากนิวรอนหนึ่งไปยังอีกนิวรอนหนึ่งโดยผ่านไซแนปส์ เมื่อกระแสประสาทผ่านเข้าไปในไซแนปส์ (เสมือนเป็นอินพุตของไซแนปส์) สารเคมีในไซแนปส์จะมีผลต่อกระแสประสาทได้สองลักษณะ คือ การกระตุ้นให้กระแสประสาทมีขนาดเพิ่มขึ้น หรือยับยั้งให้กระแสประสาทขนาดลดลง ซึ่งทั้งสองลักษณะดังกล่าวเปรียบได้กับการที่สัญญาณอินพุตถูกถ่วงน้ำหนักให้มีค่ามากขึ้นหรือลดลง เมื่อผ่านออกจากไซแนปส์กระแสประสาทที่ถูกกระตุ้นและกระแสประสาทที่ถูกยับยั้งจะถูกรวมเข้าด้วยกัน หากผลรวมของกระแสประสาทดังกล่าวมีค่ามากกว่าหรือเท่ากับศักยค่าที่ผิวของนิวรอนถัดไป หรือ เธรชโฮลด์(Threshold) กระแสประสาทดังกล่าวก็จะสามารถไหลผ่านเข้าไปในนิวรอนได้ หากผลรวมของกระแสประสาทมีค่าน้อยกว่าค่าเธรชโฮลด์ กระแสประสาทก็จะถูกยับยั้งไม่ให้ผ่านเข้าไปในนิวรอนนั้น การไหลของกระแส

ประสาทจะเป็นไปในทิศทางเดียว (One Direction) และมีความเร็วอยู่ระหว่าง 10 ถึง 120 เมตรต่อวินาที เมื่อการส่งกระแสประสาทครั้งหนึ่งสิ้นสุดลง ส่วนต่างๆของนิวรอนจะกลับเข้าสู่ภาวะเริ่มต้นอีกครั้งหนึ่ง เพื่อเตรียมพร้อมสำหรับส่งกระแสประสาทครั้งต่อไป

อาจสรุปคุณสมบัติของโครงข่ายประสาทชีวภาพได้ดังนี้

- แต่ละนิวรอนจะรับข้อมูลอินพุตเป็นจำนวนมาก และแต่ละข้อมูลอินพุตอาจถูกปรับให้เหมาะสมด้วยค่าการเชื่อมโยง (Connection Value of Weight Value of Strength) ที่ไซแนปส์ จากนั้นนำมารวมกันและทำการประมวลผล
- ภายใต้อาหารที่เหมาะสม เช่น ข้อมูลอินพุตจำนวนเพียงพอ นิวรอนจะส่งข้อมูลออกมาค่าหนึ่ง
- ข้อมูลเอาต์พุตจากนิวรอนหนึ่งอาจจะถูกส่งไปที่นิวรอนอื่นๆหลายนิวรอนผ่านทางแอกซอน ซึ่งแต่ละนิวรอนจะมีค่าเป็นเฉพาะตำแหน่งหรือเฉพาะที่ (Local)
- ความจำแบ่งออกเป็น 2 ส่วน คือ ความจำระยะยาว จะเก็บที่ไซแนปส์ของนิวรอนหรือเก็บในค่าการเชื่อมโยงและความจำระยะสั้น สัมพันธ์กับข้อมูลเอาต์พุตจากนิวรอน
- ความถูกต้องของการเชื่อมโยงขึ้นกับประสบการณ์ในการเรียนรู้ที่ผ่านมา
- การควบคุมการถ่ายโอนของไซแนปส์อาจจะเป็นการกระตุ้นแบบเสริมหรือแบบหักล้าง

นอกจากที่กล่าวมา โครงข่ายประสาทชีวภาพยังยึดหยุ่นต่อความบกพร่อง (Fault-Tolerance) 2 ประเด็นด้วยกัน ประการแรกคือ ความสามารถในการจดจำและจำแนกข้อมูลอินพุตที่แตกต่างจากข้อมูลอินพุตที่เคยได้รับไม่มากนัก ตัวอย่างเช่น ความสามารถในการจำแนกภาพบุคคลที่ไม่เคยเห็นมาก่อนหรือสามารถจดจำบุคคลที่ไม่ได้พบเห็นกันเป็นเวลานาน ประการที่สองสามารถในการซ่อมแซมหรือชดเชยความเสียหายของตัวนิวรอนเอง ในสมองมนุษย์เมื่อนิวรอนหนึ่งชำรุดเสียหายก็จะมีการสร้างนิวรอนใหม่ขึ้นมาแทนที่ นั่นคือยังคงมีการเรียนรู้อย่างต่อเนื่องแม้ว่าจะสูญเสียนิวรอนบางส่วน ในบางครั้งนิวรอนที่เหลือจะได้รับการฝึกสอนเพื่อทำหน้าที่แทนนิวรอนที่ชำรุดเสียหาย

จากจุดเด่นหลายประการของโครงข่ายประสาทชีวภาพผนวกกับเทคโนโลยีด้านคอมพิวเตอร์การทำงานแต่ละครั้งของอุปกรณ์สร้างกึ่งตัวนำในคอมพิวเตอร์จะใช้เวลาอยู่ในช่วงนาโนวินาที (10^{-9} s) ซึ่งมีความรวดเร็วมากเมื่อเปรียบเทียบกับการทำงานของสมองมนุษย์ ทำให้มีการคิดค้นแนวทางแก้ปัญหา โดยมีโครงสร้างองค์ประกอบและกลไกการแก้ปัญหาลักษณะเดียวกับโครงข่ายประสาทชีวภาพ โดยเรียกว่า โครงข่ายประสาทเทียม (Artificial Neural Networks) (ภูริทัศน์ ลักขชร, 2548)

2.2.3 เซลล์ประสาทเทียม

เมื่อทราบหลักการการทำงานของเซลล์ประสาททางชีววิทยาแล้ว ก็สามารถสร้างแบบจำลองเซลล์ประสาทที่เรียกว่า เซลล์ประสาทเทียม (Artificial Neuron), ยูนิต (Unit) หรือ โหนด (Node) ซึ่งเซลล์ประสาทที่ได้ถูกพัฒนาขึ้นจะประมวลผลข้อมูลโดยอาศัยหลักการทำงานพื้นฐานของเซลล์ประสาททางชีววิทยาที่สามารถรวบรวมความรู้ได้ด้วยการผ่านกระบวนการเรียนรู้ และจะเก็บความรู้เหล่านั้นไว้ในโครงข่าย ในรูปของค่าน้ำหนักและค่าไบอัส ซึ่งสามารถปรับเปลี่ยนได้เมื่อมีการเรียนรู้สิ่งใหม่ โดยโครงข่ายประสาทเทียมจะมีความเร็วในการประมวลผลที่สูงกว่าโครงข่ายประสาทเทียมทางชีววิทยามาก แต่อย่างไรก็ตาม สมองมนุษย์ก็สามารถทำงานหลายๆงานที่ซับซ้อนพร้อมๆกันได้เร็วกว่าคอมพิวเตอร์ เป็นเพราะโครงข่ายประสาททางชีววิทยานั้นมีโครงสร้างที่ขนานกันอย่างมาก จึงสรุปได้ว่า โครงข่ายประสาทเทียมเป็นอัลกอริทึมที่เลียนแบบการเรียนรู้ของสมองมนุษย์ โดยมีองค์ประกอบของโครงสร้างพื้นฐานที่สำคัญ

ดังนั้น คือ ยูนิต (Unit), ตัวแปรอินพุต (Input Layer) , ตัวแปรเอาต์พุต (Output Layer) และค่าถ่วงน้ำหนัก (Weighted Value) ซึ่งสามารถสรุปความสัมพันธ์ระหว่างเซลล์ประสาทกับเซลล์ประสาทเทียมได้ดังนี้

เซลล์ประสาทเทียมทางชีววิทยา

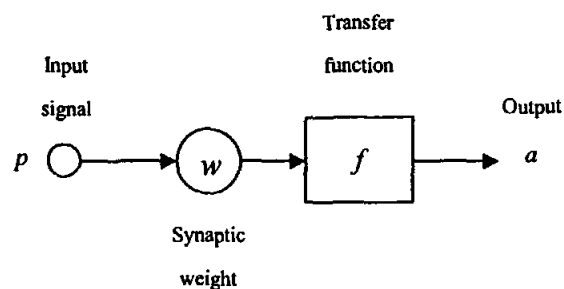
- ตัวเซลล์ (Cell Body)
- เดนไดรต์ (Dendrite)
- แอกซอน (Axon)
- ไซแนปส์ (Synapse)

เซลล์ประสาทเทียม

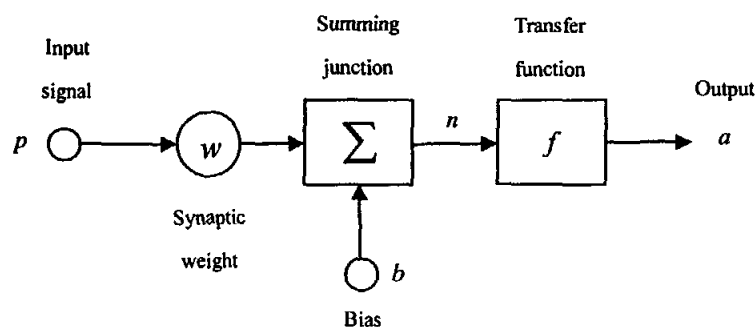
- ยูนิต (Unit)
- ตัวแปรด้านเข้า (Input)
- ตัวแปรด้านออก (Output)
- ค่าน้ำหนัก (Weight)

2.2.3.1 เซลล์ประสาทเทียมอินพุตเดียว

เซลล์ประสาทเทียมอินพุตเดียวดังแสดงในรูปที่ 2.6 (a) จะเป็นแบบไม่มีค่าไบอัส ส่วนในรูปที่ 2.6 (b) จะเป็นแบบที่มีค่าไบอัส โดยตัวแปรด้านเข้า p จะถูกส่งผ่านเข้าไประวมกับค่าน้ำหนัก (w) และค่าไบอัส (b) ซึ่งสามารถปรับค่าได้โดยใช้กฎการเรียนรู้ (Learning Rule) จากนั้นก็จะส่งสัญญาณ n ผ่านฟังก์ชันถ่ายโอน (f) เพื่อคำนวณค่าตัวแปรด้านออกของเซลล์ประสาทเทียม หรือ ยูนิต a



(a) แบบไม่มีค่าไบอัส



(b) แบบมีค่าไบอัส

รูปที่ 2.6 เซลล์ประสาทเทียมที่มีอินพุตเดียว

2.2.3.2 เซลล์ประสาทเทียมหลายอินพุท

จากรูปที่ 2.7 ตัวแปรด้านเข้า p_i โดยที่ $i = 1, 2, \dots, R$ จะถูกส่งผ่านเข้าไปรวมกับค่าน้ำหนัก (w) และค่าไบอัส (b) ที่สามารถปรับค่าได้โดยใช้กฎการเรียนรู้ ซึ่งจะ ได้ผลรวมเป็น

$$n_k = w_{k1}p_1 + w_{k2}p_2 + \dots + w_{kR}p_R + b_k$$

$$= \sum_{i=1}^R w_{ki}p_i + b_k \quad (2.1)$$

หลังจากนั้นจะส่งสัญญาณ n_k ผ่านฟังก์ชันถ่ายโอน (f) เพื่อคำนวณค่าตัวแปรด้านออกของเซลล์ประสาทเทียม หรือยูนิต a_k

$$a_k = f \left[\sum_{i=1}^R w_{ki}p_i + b_k \right] \quad (2.2)$$

กำหนดให้

p_i คือ ตัวแปรด้านเข้า โดยที่ $i = 1, 2, \dots, R$

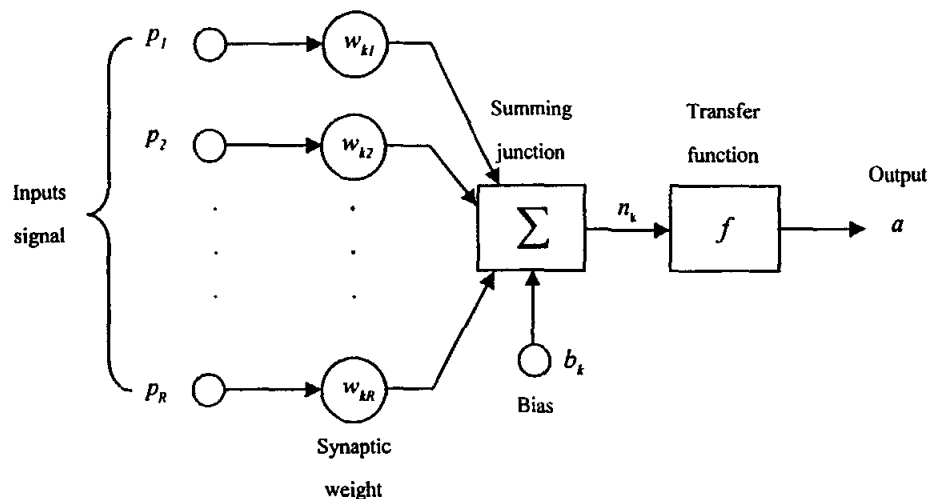
n_k คือ ผลรวมของตัวแปรด้านเข้าของยูนิตที่ k

w_{ki} คือ ค่าน้ำหนักระหว่างยูนิตที่ i และ k

b_k คือ ค่าไบอัสของยูนิตที่ k

f คือ ฟังก์ชันถ่ายโอน

a_k คือ ตัวแปรด้านออกของเซลล์ประสาทเทียมของยูนิตที่ k



รูปที่ 2.7 เซลล์ประสาทเทียมที่มีหลายอินพุท

2.2.3.3 คุณสมบัติและความสามารถของโครงข่ายประสาทเทียม

คุณสมบัติและความสามารถของโครงข่ายประสาทเทียม สามารถสรุปได้ดังนี้

1. โครงข่ายประสาทเทียมไม่ใช่โปรแกรมคอมพิวเตอร์ แต่จะโครงข่ายจะเรียนรู้จากตัวอย่าง
2. มีความยืดหยุ่นสูงจนสามารถจำลองกระบวนการของปัญหาใดๆ ได้
3. มีความสามารถในการจำชุดของค่าอินพุท เอาท์พุทที่มีความซับซ้อนมากจนไม่สามารถจำลองแบบในเชิงความน่าจะเป็นได้
4. มีความสามารถในการปรับตัวเข้ากับการเปลี่ยนแปลงของสิ่งแวดล้อม
5. มีความสามารถในการตอบสนองต่อข้อมูลที่ไม่เคยเห็น
6. ความรู้กระจายอยู่ทั่ว โครงสร้างของโครงข่ายประสาทเทียม

2.2.3.4 ลักษณะปัญหาที่เหมาะสมกับโครงข่ายประสาทเทียม

ลักษณะปัญหาที่เหมาะสมกับโครงข่ายประสาทเทียม มีดังนี้

1. ตัวอย่างของข้อมูลสำหรับการฝึกสอน ประกอบด้วยองค์ประกอบหลายๆค่า
2. เอาท์พุทที่ต้องการสามารถเป็นค่าจำนวนเต็ม หรือจำนวนจริง หรือเป็นเวกเตอร์ของจำนวนเต็มหรือจำนวนจริงก็ได้
3. ตัวอย่างของข้อมูลสามารถมีความคิดแปลปะปนอยู่ได้
4. ไม่มีปัญหาเกี่ยวกับขั้นตอนการฝึกหัดซึ่งอาจใช้เวลานานมาก
5. ต้องการประเมินผลเอาท์พุทอย่างรวดเร็ว
6. ไม่ต้องการเข้าใจวิธีการแก้ปัญหา ต้องการให้ปัญหาได้รับการแก้ไขเท่านั้น

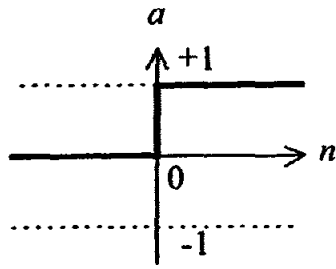
2.2.4 ชนิดของฟังก์ชันถ่ายโอน

ฟังก์ชันถ่ายโอน (Transfer function) หรือ ฟังก์ชันกระตุ้น (Activation function) จะเป็นตัวกำหนดค่าเอาท์พุทของโครงข่ายประสาทเทียม โดยปกติฟังก์ชันถ่ายโอนที่จะนำไปใช้กับโครงข่ายประสาทเทียมนั้น จะเลือกโดยผู้ออกแบบ ซึ่งอาจเป็นฟังก์ชันเชิงเส้น (Linear) หรือ ไม่เชิงเส้น (Nonlinear) ก็ได้ โดยเฉพาะอย่างยิ่งฟังก์ชันถ่ายโอนที่เลือกมาใช้งานนั้นจะต้องเพียงพอกับลักษณะของปัญหาที่โครงข่ายประสาทเทียมกำลังพยายามหาคำตอบอยู่ (โอพารธรรมชาติ, 2548)

ฟังก์ชันถ่ายโอนนั้นมีอยู่หลายชนิด แต่ที่นิยมใช้กันมากจะมีอยู่ 4 ชนิด ได้แก่

1. ฟังก์ชันถ่ายโอนชนิด Hard limit

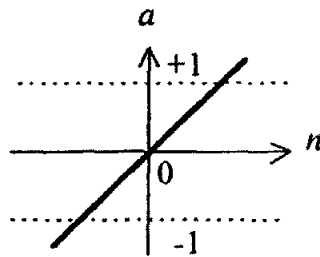
$$a(n) = \begin{cases} 0 & \text{เมื่อ } n < 0 \\ 1 & \text{เมื่อ } n \geq 0 \end{cases} \quad (2.3)$$



รูปที่ 2.8 ฟังก์ชันถ่ายโอนชนิด Hard limit

2. ฟังก์ชันถ่ายโอนชนิด Linear

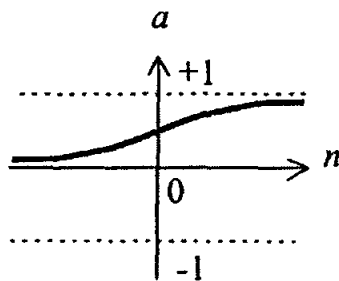
$$a(n) = n \tag{2.4}$$



รูปที่ 2.9 ฟังก์ชันถ่ายโอนชนิด Linear

3. ฟังก์ชันถ่ายโอนชนิด Log-sigmoid

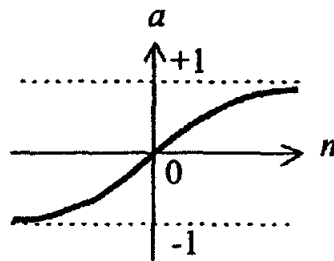
$$a(n) = \frac{1}{1 + e^{-n}} \tag{2.5}$$



รูปที่ 2.10 ฟังก์ชันถ่ายโอนชนิด Log-sigmoid

4. ฟังก์ชันถ่ายโอนชนิด Tan-sigmoid

$$a(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \quad (2.6)$$



รูปที่ 2.11 ฟังก์ชันถ่ายโอนชนิด Tan-sigmoid

2.2.5 โครงสร้างของโครงข่ายประสาทเทียม

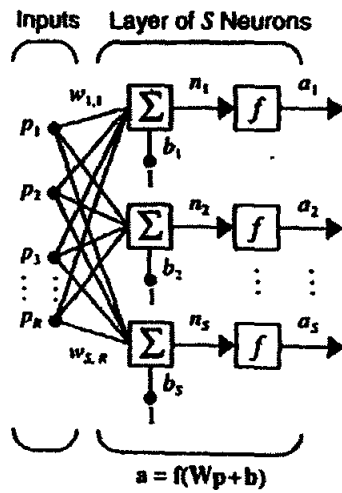
โครงข่ายประสาทเทียม ประกอบด้วยหน่วยประมวลผลย่อยที่เรียกว่า เซลล์ประสาทเชื่อมต่อกันเป็นโครงข่าย ซึ่งเป็นลักษณะการเชื่อมต่อกันของเซลล์ประสาทนี้ จะเป็นตัวกำหนดโครงสร้างของโครงข่ายประสาทเทียม การเชื่อมต่อกันระหว่างเซลล์ประสาทระนั้นจะมีค่าน้ำหนักเข้ามาเกี่ยวข้องหรือไม่ก็ได้ โดยที่ค่าน้ำหนักนี้จะสามารถปรับเปลี่ยนได้ตามเหมาะสมของกฎการเรียนรู้ (Learning Rules) ซึ่งโครงสร้างของโครงข่ายประสาทเทียมอาจพิจารณาได้จากทิศทางเคลื่อนที่ของสัญญาณข้อมูล (โอพาร ธรรมานนท์, 2548) จำแนกความแตกต่างที่สำคัญได้ 2 ชนิด ดังนี้

2.2.5.1 Feedforward Networks

โครงข่ายประสาทเทียมชนิดนี้ อาจจะมีจำนวนเซลล์ประสาทตั้งแต่ 2 เซลล์ขึ้นไปประกอบอยู่ในแต่ละชั้นและสามารถมีจำนวนชั้นได้มากกว่า 1 ชั้นขึ้นไป โดยเซลล์ประสาทจะถูกจัดกลุ่มลงในชั้น (Layer) ที่เชื่อมต่อกันในลักษณะจากชั้นก่อนหน้าไปยังชั้นถัดไป การเชื่อมต่อของแต่ละชั้นจะเป็นแบบถึงกันหมด (Fully Connect) จึงทำให้สัญญาณข้อมูลเคลื่อนที่ไปในทิศทางเดียวเท่านั้น คือ เคลื่อนที่ไปข้างหน้า (Feedforward) จากอินพุทไปสู่ชั้นเอาต์พุท (Output Layer) ตัวอย่างของโครงข่ายประสาทเทียมชนิดนี้มี 2 แบบ ดังนี้

1. โครงข่ายประสาทเทียมแบบชั้นเดียว (Single Layer Neural Networks)

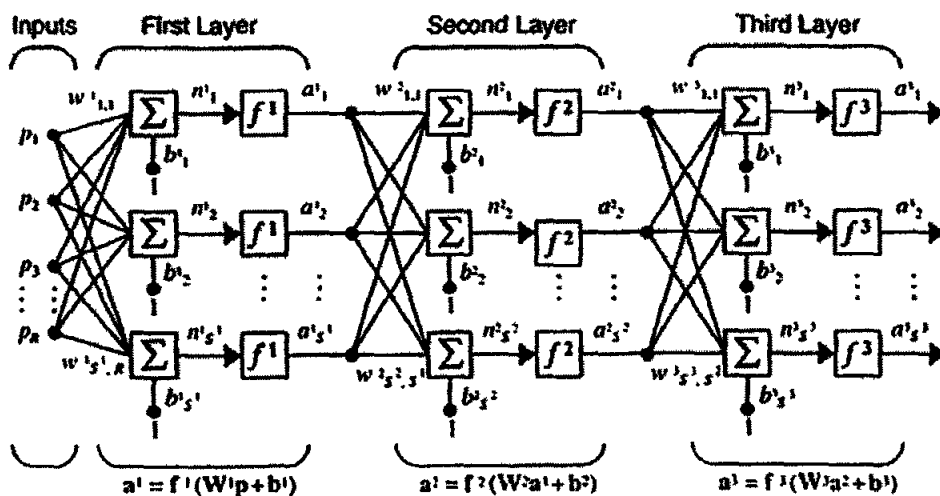
โครงข่ายประสาทเทียมแบบชั้นเดี่ยวดังแสดงในรูปที่ 2.12 จะประกอบด้วยชั้นอินพุท (Input Layer) และชั้นเอาต์พุท (Output Layer) ในชั้นอินพุทนั้นเราจะไม่นับเนื่องจากเป็นชั้นที่ไม่มีผลการประมวลผลใดๆ โครงข่ายประสาทเทียมแบบชั้นเดี่ยวนี้อาจมีความสามารถในการคำนวณที่ค่อนข้างน้อย แต่เวลาที่ใช้ในการคำนวณนั้นจะเร็ว



รูปที่ 2.12 โครงข่ายประสาทเทียมแบบชั้นเดียว

2. โครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Neural Networks)

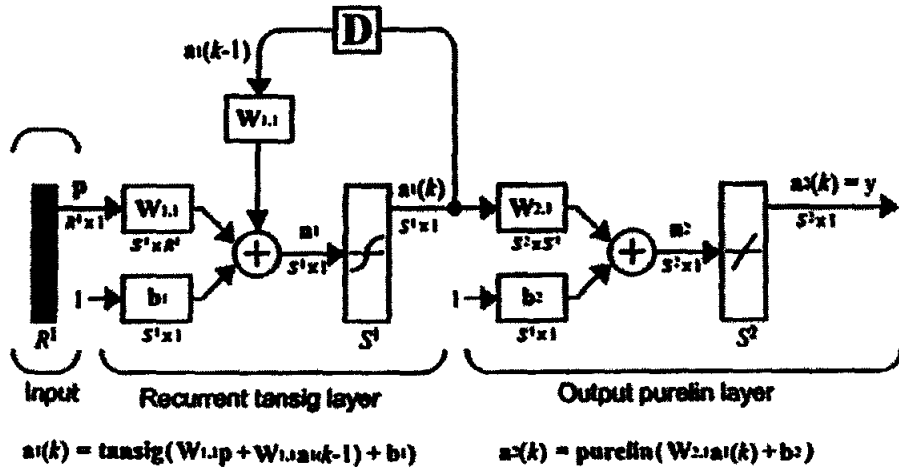
โครงข่ายประสาทเทียมแบบหลายชั้นดังแสดงในรูปที่ 2.13 จะมีชั้นที่เพิ่มขึ้นมาจากโครงข่ายประสาทเทียมแบบชั้นเดียวที่มีแค่ชั้นอินพุตและชั้นเอาต์พุต นั่นคือจะมีชั้นซ่อน (Hidden Layer) เพิ่มขึ้นมา โดยในชั้นซ่อนนี้สามารถมีได้ตั้งแต่ 1 ชั้นขึ้นไป ขึ้นอยู่กับความซับซ้อนของปัญหาที่นำมาให้โครงข่ายประสาทเทียมทำการวิเคราะห์ แต่โดยทั่วไปแล้วก็มักจะมีชั้นซ่อนไม่เกิน 2 ชั้น เนื่องจากเวลาที่ใช้ในการคำนวณ เพราะถ้าจำนวนชั้นซ่อนมาก เวลาที่ใช้ในการคำนวณก็จะมากตามไปด้วย แต่จะมีความสามารถในการคำนวณมากกว่าโครงข่ายประสาทเทียมแบบชั้นเดียว



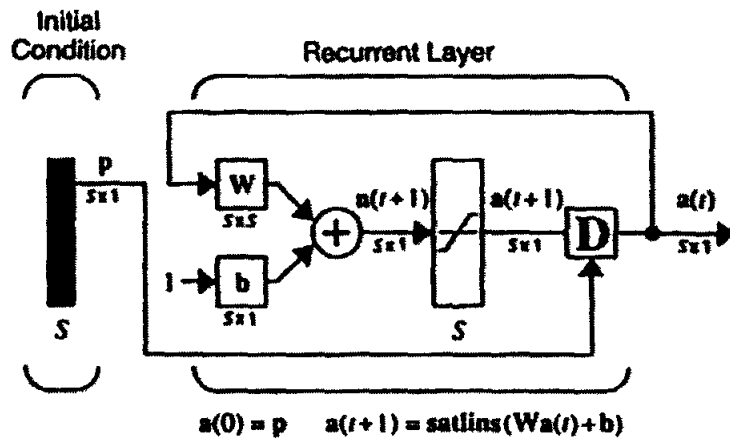
รูปที่ 2.13 โครงข่ายประสาทเทียมแบบหลายชั้น

2.2.5.2 Recurrent Networks

โครงข่ายประสาทเทียมชนิดนี้ ลักษณะเอาที่ทุกของเซลล์ประสาทจะมีการเคลื่อนที่ย้อนกลับ ไปยังชั้นก่อนหน้า หรือย้อนกลับไปเป็นอินพุตของตัวเอง จึงทำให้ลักษณะสัญญาณข้อมูลมีการเคลื่อนที่ 2 ทิศทาง คือ เคลื่อนที่ไปข้างหน้า (feedforward) และเคลื่อนที่ย้อนกลับ (Feedback) ซึ่งตัวอย่างของโครงข่ายประสาทเทียมชนิดนี้ ได้แก่ Elman Network และ Hopfield Network สามารถแสดงได้ดังรูปที่ 2.14 (a) และ 2.14 (b) ตามลำดับ



(a) Elman Network



(b) Hopfield Network

รูปที่ 2.14 Recurrent Networks

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

2.2.6 การเรียนรู้ของโครงข่ายประสาทเทียม

การเรียนรู้ของโครงข่ายประสาทเทียมจะใช้กฎการเรียนรู้ (Learning Rule) เพื่อหาค่าถ่วงน้ำหนักที่เหมาะสมให้แก่โครงข่ายนั้นๆ ซึ่งโครงข่ายประสาทเทียมสามารถแบ่งออกเป็น 2 ประเภทตามลักษณะการเรียนรู้ คือ การเรียนรู้แบบควบคุม (Supervised Learning) และการเรียนรู้แบบไม่มีการควบคุม (Unsupervised Learning)

2.2.6.1 การเรียนรู้แบบควบคุม (Supervised Learning)

การเรียนรู้วิธีนี้จะมีข้อมูลที่เรียกว่า ชุดฝึกสอน (Training Pattern) โดยในชุดฝึกสอนนี้จะมีข้อมูลอินพุตและเอาต์พุต ซึ่งโครงข่ายประสาทเทียมจะถูกฝึกสอนจากข้อมูลในชุดฝึกสอนเพื่อหาค่าถ่วงน้ำหนักที่เหมาะสม โดยนำเอาต์พุตที่ได้จากโครงข่ายประสาทเทียมไปเปรียบเทียบกับเอาต์พุตของชุดฝึกสอน จากนั้นค่าผิดพลาดที่ได้จากความแตกต่างของเอาต์พุตจากโครงข่ายประสาทเทียมและเอาต์พุตจากชุดฝึกสอนก็จะถูกป้อนกลับไปใหม่ แล้วคำนวณหาค่าถ่วงน้ำหนักใหม่ เพื่อให้ค่าผิดพลาดระหว่างเอาต์พุตของโครงข่ายประสาทเทียมกับเอาต์พุตของชุดฝึกสอนมีค่าน้อยที่สุด โครงข่ายประสาทเทียมประเภทนี้จะแบ่งออกเป็น Feedforward Neural Networks และ Radial Basis Function Networks ซึ่งทั้งสองชนิดนิยมใช้ในการแก้ปัญหาลักษณะการประมาณค่าฟังก์ชัน (Function Approximation), Forecasting, Classification และ Recognition ตัวอย่างของการเรียนรู้ชนิดนี้ ได้แก่ Delta Rule และ Back-propagation

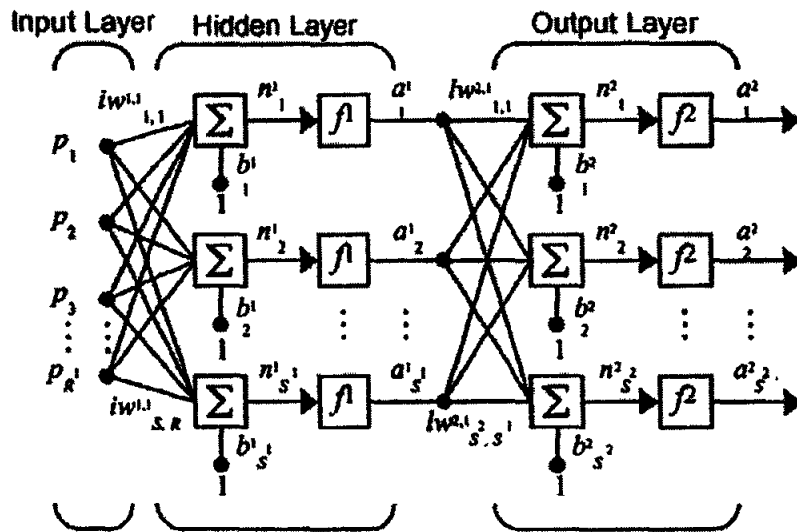
2.2.6.2 การเรียนรู้แบบไม่มีการควบคุม (Unsupervised Learning)

การเรียนรู้วิธีนี้จะป้อนข้อมูลอินพุตเข้าสู่โครงข่ายและภายในโครงข่ายจะมีโหนดเอาต์พุตจำนวนหลายโหนดด้วยกัน โดยแต่ละโหนดจะแทนกลุ่มของข้อมูลที่มีคุณสมบัติเหมือนกัน เมื่อป้อนข้อมูลอินพุตเข้าสู่โครงข่าย โครงข่ายจะคำนวณค่าความสัมพันธ์ที่มีอยู่ภายในเซตของอินพุต โดยอาศัยค่าน้ำหนักเป็นตัวแยกความแตกต่างของข้อมูลอินพุตไปเก็บไว้ในโหนดเอาต์พุตของโครงข่าย การสอนโดยวิธีนี้จะไม่สามารถระบุได้ว่าข้อมูลเอาต์พุตโหนดใดเป็นข้อมูลกลุ่มไหน ผู้ใช้ต้องกำหนดเอง ซึ่งต่างจากการสอนแบบชี้หน้าที่โครงข่ายสามารถระบุกลุ่มของข้อมูลเอาต์พุตได้อย่างแน่นอน โครงข่ายประสาทเทียมประเภทนี้แบ่งออกเป็น 2 ประเภท ได้แก่ Kohonen Networks ซึ่งนิยมใช้ในการแก้ปัญหาลักษณะการแยกแยะ (Classification Problem) และ Hopfield Networks ซึ่งนิยมใช้ในการแก้ปัญหาลักษณะการหาค่าที่เหมาะสม (Optimization Problem)

2.2.7 โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีหลายชั้น

การทำนายรูปร่างรอยเชื่อมจะใช้โครงข่ายประสาทเทียมที่มีโครงสร้างแบบ Feedforward Neural Networks ซึ่งมีลักษณะดังนี้

โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีหลายชั้น ประกอบด้วย ส่วนรับข้อมูลที่เรียกว่า ตัวแปรด้านเข้า (Inputs), ชั้นซ่อน (Hidden Layer) ที่สามารถมีได้มากกว่าหนึ่งชั้นขึ้นไปและชั้นตัวแปรด้านออก (Output Layer) โดยเซลล์ประสาทตั้งแต่ 2 เซลล์ขึ้นไปจะถูกจัดกลุ่มลงในชั้นที่เชื่อมต่อกัน ในลักษณะจากชั้นก่อนหน้าไปยังชั้นถัดไป ทุกเซลล์ประสาทในตัวแปรด้านเข้าจะส่งสัญญาณไปยังทุกเซลล์ประสาทในชั้นซ่อน และทุกเซลล์ประสาทในชั้นซ่อนจะส่งสัญญาณไปยังทุกเซลล์ประสาทในชั้นตัวแปรด้านออก ดังแสดงในรูปที่ 2.15



รูปที่ 2.15 โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีสองชั้น

จากรูปที่ 2.15 แสดงโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีสองชั้น โดยมีตัวแปรด้านเข้าเท่ากับ R มีเซลล์ประสาทในชั้นซ่อนเท่ากับ S' และในชั้นตัวแปรด้านออกเท่ากับ S''

โดยตัวแปรด้านเข้า p จะถูกเชื่อมต่อกับแต่ละเซลล์ประสาทในชั้นซ่อน ด้วยค่าน้ำหนัก W' ที่เป็นเมตริกซ์ขนาด $S' \times R$ จากนั้นก็จะส่งสัญญาณ n' ที่เป็นเมตริกซ์ขนาด $S' \times 1$ ผ่านฟังก์ชันถ่ายโอน f' ซึ่งจะได้ผลลัพธ์ของชั้นซ่อน a' เป็นเมตริกซ์ขนาด $S' \times 1$ จะไปเป็นอินพุตของชั้นข้อมูลเอาต์พุต ซึ่งจะถูกเชื่อมต่อกับแต่ละเซลล์ประสาทในชั้นข้อมูลเอาต์พุตด้วยค่าน้ำหนัก W'' ที่เป็นเมตริกซ์ขนาด $S'' \times S'$ และจะถูกรวมกับค่าไบอัส b'' ที่เป็นเมตริกซ์ขนาด $S'' \times 1$ จากนั้นก็จะส่งสัญญาณ n'' ที่เป็นเมตริกซ์ขนาด $S'' \times 1$ ผ่านฟังก์ชันถ่ายโอน f'' ซึ่งจะได้ผลลัพธ์ของโครงข่ายประสาทเทียม a'' เป็นเมตริกซ์ขนาด $S'' \times 1$ ดังแสดงในสมการ 2.7 และ 2.8 ตามลำดับ

$$a' = f'(W'p + b') \quad (2.7)$$

$$a'' = f''(W''a' + b'') \quad (2.8)$$

หลังจากนั้นก็将会ใช้การเรียนรู้แบบแพร่ค่าย้อนกลับมาทำการปรับสอนโครงข่ายประสาทเทียม ประกอบด้วย การส่งข้อมูลระหว่างชั้นของโครงข่ายประสาทเทียมสองแบบ คือ การส่งไปข้างหน้า (Forward Pass) และการส่งย้อนกลับ (Backward Pass) การส่งไปข้างหน้าข้อมูลจะถูกป้อนเข้าโครงข่ายประสาทเทียมทางตัวแปรด้านเข้าไปยังชั้นซ่อนเพื่อทำการคำนวณ และจะได้ผลลัพธ์ออกมายังตัวแปรด้านออก โดยที่ค่าน้ำหนักและค่าไบอัสที่เชื่อมต่อระหว่างชั้นของตัวแปรในโครงข่ายประสาทเทียมทั้งหมดจะยังไม่ถูกเปลี่ยนแปลง ส่วนการส่งย้อนกลับก็จะตรงกันข้ามคือ ค่าน้ำหนักและค่าไบอัส ที่เชื่อมต่อระหว่างชั้นของโครงข่ายประสาทเทียมจะถูกปรับเปลี่ยนไปตามกฎการเรียนรู้ ซึ่งค่าความแตกต่างระหว่างผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมกับเอาต์พุตเป้าหมาย หรือเรียกว่า ค่าผิดพลาด (error) จะถูกส่งกลับไปยังโครงข่ายประสาทเทียมตามทิศทางของการเชื่อมต่อ เพื่อนำไปใช้ปรับค่าน้ำหนักและค่าไบอัส โดยมี

จุดประสงค์เพื่อให้ค่าผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมเข้าใกล้กับค่าเอาต์พุตเป้าหมาย กล่าวคือ ทำให้ค่าผิดพลาดที่เกิดจากโครงข่ายประสาทเทียมมีค่าลดลง

โดยค่าผิดพลาด e ที่เกิดขึ้นในชั้นข้อมูลเอาต์พุต ซึ่งเป็นความแตกต่างระหว่างผลลัพธ์ที่ได้ของโครงข่ายประสาทเทียม a^2 กับเอาต์พุตเป้าหมาย t ซึ่งสามารถแสดงได้ดังสมการ 2.9

$$e_k = t_k - a_k^2 \quad \text{โดยที่ } k = 1, 2, \dots, S^2 \quad (2.9)$$

สำหรับการแก้ปัญหาทั่วไปที่ไม่มีความซับซ้อนมากนัก โครงข่ายประสาทเทียมที่ใช้การเรียนรู้แบบแพร่ค่าย้อนกลับมักนิยมใช้ฟังก์ชันถ่ายโอนชนิด Log-sigmoid ในชั้นซ่อน แต่ถ้าหากปัญหาที่ซับซ้อนมากขึ้นก็ต้องใช้ฟังก์ชันถ่ายโอนชนิด Tan-sigmoid ส่วนในชั้นข้อมูลเอาต์พุตมักนิยมใช้ฟังก์ชันถ่ายโอนชนิด Linear เหตุผลที่นิยมนำฟังก์ชันถ่ายโอนทั้งสองชนิดนี้มาใช้ก็เพราะสามารถหาอนุพันธ์ได้ และสร้างขอบเขตการตัดสินใจในลักษณะโค้งมนได้ โครงข่ายประสาทเทียมที่ใช้ฟังก์ชันถ่ายโอนแบบนี้จะสามารถนำไปใช้กับปัญหาที่มีความซับซ้อนและมีลักษณะไม่เชิงเส้นได้ดี ในกรณีที่ปัญหาที่มีลักษณะเชิงเส้นหรือไม่ซับซ้อนมากนัก การใช้โครงข่ายประสาทเทียมที่มีชั้นซ่อนเพียงชั้นเดียวก็เพียงพอสำหรับการแก้ปัญหานั้น แต่ถ้าหากปัญหานั้นมีความซับซ้อนและไม่เชิงเส้นมาก การแก้ปัญหาดังกล่าวก็จำเป็นต้องใช้โครงข่ายประสาทเทียมที่มีความซับซ้อนมากขึ้น เพราะว่าการเพิ่มชั้นซ่อนหนึ่งชั้นหรือมากกว่าจะทำให้โครงข่ายประสาทเทียมสามารถที่จะแยก High-order Statistics ได้ ซึ่งมีประโยชน์มากเมื่อตัวแปรด้านเข้ามีขนาดใหญ่ขึ้น

2.2.8 กฎการเรียนรู้แบบแพร่ค่าย้อนกลับ (Back-propagation Algorithm)

กฎการเรียนรู้ที่ใช้กับโครงข่ายประสาทเทียมแบบหลายชั้นมีหลายชนิด แต่จะกล่าวถึงเฉพาะกฎการเรียนรู้ที่เกี่ยวข้องกับปริภูมิพารามิเตอร์ฉบับนี้เท่านั้น

กฎการเรียนรู้แบบแพร่ค่าย้อนกลับ ในที่นี้จะพิจารณาโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มีสองชั้น ซึ่งประกอบด้วย ชั้นตัวแปรด้านออก 1 ชั้น, ชั้นซ่อน 1 ชั้น และตัวแปรด้านเข้าที่ไม่นับเป็นชั้น ดังแสดงในรูปที่ 2.12

1. การส่งไปข้างหน้า (Forward Pass)

เมื่อตัวแปรด้านเข้า p_i โดยที่ $i = 1, 2, \dots, R$ ได้ถูกส่งผ่านเข้าไปรวมกับค่าน้ำหนัก w'_{ij} และ ค่าไบอัส b'_j โดยที่ $j = 1, 2, \dots, S'$ ซึ่งจะได้ผลรวมของข้อมูลด้านเข้าของชั้นซ่อนที่ j เป็นดังนี้

$$n'_j = \sum_{i=1}^R w'_{ij} p_i + b'_j \quad (2.10)$$

หลังจากนั้นก็ส่งสัญญาณ n'_j ผ่านฟังก์ชันถ่ายโอน f' เพื่อคำนวณค่าข้อมูลด้านออกของชั้นซ่อนที่ j หรือ ยูนิต a'_j เป็นดังนี้

$$a'_j = f'(n'_j) = f' \left[\sum_{i=1}^R w'_{ij} p_i + b'_j \right] \quad (2.11)$$

ข้อมูลด้านออกของชั้นซ่อนที่ j หรือยูนิต a_j^1 ที่ได้จะถูกส่งผ่าน ไปรวมกับค่าน้ำหนัก w_{kj}^2 และ ค่าไบอัส b_k^2 โดยที่ $k = 1, 2, \dots, S^2$ ซึ่งจะได้ผลรวมของข้อมูลด้านเข้าของชั้นซ่อนที่ k เป็นดังนี้

$$n_k^2 = \sum_{j=1}^{S^1} w_{kj}^2 a_j^1 + b_k^2 \quad (2.12)$$

หลังจากนั้นก็ส่งสัญญาณ n_k^2 ผ่านฟังก์ชันถ่ายโอน f^2 เพื่อคำนวณค่าข้อมูลด้านออกของชั้นซ่อนที่ k หรือยูนิต a_k^2 เป็นดังนี้

$$a_k^2 = f^2(n_k^2) = f^2\left[\sum_{j=1}^{S^1} w_{kj}^2 a_j^1 + b_k^2\right] \quad (2.13)$$

2. การส่งย้อนกลับ (Backward Pass)

โดยจะนิยามค่าผิดพลาดที่เกิดขึ้นในชั้นข้อมูลด้านออกของตัวแปรด้านออกที่ k ด้วยฟังก์ชันค่าผิดพลาด ดังนี้

$$E = \frac{1}{2} (e_k)^2 = \frac{1}{2} \sum_{k=1}^{S^2} (t_k - a_k^2)^2 \quad (2.14)$$

ฟังก์ชันค่าผิดพลาด E เป็นสมการกำลังสองที่อยู่ในรูปของค่าน้ำหนักและค่าผิดพลาด ซึ่งมีพื้นผิวเป็นรูปพาราโบลาที่ประกอบด้วยค่าค่าสุดค่าหนึ่ง และจะใช้กฎ Gradient Descent ทำการปรับเวกเตอร์ของค่าน้ำหนักในทิศทางของเวกเตอร์ตัวแปรด้านเข้า โดย E จะมีค่าเป็นบวกเสมอ ถ้าค่าน้ำหนักและค่าไบอัสถูกปรับเข้าสู่ค่าที่ต้องการ ค่าของฟังก์ชันก็จะค่อยๆ ลดลงในลักษณะที่ลู่อเข้าสู่ศูนย์ เมื่อผ่านกระบวนการทำซ้ำไปเรื่อยๆ ซึ่งกระบวนการปรับค่าน้ำหนักและค่าไบอัสนี้จะทำการปรับค่าฟังก์ชันลงมายังจุดต่ำสุดบนพื้นผิวที่นิยามใน w space และ b space สมการ 2.14 สามารถเขียนใหม่ได้เป็น

$$E = \frac{1}{2} \sum_{k=1}^{S^2} \left[t_k - f^2\left(\sum_{j=1}^{S^1} w_{kj}^2 a_j^1 + b_k^2\right) \right]^2 \quad (2.15)$$

จากนั้นใช้กฎลูกโซ่ (Chain Rule) เพื่อคำนวณหาค่า Gradient ซึ่งที่ตัวแปรด้านออกจะได้

$$\frac{\partial E}{\partial w_{kj}^2} = -(t_k - a_k^2) f'^2(n_k^2) a_j^1 \quad (2.16)$$

$$\frac{\partial E}{\partial b_k^2} = -(t_k - a_k^2) f'^2(n_k^2) \quad (2.17)$$

และชั้นซ่อนจะได้

$$\frac{\partial E}{\partial w_{ji}^1} = -\sum_{k=1}^{S^2} [(t_k - a_k^2) f^{2'}(n_k^2) w_{kj}^2] f^{1'}(n_j^1) p_i \quad (2.18)$$

$$\frac{\partial E}{\partial b_j^1} = -\sum_{k=1}^{S^2} [(t_k - a_k^2) f^{2'}(n_k^2) w_{kj}^2] f^{1'}(n_j^1) \quad (2.19)$$

กำหนดให้

$$\delta_k = (t_k - a_k^2) f^{2'}(n_k^2) \quad (2.20)$$

$$\delta_j = f^{1'}(n_j^1) \sum_{k=1}^{S^2} (\delta_k w_{kj}^2) \quad (2.21)$$

ถ้า E เป็นฟังก์ชันที่สามารถหาอนุพันธ์ได้อย่างต่อเนื่องในทุกค่าของค่าน้ำหนักและค่าไบอัส จะสามารถคำนวณหาค่าน้ำหนักและค่าไบอัสที่เหมาะสมโดยใช้กฎ Gradient Descent ได้ดังนี้
ที่เซลล์ประสาทของตัวแปรด้านออกจะได้

$$\Delta w_{kj}^2 = -\alpha \frac{\partial E}{\partial w_{kj}^2} = \alpha \delta_k a_j^1 \quad (2.22)$$

$$\Delta b_k^2 = -\alpha \frac{\partial E}{\partial b_k^2} = \alpha \delta_k \quad (2.23)$$

และที่เซลล์ประสาทของชั้นซ่อนจะได้

$$\Delta w_{ji}^1 = -\alpha \frac{\partial E}{\partial w_{ji}^1} = \alpha \delta_j p_i \quad (2.24)$$

$$\Delta b_j^1 = -\alpha \frac{\partial E}{\partial b_j^1} = \alpha \delta_j \quad (2.25)$$

ดังนั้น การปรับค่าน้ำหนักและค่าไบอัสของตัวแปรด้านออก (Weight Update) จะได้

$$w_{kj}^{2(new)} = w_{kj}^{2(old)} + \Delta w_{kj}^2 = w_{kj}^{2(old)} + \alpha \delta_k a_j^1 \quad (2.26)$$

$$b_k^{2(new)} = b_k^{2(old)} + \Delta b_k^2 = b_k^{2(old)} + \alpha \delta_k \quad (2.27)$$

และการปรับค่าน้ำหนักและค่าไบอัสของชั้นซ่อนจะได้

$$w_{ji}^{l(new)} = w_{ji}^{l(old)} + \Delta w_{ji}^l = w_{ji}^{l(old)} + \alpha \delta_j p_i \quad (2.28)$$

$$b_{ji}^{l(new)} = w_{ji}^{l(old)} + \Delta b_j^l = b_j^{l(old)} + \alpha \delta_j \quad (2.29)$$

โดยที่ α คือ อัตราการเรียนรู้ (Learning Rule) ปกติจะกำหนดค่าเริ่มต้นที่ค่าน้อย เช่น 0.01

กระบวนการนี้จะนำค่าผิดพลาด ($e_k - a_k^2$) ที่ได้ในชั้นตัวแปรด้านนอกกลับมาคำนวณหาค่าที่ใช้ในการปรับค่าน้ำหนักและค่าไบอัสของแต่ละเซลล์ประสาท โดยจะเริ่มต้นจากเซลล์ประสาทในชั้นตัวแปรด้านนอก จนถึงเซลล์ประสาทในชั้นซ่อนแรก ซึ่งจะเรียกกระบวนการนี้ว่า กระบวนการแพร่ค่าผิดพลาดย้อนกลับ ซึ่งค่าน้ำหนักและค่าไบอัสของแต่ละเซลล์ประสาทจะถูกปรับด้วยค่า Δw และค่า Δb จนกระทั่งฟังก์ชันค่าผิดพลาดมีค่าลดลงถึงจุดที่ยอมรับได้ โดยค่า α ที่ใช้มีค่าน้อยก็จะทำให้การคำนวณใช้เวลาาน ในทางตรงข้ามหากค่า α มีค่ามากก็จะทำให้การคำนวณใช้เวลาเร็วขึ้น แต่อาจเกิดการแกว่งขึ้นได้ เมื่อต้องปรับค่าฟังก์ชันมายังจุดต่ำสุดบน Flat Surface (โอพาร ธรรมานนท์, 2548)

ในขั้นตอนการปรับค่าน้ำหนักที่กล่าวมาข้างต้น เป็นหลักการกฎการเรียนรู้แบบดั้งเดิม ซึ่งต่อมาได้รับการปรับปรุงให้สามารถลู่เข้าคำตอบได้เร็วยิ่งขึ้น เกิดเป็นหลายวิธี เช่น Gradient Descent with Adaptive Learning Rate, Gradient Descent with Momentum, Gradient Descent with Momentum and Adaptive Learning Rate, Resilient, Bayesian Regularization และ Levenberg-Marquardt Algorithm ซึ่งปัญญาประดิษฐ์ฉบับนี้จะเลือกใช้ Levenberg-Marquardt Algorithm เนื่องจากพบว่าเป็นกฎการเรียนรู้ที่สามารถปรับสอนโครงข่ายประสาทเทียมได้รวดเร็วกว่าวิธีการอื่นๆ และเหมาะกับ โครงข่ายประสาทเทียมที่มีจำนวนตัวแปรขนาดกลาง

Levenberg-Marquardt Algorithm เป็นกฎการเรียนรู้แบบแพร่ค่าย้อนกลับที่ปรับปรุงมาจากวิธี Gauss-Newton Nonlinear Least Squares Optimization Algorithm โดยโครงข่ายประสาทเทียมจะปรับค่าน้ำหนักและค่าไบอัสตามสมการ 2.30 และ 2.31 ซึ่งเขียนในรูปของเมตริกซ์ ดังนี้

$$w^{new} = w^{old} + [J^T(w) \cdot J(w) + \mu I]^{-1} J^T(w) \cdot e \quad (2.30)$$

$$b^{new} = b^{old} + [J^T(b) \cdot J(b) + \mu I]^{-1} J^T(b) \cdot e \quad (2.31)$$

โดยที่ μ คือ Learning Factor ซึ่งมีค่าเริ่มต้นเท่ากับ 0.001

I คือ Identity Matrix

$J(w)$ คือ จาโคเบียนเมตริกซ์ (Jacobian matrix) ของอนุพันธ์ค่าผิดพลาดเทียบกับค่าน้ำหนัก

$J^T(w)$ คือ Transpose matrix ของ $J(w)$

$J(b)$ คือ จาโคเบียนเมตริกซ์ (Jacobian matrix) ของอนุพันธ์ค่าผิดพลาดเทียบกับค่าไบอัส

$J^T(b)$ คือ Transpose matrix ของ $J(b)$

e คือ เมตริกซ์ของค่าผิดพลาดระหว่างเอาท์พุทเป้าหมายกับตัวแปรด้านออก

$$J(w) = \begin{bmatrix} \frac{\partial e_1}{\partial w_{1,1}^2} & \frac{\partial e_1}{\partial w_{1,2}^2} & \dots & \frac{\partial e_1}{\partial w_{1,s^1}^2} \\ \frac{\partial e_2}{\partial w_{2,1}^2} & \frac{\partial e_2}{\partial w_{2,2}^2} & \dots & \frac{\partial e_2}{\partial w_{2,s^1}^2} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \frac{\partial e_{s^2}}{\partial w_{s^2,1}^2} & \frac{\partial e_{s^2}}{\partial w_{s^2,2}^2} & \dots & \frac{\partial e_{s^2}}{\partial w_{s^2,s^1}^2} \end{bmatrix} \quad (2.32)$$

$$J(b) = \begin{bmatrix} \frac{\partial e_1}{\partial b_1^2} \\ \frac{\partial e_2}{\partial b_2^2} \\ \cdot \\ \cdot \\ \frac{\partial e_{s^2}}{\partial b_{s^2}^2} \end{bmatrix} \quad (2.33)$$

$$\begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_{s^2} \end{bmatrix} = \begin{bmatrix} t_1 - a_1^2 \\ t_2 - a_2^2 \\ \cdot \\ \cdot \\ t_{s^2} - a_{s^2}^2 \end{bmatrix} \quad (2.34)$$

ค่า Learning Factor (μ) จะมีค่าเริ่มต้นของการฝึกสอนด้วยค่าน้อยๆ เช่น เท่ากับ 0.001 และจะปรับเปลี่ยนในช่วงเวลาระหว่างกระบวนการเรียนรู้ ขณะเริ่มกระบวนการเรียนรู้ อินพุททั้งหมดจะเข้าสู่โครงข่ายประสาทเทียม ซึ่งในทุกรอบของการเรียนรู้ โครงข่ายประสาทเทียมจะคำนวณหาค่าผิดพลาดที่เกิดขึ้น และนำมาปรับเปลี่ยนค่าน้ำหนักและค่าไบอัส ถ้าค่า E ใหม่มีค่าน้อยกว่าค่า E เก่า ก็จะปรับลดค่า μ โดยคูณด้วย Decreased Factor ที่มีค่าเท่ากับ 0.1 และทำการฝึกสอนต่อไป แต่ถ้าค่า E ใหม่มากกว่าค่า E เก่า ดังนั้นค่า Δw และ Δb ในสมการที่ 2.30 และ 2.31 ก็จะถูกปรับเพิ่มค่า μ โดยการคูณด้วย Increased Factor ที่มีค่าเท่ากับ 10 จนกระทั่ง ค่า E มีค่าลดลง การฝึกสอนโครงข่ายประสาทเทียมจะสิ้นสุดลง ถ้าหากค่า Gradient Factor (ค่า $J^T(w) \cdot e$ หรือ $J^T(b) \cdot e$) มีค่าน้อยกว่าค่าสูงสุดที่กำหนดไว้ เช่น $1 \cdot e^{-10}$ หรือเมื่อค่า μ มีค่ามากกว่าค่าสูงสุดที่กำหนดไว้ เช่น $1 \cdot e^{10}$

โครงข่ายประสาทเทียมแบบแพร่ค่าย้อนกลับนั้น จะมีจำนวนนิวรอนในชั้นอินพุตและชั้นเอาต์พุตคงที่ ซึ่งเท่ากับจำนวนอินพุตและเอาต์พุตของปัญหานั้นเอง ส่วนจำนวนนิวรอนในชั้นซ่อนนั้นสามารถปรับเปลี่ยนได้ตามความเหมาะสมและไม่มีจำนวนที่เป็นสูตรสำเร็จ ดังนั้น ในการฝึกสอนของโครงข่ายประสาทเทียมจึงต้องมีการปรับเปลี่ยนจำนวนนิวรอนในชั้นซ่อน เพื่อหาโครงสร้างที่เหมาะสมที่ให้ค่าผิดพลาดระหว่างคำตอบที่แท้จริงและผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมนั้นมีค่าน้อยที่สุด จำนวนนิวรอนในชั้นซ่อนนอกจากจะมีผลต่อประสิทธิภาพของโครงข่ายประสาทเทียมแล้ว ยังส่งผลกระทบต่อระยะเวลาในการฝึกสอนอีกด้วย นั่นคือ ถ้ามีจำนวนนิวรอนในชั้นซ่อนมาก เวลาที่ใช้ในการฝึกสอนก็จะมากกว่าโครงข่ายประสาทเทียมที่มีจำนวนนิวรอนในชั้นซ่อนน้อยกว่า เมื่อฝึกสอนด้วยเงื่อนไขเดียวกัน อย่างไรก็ตาม ไม่สามารถสรุปได้ว่าโครงข่ายประสาทเทียมที่มีจำนวนนิวรอนในชั้นซ่อนมากจะมีประสิทธิภาพสูงกว่าโครงข่ายประสาทเทียมที่มีจำนวนนิวรอนในชั้นซ่อนน้อย ซึ่งต้องทำการลองปรับเปลี่ยนเพื่อหาโครงข่ายที่เหมาะสมที่สุด (ฉัตรชัย สานติสุขรัตน์, 2548)

2.2.9 การ Normalization และ Unnormalization ชุดข้อมูล

การ Normalization หรือการจัดระเบียบข้อมูล เป็นวิธีการที่นำมาใช้เพิ่มความแม่นยำของผลลัพธ์ (Output) ที่ได้จากโครงข่ายประสาทเทียม และลดเวลาในการเรียนรู้ของโครงข่ายประสาทเทียม การทำ Normalization ชุดข้อมูลของปัจจัยนำเข้า (Input) และผลลัพธ์ (Output) ต้องทำก่อนเริ่มกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม เพื่อให้ค่าข้อมูลที่ได้อยู่บนพื้นฐานการแจกแจงแบบปกติ (Hong Tae Kang และ Jung Me Park, 2006)

โดยการเลือกขอบเขตของการ Normalization ขึ้นอยู่กับฟังก์ชันถ่ายโอน สำหรับการทำนายรูปร่างรอยเชื่อมนี้ จะทำการ Normalization ข้อมูลให้อยู่ในช่วง $[0,1]$ เนื่องจากเลือกใช้ฟังก์ชันถ่ายโอนชนิด Log-sigmoid ในชั้นซ่อนและใช้ฟังก์ชันถ่ายโอนชนิด Linear ในชั้นเอาต์พุต โดยสมการที่ใช้ในการปรับค่าในช่วง $a-b$ คือ

$$f_i(x) = \left[(x - Min_i) \times \left(\frac{b - a}{Max_i - Min_i} \right) \right] + a \quad (2.35)$$

$$x = \left[(f_i(x) - a) \times \left(\frac{Max_i - Min_i}{b - a} \right) \right] + Min_i \quad (2.36)$$

โดยที่ Max_i คือ ค่าที่มากที่สุดของแต่ละตัวแปร

Min_i คือ ค่าที่น้อยที่สุดของแต่ละตัวแปร

x คือ ค่าของตัวแปรที่จะทำการ Normalization

$f_i(x)$ คือ ค่าของตัวแปรที่ผ่านการ Normalization โดยมีค่าอยู่ในช่วง $[0,1]$

ซึ่งการ Normalization จะทำทั้งชุดข้อมูลของปัจจัยด้านเข้า (Input) และผลลัพธ์ (Output) ตามสมการ 2.35 ส่วนการทดสอบจะนำข้อมูลเอาต์พุตจากการเรียนรู้โดยใช้โครงข่ายประสาทเทียม นำไปทำการ Unnormalization ตามสมการ 2.36 เพื่อแปลงค่ากลับเป็นค่าในขอบเขตเริ่มต้น แล้วนำไปเปรียบเทียบกับข้อมูลต้นแบบต่อไป

2.3 การวิเคราะห์ความถดถอยเชิงพหุ (Multiple Regression Analysis)

ถ้ามีตัวแปรอิสระ k ตัว (X_1, X_2, \dots, X_k) ที่มีความสัมพันธ์กับตัวแปรตาม Y โดยที่ความสัมพันธ์อยู่รูปเชิงเส้น จะได้ความถดถอยเชิงพหุ ซึ่งแสดงความสัมพันธ์ระหว่าง Y และ X_1, X_2, \dots, X_k ดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e \quad (2.37)$$

โดยที่ β_0 คือ ส่วนตัดแกน Y เมื่อกำหนดให้ $X_1 = X_2 = \dots = X_k = 0$
 $\beta_1, \beta_2, \dots, \beta_k$ คือ สัมประสิทธิ์ความถดถอยเชิงส่วน (Partial Regression Coefficient) โดยที่ β_j เป็นค่าที่แสดงถึงการเปลี่ยนแปลงของตัวแปรตาม Y เมื่อตัวแปรอิสระ X_j เปลี่ยนไป 1 หน่วย โดยที่ตัวแปรอิสระ X ตัวอื่นๆ มีค่าคงที่
 e คือ ค่าความคลาดเคลื่อน

2.3.1 การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยสุด

จากสมการความถดถอยเชิงพหุ ซึ่งมีพารามิเตอร์ $k + 1$ ตัว คือ $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ การประมาณค่า $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ จะต้องใช้ข้อมูลตัวอย่างของ Y, X_1, X_2, \dots, X_k โดยใช้ตัวอย่างขนาด n จากสมการความถดถอยเชิงพหุในสมการที่ 2.37 และจะประมาณค่า Y หรือจะประมาณสมการที่ 2.37 ด้วยสมการที่ 2.38

$$\hat{Y}_i = a + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki} \quad (2.38)$$

โดยที่ $\beta_0 = a, \beta_1 = b_1, \beta_2 = b_2, \dots, \beta_k = b_k$

ดังนั้นค่าความคลาดเคลื่อนในการประมาณค่า Y_i ด้วย \hat{Y}_i คือ $Y_i - \hat{Y}_i = e_i$ หรือเรียกว่า Residual หรือ Error การประมาณค่า $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ ด้วยค่า a, b_1, b_2, \dots, b_k นั้นมีจุดประสงค์เพื่อทำให้ผลบวกของค่าความคลาดเคลื่อนยกกำลังสองมีค่าน้อยที่สุด โดยใช้วิธีกำลังสองน้อยที่สุด กรณีที่มีตัวแปรอิสระ 2 ตัว (X_1, X_2) ที่มีความสัมพันธ์กับ Y ดังนั้นสมการความถดถอยจะเป็นดังนี้

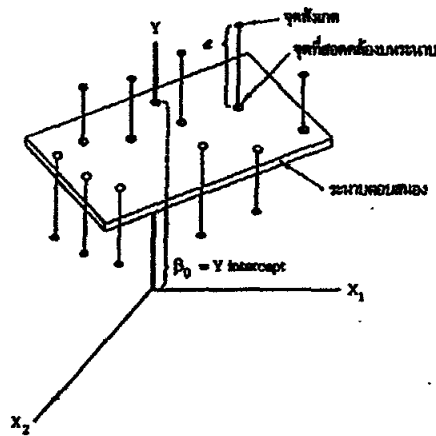
$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i \quad (2.39)$$

และค่าประมาณของ Y_i คือ

$$\hat{Y}_i = a + b_1 X_{1i} + b_2 X_{2i} \quad (2.40)$$

และค่าความคลาดเคลื่อน $e_i = Y_i - \hat{Y}_i$

ซึ่งกราฟของการถดถอยเชิงพหุ จะเป็นผิวตอบสนอง (Response surface) แทนเส้นตรง ในกรณีที่มีสมการมีการถดถอยมี 3 ตัวแปร คือ Y, X_1 , และ X_2 กราฟจะอยู่ใน 3 มิติ ดังรูปที่ 2.16 ซึ่งแสดงแผนภาพการกระจายของผิวตอบสนอง เมื่อตัวแปรอิสระมี 2 ตัว สมการถดถอยเชิงพหุก็จะสร้างระนาบตอบสนอง (Response plane) จากรูปจะเห็นอนุกรมของจุดรอบๆ ระนาบ บางจุดอยู่เหนือระนาบ บางจุดอยู่ใต้ระนาบ โดยระยะแนวตั้งจาก Y_i คือระยะความคลาดเคลื่อน e_i



รูปที่ 2.16 ระนาบถดถอยเชิงพหุที่มีตัวแปรตาม Y และตัวแปรอิสระ X_1 และ X_2

เนื่องจากต้องการ $\min \sum_{i=1}^n e_i^2 = \min \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ ดังนั้นจึงใช้อนุพันธ์ย่อยเทียบกับ a, b_1 และ b_2 แล้วทำให้เท่ากับศูนย์ดังนี้

$$\frac{\partial}{\partial a} \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] = \frac{\partial}{\partial a} \left[\sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i})^2 \right] = 0$$

$$-2 \sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i}) = 0$$

$$-2 \sum_{i=1}^n Y_i + 2na + 2b_1 \sum_{i=1}^n X_{1i} + 2b_2 \sum_{i=1}^n X_{2i} = 0$$

$$na + b_1 \sum_{i=1}^n X_{1i} + b_2 \sum_{i=1}^n X_{2i} = \sum_{i=1}^n Y_i$$

(2.41)

$$\frac{\partial}{\partial b_1} \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] = \frac{\partial}{\partial b_1} \left[\sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i})^2 \right] = 0$$

$$\begin{aligned}
& -2X_{1i} \sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i}) = 0 \\
& -2 \sum_{i=1}^n X_{1i} Y_i + 2a \sum_{i=1}^n X_{1i} + 2b_1 \sum_{i=1}^n X_{1i}^2 + 2b_2 \sum_{i=1}^n X_{1i} X_{2i} = 0 \\
& a \sum_{i=1}^n X_{1i} + b_1 \sum_{i=1}^n X_{1i}^2 + b_2 \sum_{i=1}^n X_{1i} X_{2i} = \sum_{i=1}^n X_{1i} Y_i
\end{aligned} \tag{2.42}$$

$$\begin{aligned}
& \frac{\partial}{\partial b_2} \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] = \frac{\partial}{\partial b_2} \left[\sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i})^2 \right] = 0 \\
& -2X_{2i} \sum_{i=1}^n (Y_i - a - b_1 X_{1i} - b_2 X_{2i}) = 0
\end{aligned}$$

$$\begin{aligned}
& -2 \sum_{i=1}^n X_{2i} Y_i + 2a \sum_{i=1}^n X_{2i} + 2b_1 \sum_{i=1}^n X_{1i} X_{2i} + 2b_2 \sum_{i=1}^n X_{2i}^2 = 0 \\
& a \sum_{i=1}^n X_{2i} + b_1 \sum_{i=1}^n X_{1i} X_{2i} + b_2 \sum_{i=1}^n X_{2i}^2 = \sum_{i=1}^n X_{2i} Y_i
\end{aligned} \tag{2.43}$$

สมการที่ 2.41, 2.42, 2.43 เรียกว่า ชุดของสมการปกติ ดังนี้

$$\left. \begin{aligned}
& na + b_1 \sum_{i=1}^n X_{1i} + b_2 \sum_{i=1}^n X_{2i} = \sum_{i=1}^n Y_i \\
& a \sum_{i=1}^n X_{1i} + b_1 \sum_{i=1}^n X_{1i}^2 + b_2 \sum_{i=1}^n X_{1i} X_{2i} = \sum_{i=1}^n X_{1i} Y_i \\
& a \sum_{i=1}^n X_{2i} + b_1 \sum_{i=1}^n X_{1i} X_{2i} + b_2 \sum_{i=1}^n X_{2i}^2 = \sum_{i=1}^n X_{2i} Y_i
\end{aligned} \right\} \tag{2.44}$$

ซึ่งสามารถคำนวณหาค่า a , b_1 และ b_2 โดยใช้เมทริกซ์ได้ ดังนี้

$$\begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1 X_2 \\ \sum X_2 & \sum X_1 X_2 & \sum X_2^2 \end{bmatrix} \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \end{bmatrix} \tag{2.45}$$

หรืออาจใช้สัญลักษณ์ $X'Xb = XY$ (2.46)

$$\text{โดยที่ } X = \begin{bmatrix} 1 & X_{11} & X_{21} \\ 1 & X_{12} & X_{22} \\ \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} \end{bmatrix}, b = \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix}, Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, X'Y = \begin{bmatrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \end{bmatrix}$$

$$\text{และ } b = \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = (X'X)^{-1}X'Y$$

$$\therefore b = \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1 X_2 \\ \sum X_2 & \sum X_1 X_2 & \sum X_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \end{bmatrix}$$

$$\text{ถ้าให้ } X'X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

$$\text{จะได้ } (X'X)^{-1} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}^{-1} = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix}$$

$$\text{โดยที่ } \begin{array}{lll} A = \frac{(ek-fh)}{z} & B = \frac{(bk-ch)}{z} & C = \frac{(bf-ce)}{z} \\ D = \frac{(dk-fg)}{z} & E = \frac{(ak-cg)}{z} & F = \frac{(af-cd)}{z} \\ G = \frac{(dh-eg)}{z} & H = \frac{(ah-bg)}{z} & K = \frac{(ae-bd)}{z} \end{array}$$

และ z เป็นดีเทอร์มิแนนต์ของเมทริกซ์ $X'X$ โดย $z = a(ek - fh) - b(dk - fg) + c(dh - eg)$

$$\text{ดังนั้น } \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix} \begin{bmatrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \end{bmatrix} \quad (2.47)$$

2.3.2 การวิเคราะห์ความแปรปรวน (Analysis of Variance)

เมื่อมีตัวแปรอิสระหลายตัวในสมการถดถอย ซึ่งสามารถทดสอบความมีนัยสำคัญทั้งหมดของความสัมพันธ์ของการถดถอยเชิงเส้นแบบพหุ หรือความสัมพันธ์ระหว่างตัวแปรตาม Y และเซตของตัวแปรอิสระ X_1, X_2, \dots, X_k โดยตั้งสมมติฐานดังนี้

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{มี } \beta_i \neq 0 \text{ อย่างน้อย 1 ค่า, } i = 1, 2, \dots, k$$

สถิติสำหรับการทดสอบ คือ

$$F = \frac{MSR}{MSE} = \frac{(b'X'Y - n\bar{y}^2)/k}{(Y'Y - b'XY)/(n - k - 1)} \quad (2.48)$$

ซึ่งจะปฏิเสธสมมติฐาน H_0 ถ้า $F > F_{k, n-k-1, \alpha}$

ผลการทดสอบสมมติฐานอาจเป็นดังนี้

1. ยอมรับสมมติฐาน $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ ซึ่งสรุปได้ว่า Y ไม่มีความสัมพันธ์กับ X ทั้ง k ตัว (X_1, X_2, \dots, X_k) ในรูปเชิงเส้น
2. ปฏิเสธสมมติฐาน H_0 หรือยอมรับสมมติฐาน H_1 ซึ่งสรุปได้ว่ามี X_i อย่างน้อย 1 ตัวที่มีความสัมพันธ์กับ Y ในรูปเชิงเส้น ซึ่งต้องทดสอบต่อไปว่า X_i ตัวใดมีความสัมพันธ์กับ Y โดยใช้สถิติทดสอบ t (t-Test)

การทดสอบ F (F-Test) นั้นจำเป็นต้องสร้างตารางวิเคราะห์ความแปรปรวน (ANOVA Table) ดังตารางที่ 2.1 ส่วนค่า SST, SSR และ SSE จะคำนวณได้จากสมการที่ 2.49 ถึง 2.51 ดังนี้

$$\text{จากสมการถดถอยเชิงเส้นแบบพหุ } Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e$$

ค่าแปรปรวนของ Y = ค่าแปรปรวนที่เกิดจากอิทธิพลของ X_1, \dots, X_k + ค่าแปรปรวนอย่างสุ่ม หรือ $SST = SSR + SSE$ หรือ

$$SST = \sum_{i=1}^n (Y_i - \bar{y})^2$$

$$SST = \sum_{i=1}^n (Y_i - \bar{y})^2 = Y'Y - n\bar{y}^2 \quad (2.49)$$

$$SSR = b'X'Y - n\bar{y}^2 \quad (2.50)$$

$$SSE = \sum_{i=1}^n [Y_i - (a + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki})]^2 \quad (2.51)$$

$$\text{หรือ } SSE = SST - SSR = Y'Y - b'X'Y$$

โดยที่	SST (Sum Square of Total)	คือ ค่าแปรปรวนทั้งหมดของ Y
	SSR (Sum Square of Regression)	คือ ค่าแปรปรวนของ Y เนื่องจากอิทธิพลของ X_1, X_2, \dots, X_k
	SSE (Sum Square of Error)	คือ ค่าแปรปรวนของ Y เนื่องจากอิทธิพลอื่นๆ

ตารางที่ 2.1 การวิเคราะห์ความแปรปรวนของการวิเคราะห์เชิงพหุ

แหล่งแปรปรวน (SV)	องศาอิสระ (DF)	ผลบวกกำลังสอง (SS)	ผลบวกกำลังสองเฉลี่ย (MS)	F
ความถดถอย (Regression)	K	SSR	$MSR = \frac{SSR}{K}$	$\frac{MSR}{MSE}$
ความคลาดเคลื่อน (Error) หรือ Residual	$n - k - 1$	SSE	$MSE = \frac{SSE}{n - k - 1}$	
ผลรวม	$n - 1$	SST		

2.3.3 การประมาณค่าความคลาดเคลื่อน

การประมาณค่าความคลาดเคลื่อนของการถดถอย หรือการประมาณค่าความแปรปรวนของการประมาณค่าความคลาดที่เกิดจากการประมาณค่า $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i$ ด้วย $\hat{Y}_i = a + b_1 X_{1i} + b_2 X_{2i}$ เป็นค่าความคลาดเคลื่อน $e_i = Y_i - \hat{Y}_i$

กรณีที่มีตัวแปรอิสระ k ตัว จะได้ค่าความแปรปรวนของการประมาณคือ $S_e^2 = S_{\hat{Y}_{1,2,\dots,k}}^2 = S^2$ โดยที่

$$S^2 = \frac{SSE}{n - k - 1} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - k - 1} \quad (2.52)$$

ดังนั้นค่าความคลาดเคลื่อน หรือค่าเบี่ยงเบนมาตรฐานของการประมาณค่า คือ

$$S = \sqrt{S^2} = \sqrt{\frac{SSE}{(n - k - 1)}} = \sqrt{MSE} \quad (2.53)$$

กรณีที่มีตัวแปรอิสระ 2 ตัว คือ (X_1, X_2) จะได้

$$S^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 2 - 1} = MSE \quad (2.54)$$

2.3.4 การประมาณค่าเบี่ยงเบนมาตรฐานของสัมประสิทธิ์การถดถอย

ในการประมาณค่า β_i ด้วยค่า b_i แบบข้างนั้น จะต้องทราบค่าเบี่ยงเบนมาตรฐานของ $b_i, i = 1, 2, \dots, k$ โดยที่ค่าความแปรปรวนของ b_i คือ $(b) = S^2(b) = (X'X)^{-1}S^2$

กรณีที่ $k = 2$ หรือมีตัวแปรอิสระ 2 ตัว คือ (X_1, X_2) จะได้

$$b = \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix}, S_b^2 = \begin{bmatrix} S_a^2 & S_{a,b_1} & S_{a,b_2} \\ S_{a,b_1} & S_{b_1}^2 & S_{b_1,b_2} \\ S_{a,b_2} & S_{b_1,b_2} & S_{b_2}^2 \end{bmatrix}$$

และค่าเบี่ยงเบนมาตรฐานของ b_1 คือ

$$S_{b_1} = \sqrt{S_{b_1}^2} = \frac{S}{\sqrt{\sum X_1^2 (1 - r_{12}^2)}} \quad (2.55)$$

$$S_{b_2} = \sqrt{S_{b_2}^2} = \frac{S}{\sqrt{\sum X_2^2 (1 - r_{12}^2)}} \quad (2.56)$$

$$r_{12} = \frac{\sum X_1 X_2}{\sqrt{(\sum X_1^2)(\sum X_2^2)}} \quad (2.57)$$

โดยที่ $x_i = (X_i - \bar{X}_i)$

r_{12} = สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X_1 กับ X_2

2.3.5 ความหมายของสัมประสิทธิ์ความถดถอยเชิงส่วน

ถ้ามีตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (Y) 3 ตัว คือ X_1, X_2 และ X_3 สมการความถดถอยเชิงพหุ คือ

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$$

$$\text{ค่าประมาณของ } Y \text{ คือ } \hat{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3 \quad (2.58)$$

จากสมการที่ 2.58 ค่าประมาณของ $\beta_0, \beta_1, \beta_2$ และ β_3 คือ a, b_1, b_2 และ b_3 ตามลำดับโดยที่

a คือ ส่วนหรือระยะตัดแกน Y ซึ่งเมื่อกำหนดให้ $X_1 = X_2 = X_3 = 0$

b_1, b_2 และ b_3 เป็นค่าประมาณของสัมประสิทธิ์ความถดถอยเชิงส่วน ซึ่งมีหน่วยเหมือน Y และมีความหมายดังนี้

- b_1 เป็นค่าซึ่งแสดงความสัมพันธ์ระหว่าง Y และ X_1 หมายถึง ถ้า X_1 เพิ่มขึ้น 1 หน่วย จะทำให้ Y เปลี่ยนไป b_1 หน่วย (ขึ้นอยู่กับเครื่องหมายของ b_1) โดยที่ที่กำหนดให้ตัวแปรอิสระอื่นๆ คือ X_2 และ X_3 มีค่าคงที่
- b_2 เป็นค่าที่แสดงความสัมพันธ์ระหว่าง Y และ X_2 หมายถึง ถ้า X_2 เพิ่มขึ้น 1 หน่วย จะทำให้ Y เปลี่ยนไป b_2 หน่วย โดยที่ที่กำหนดให้ตัวแปรอิสระอื่นๆ คือ X_1 และ X_3 มีค่าคงที่

ในทำนองเดียวกัน b_3 จะแสดงถึงความสัมพันธ์ระหว่าง Y กับ X_3 โดยที่ถ้า X_3 เพิ่มขึ้น 1 หน่วย จะทำให้ Y เปลี่ยนไป b_3 หน่วย โดยควบคุมให้ X_1 และ X_2 มีค่าคงที่

การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอย

ในการทดสอบเกี่ยวกับค่า β_i เมื่อมีตัวแปรอิสระมากกว่า 1 ตัว จะมี 2 ขั้นตอนดังนี้

ขั้นที่ 1 $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$

H_1 : มี β_i อย่างน้อย 1 ค่าที่ $\neq 0$; $i = 1, 2, \dots, k$

การทดสอบจะใช้ F-test จากตาราง ANOVA

- ถ้ายอมรับ H_0 จะสรุปได้ว่า X_1, X_2, \dots, X_k ไม่มีความสัมพันธ์กับ Y
- ถ้าปฏิเสธ H_0 จะสรุปได้ว่ามี X_i อย่างน้อย 1 ตัวที่มีความสัมพันธ์กับ Y จึงต้องทดสอบขั้นต่อไปในขั้นที่ 2 ว่ามี X ตัวใดบ้างที่มีความสัมพันธ์กับ Y

ขั้นที่ 2 สมมติฐาน $H_0: \beta_i = 0$

$H_1: \beta_i \neq 0$; $i = 1, 2, \dots, k$

สถิติทดสอบ $t = \frac{b_i - c}{s_{b_i}}$ หรือใช้สถิติทดสอบ z ถ้า n มีค่ามาก

เขตปฏิเสธสมมติฐาน H_0

จะปฏิเสธสมมติฐาน H_0 เมื่อ $t > t_{1-\alpha/2; n-k-1}$ หรือ $t < -t_{1-\alpha/2; n-k-1}$

หรือกล่าวว่าจะปฏิเสธ H_0 ถ้า $|t| > t_{1-\alpha/2; n-k-1}$ หรือ Significance ของสถิติ t < ระดับนัยสำคัญที่กำหนด (α)

ตารางที่ 2.2 การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอย β_i ในขั้นที่ 2

การทดสอบแบบด้านเดียว		การทดสอบแบบสองด้าน
สมมติฐาน	$H_0: \beta_i \leq 0$ $H_1: \beta_i > 0$	$H_0: \beta_i \geq 0$ $H_1: \beta_i < 0$
สถิติทดสอบ	$t = b_i / s_{b_i}$	
เขตปฏิเสธ H_0	$t > t_{1-\alpha/2; n-k-1}$ หรือ 1. $\frac{sig.}{2} < \alpha$ และ 2. $t > 0$	$t < -t_{1-\alpha/2; n-k-1}$ หรือ 1. $\frac{sig.}{2} < \alpha$ และ 2. $t < 0$
	$ t > t_{1-\alpha/2; n-k-1}$ หรือ Sig. ของ $t < \alpha$	

2.3.6 สัมประสิทธิ์การตัดสินใจเชิงพหุ (Multiple Coefficient of Determination : R^2 หรือ r^2)

สัมประสิทธิ์การตัดสินใจเชิงพหุ เป็นสัดส่วนหรือเปอร์เซ็นต์ที่ตัวแปรอิสระ (X_1, X_2, \dots, X_k) สามารถอธิบายการเปลี่ยนแปลงของ Y ได้ หรือกล่าวได้ว่าสัมประสิทธิ์การตัดสินใจเชิงพหุเป็นสัดส่วนหรือเปอร์เซ็นต์ความผันแปร Y ที่มีสาเหตุเนื่องจากความผันแปรของ $X_1, X_2,$ และ X_k โดยที่สัมประสิทธิ์การตัดสินใจเชิงพหุจะใช้สัญลักษณ์ $R^2_{Y,1,2,\dots,k}$ แต่โดยทั่วไปจะใช้ R^2

$$R^2 = r^2 = \frac{\text{ความผันแปรของ } Y \text{ เนื่องจากอิทธิพลของ } X_1, X_2, \dots, X_k}{\text{ความแปรผันทั้งหมด}} = \frac{SSR}{SST} \quad (2.59)$$

โดยที่ $0 \leq R^2, r^2 \leq 1$

ถ้า R^2 ที่ใกล้ 1 จะหมายถึง X_1, X_2, \dots, X_k มีความสัมพันธ์กับ Y มาก แต่ถ้า R^2 เข้าใกล้ 0 หมายถึงค่า X_1, X_2, \dots, X_k มีความสัมพันธ์กับ Y น้อย

เนื่องจาก SSR จะเพิ่มขึ้นถ้าเพิ่มตัวแปรอิสระ เช่น เดิมมี X_1 และ X_2 มีความสัมพันธ์กับ Y แต่ถ้าเพิ่มตัวแปรอิสระ X_3 เข้าในสมการถดถอย จะได้ว่า

$$SSR(X_1, X_2, X_3) > SSR(X_1, X_2)$$

โดยที่ $SSR(X_1, X_2, X_3)$ หมายถึง SSR ของสมการความถดถอยที่มีตัวแปรอิสระ X_1, X_2 และ X_3

และ $SSR(X_1, X_2)$ หมายถึง SSR ของสมการความถดถอยที่มีตัวแปรอิสระ X_1 และ X_2

ดังนั้น เมื่อเพิ่มตัวแปรอิสระเข้าสมการความถดถอยจะทำให้ค่า R^2 มากขึ้นทั้งที่ตัวแปรอิสระ X ที่เพิ่มอาจจะไม่มีความสัมพันธ์กับ Y เลยก็ได้ จึงได้มีการปรับค่า R^2 ให้ถูกต้องขึ้น เรียกว่า *Adjust R^2* โดยที่

$$R_a^2 = Adjust R^2 = 1 + \frac{(n-1)}{(n-k-1)}(R^2 - 1) \quad (2.60)$$

กรณีที่มีตัวแปรอิสระหลายตัว จะพิจารณาจากค่า R_a^2 มากกว่าค่า R^2

2.3.7 สัมประสิทธิ์สหสัมพันธ์เชิงพหุ (Multiple Coefficient of Correlation)

ค่าของสัมประสิทธิ์สหสัมพันธ์เชิงพหุ ได้จากการถอดรากที่สองของสัมประสิทธิ์การตัดสินใจเชิงพหุ

ดังนั้น สัมประสิทธิ์สหสัมพันธ์เชิงพหุ $= R_{Y,12\dots k} = R = \sqrt{R^2_{Y,12\dots k}}$ โดยที่ $0 \leq R \leq 1$

โดยที่ สัมประสิทธิ์สหสัมพันธ์เชิงพหุ หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง Y กับ \hat{Y} ถ้าสัมประสิทธิ์สหสัมพันธ์เชิงพหุมีค่ามากกว่าแสดงว่าค่า Y ใกล้เคียงกับค่า \hat{Y} มาก

1. R มีค่าเข้าใกล้ 0 แสดงว่า Y มีความสัมพันธ์กับ X_1, X_2, \dots, X_k น้อยมาก และถ้า $R = 0$ แสดงว่า Y ไม่มีความสัมพันธ์กับ X_1, X_2, \dots, X_k เลย
2. R มีค่าเข้าใกล้ 1 แสดงว่า Y มีความสัมพันธ์กับตัวแปรอิสระทั้ง k ตัวมีมาก

2.3.8 สัมประสิทธิ์สหสัมพันธ์บางส่วน (Coefficient of Partial Correlation)

เป็นการวัดค่าระดับความสัมพันธ์เชิงเส้นระหว่าง Y และ X ตัวใดตัวหนึ่ง โดยให้ X ตัวอื่น ๆ มีค่าคงที่ ถ้ามีตัวแปรอิสระ 2 ตัว คือ X_1 และ X_2 ที่มีความสัมพันธ์กับตัวแปรตาม Y ซึ่งสมการสำหรับคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรคู่ใดคู่หนึ่ง เป็นดังนี้

$$r_{12} = \frac{\sum (X_1 - \bar{X}_1)(X_2 - \bar{X}_2)}{\sqrt{\sum (X_1 - \bar{X}_1)^2 \sum (X_2 - \bar{X}_2)^2}} = \frac{\sum x_1 x_2}{\sqrt{(\sum x_1^2)(\sum x_2^2)}}$$

โดยที่ $y = Y - \bar{Y}$, $x_i = X_i - \bar{X}_i$, $i=1,2$

$$r_{Y1} = \frac{\sum (Y - \bar{Y})(X_1 - \bar{X}_1)}{\sqrt{\sum (Y - \bar{Y})^2 \sum (X_1 - \bar{X}_1)^2}} = \frac{\sum yx_1}{\sqrt{(\sum y^2)(\sum x_1^2)}}$$

$$r_{Y2} = \frac{\sum (Y - \bar{Y})(X_2 - \bar{X}_2)}{\sqrt{\sum (Y - \bar{Y})^2 \sum (X_2 - \bar{X}_2)^2}} = \frac{\sum yx_2}{\sqrt{(\sum y^2)(\sum x_2^2)}}$$

ดังนั้นจึงสามารถคำนวณหาสัมประสิทธิ์สหสัมพันธ์บางส่วนได้ดังนี้

$$r_{Y1.2} = \frac{r_{Y1} - r_{Y2} \cdot r_{12}}{\sqrt{(1 - r_{Y2}^2)(1 - r_{12}^2)}} \quad (2.61)$$

$$r_{Y2.1} = \frac{r_{Y2} - r_{Y1} \cdot r_{12}}{\sqrt{(1 - r_{Y1}^2)(1 - r_{12}^2)}} \quad (2.62)$$

โดยที่ $r_{Y1.2}$ = สัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่าง Y กับ X_1 เมื่อกำหนดให้ X_2 คงที่

$r_{Y2.1}$ = สัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่าง Y กับ X_2 เมื่อกำหนดให้ X_1 คงที่

2.3.9 การตรวจสอบสมมติฐานของการวิเคราะห์การถดถอย

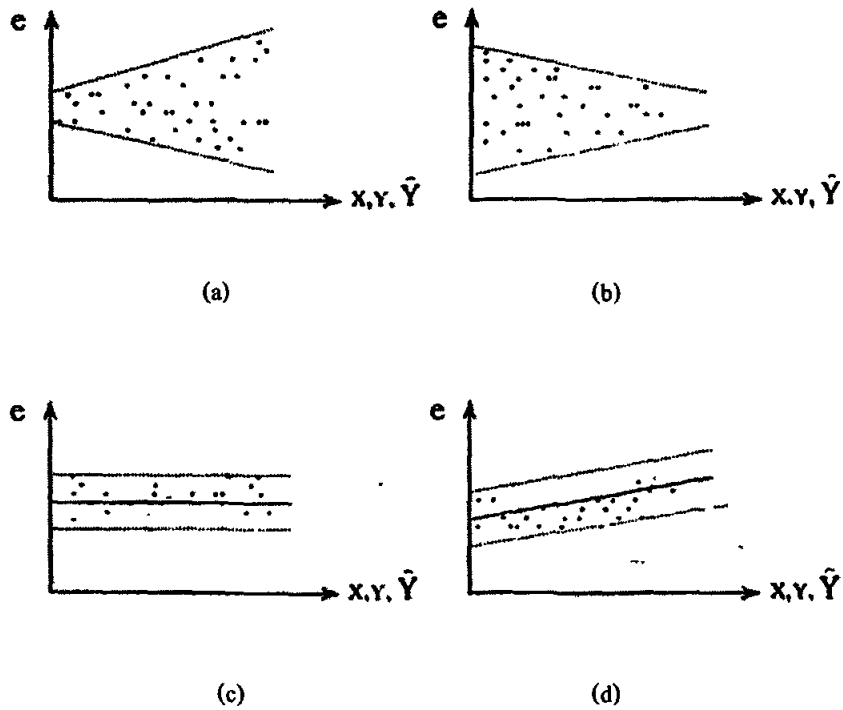
ก่อนที่จะพยายามค่า Y ด้วยสมการถดถอยโดยกำหนด X นั้น ค่าความคลาดเคลื่อน (e) จะต้องมีคุณสมบัติตามสมมติฐานทั้ง 4 ข้อ ดังนี้

2.3.9.1 ค่า X ต้องเป็นค่าที่กำหนดไว้ล่วงหน้าหรือทราบค่า

2.3.9.2 ค่าความคลาดเคลื่อน e มีค่าเฉลี่ย = 0 $E(e) = 0$ หรือ $V(e) = \sigma^2$

การตรวจสอบว่า $E(e) = 0$ เนื่องจากใช้วิธีกำลังสองน้อยที่สุดในการประมาณ β_0 ด้วย a และ β_1 ด้วย b , ซึ่งจะทำให้ $\sum e_i = 0$ ดังนั้น $E(e) = 0$

การตรวจสอบว่า $V(e) = V(Y) = \sigma^2$ = ค่าคงที่ จะพิจารณาจากกราฟแสดงความสัมพันธ์ระหว่างค่า e กับ \hat{Y} ถ้า $V(e)$ ไม่เท่ากับค่าคงที่ เรียกว่า เกิดปัญหา Heteroscedastic ดังแสดงในรูปที่ 2.17 (a) และ 2.17 (b) แต่ถ้า $V(e) = \sigma^2$ ค่าคงที่ตามสมมติฐานที่ตั้งไว้ ดังรูปที่ 2.17 (c) และ 2.17 (d) คือ ค่า e จะกระจายอยู่รอบๆ ศูนย์ หรือมีค่าในช่วงหนึ่งแคบๆ ไม่ว่า X หรือ Y เปลี่ยนไปอย่างไร



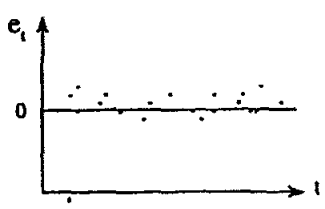
รูปที่ 2.17 กราฟแสดงความสัมพันธ์ระหว่างค่า e กับ \hat{Y}

จากรูปที่ 2.17 (a) จะพบว่า σ_e^2 จะมีค่าน้อยเมื่อ \hat{Y} มีค่าน้อย และเมื่อ \hat{Y} มีค่ามากจะทำให้ σ_e^2 มีค่ามากด้วย ขณะที่รูปที่ 2.17 (c) และ 2.17 (d) ค่า σ_e^2 จะคงที่เมื่อ \hat{Y} เปลี่ยนไป ส่วนรูปที่ 2.17 (b) ค่า σ_e^2 จะมีค่าน้อยเมื่อ \hat{Y} มีค่ามาก

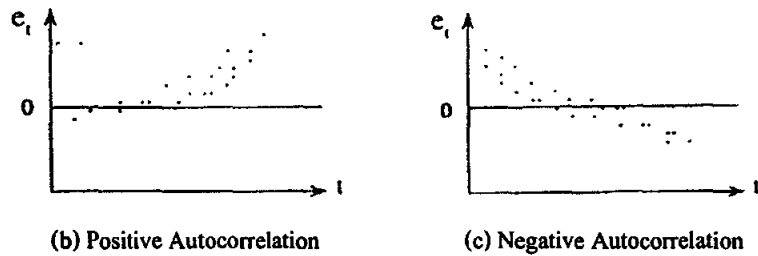
2.3.9.3 การตรวจสอบ e_t และ e_{t-1} เป็นอิสระต่อกัน

การตรวจสอบความเป็นอิสระต่อกันของ e_t และ e_{t-1} โดยที่ $e_t = Y_t - \hat{Y}_t$ และ $e_{t-1} = Y_{t-1} - \hat{Y}_{t-1}$ ซึ่งสามารถทำได้ 2 วิธี คือ

1. เขียนกราฟแสดงความสัมพันธ์ระหว่าง e_t กับ t



(a) No Autocorrelation



รูปที่ 2.18 กราฟแสดงความสัมพันธ์ระหว่าง e_t กับ t

ถ้า e_t และ e_{t-1} มีความสัมพันธ์จะเรียกว่า เกิด Autocorrelation ดังแสดงในรูปที่ 2.18 (b) และ รูปที่ 2.18 (c) โดยรูปที่ 2.18 (b) แสดงความสัมพันธ์ระหว่าง e_t กับ e_{t-1} ในทิศทางบวก ที่เรียกว่า Positive Autocorrelation ส่วนรูปที่ 2.18 (c) แสดงความสัมพันธ์ในทิศทางลบ เรียกว่า Negative Autocorrelation และในรูปที่ 2.18 (a) แสดงความสัมพันธ์ของ e_t กับ e_{t-1} ถ้าเขียนกราฟระหว่าง e_t กับ t ได้ดังรูปที่ 2.18 (a) แสดงว่าค่าความคลาดเคลื่อนเป็นอิสระกัน

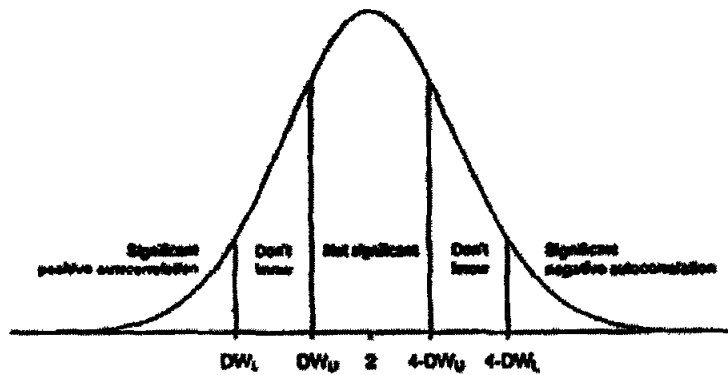
2. ใช้สถิติทดสอบ Durbin-Watson (DW)

เป็นการทดสอบความสัมพันธ์ของ e_t กับ e_{t-1} โดยที่ t เป็นช่วงเวลา สถิติสำหรับการทดสอบ คือ

$$DW = \frac{\sum_{t=1}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (3.60)$$

โดยที่ $0 \leq DW \leq 4$ และมีคุณสมบัติดังนี้

- ถ้าค่า DW มีค่าอยู่ระหว่าง 1.5-2.5 แสดงว่าค่าความคลาดเคลื่อน e_t เป็นอิสระต่อกัน
- ถ้าค่า $DW < 1.5$ แสดงว่าค่าความคลาดเคลื่อนแต่ละค่ามีความสัมพันธ์กันในทางบวก และ ถ้ายิ่งเข้าใกล้ 0 แสดงว่าค่าความคลาดเคลื่อนแต่ละค่ามีความสัมพันธ์กันอย่างมาก
- ถ้าค่า $DW > 2.5$ แสดงว่าค่าความคลาดเคลื่อนแต่ละค่ามีความสัมพันธ์กันในทางลบ และ ถ้ายิ่งเข้าใกล้ 4 แสดงว่าค่าความคลาดเคลื่อนแต่ละค่ามีความสัมพันธ์กันอย่างมาก



รูปที่ 2.19 การแจกแจงของ Durbin-Watson

จากรูปที่ 2.19 ถ้าเป็นการทดสอบแบบ 2 ข้าง จะใช้ $\frac{\alpha}{2}$ แต่ถ้าทดสอบแบบข้างเดียวจะใช้ α ซึ่งการแจกแจงของ Durbin-Watson จะเป็น 5 ช่วง ดังนี้

ช่วงที่ 1 น้อยกว่า DW_L

ช่วงที่ 2 อยู่ระหว่าง DW_L และ DW_U

ช่วงที่ 3 อยู่ระหว่าง DW_U และ $4-DW_U$

ช่วงที่ 4 อยู่ระหว่าง $4-DW_U$ และ $4-DW_L$

ช่วงที่ 5 มากกว่า $4-DW_L$

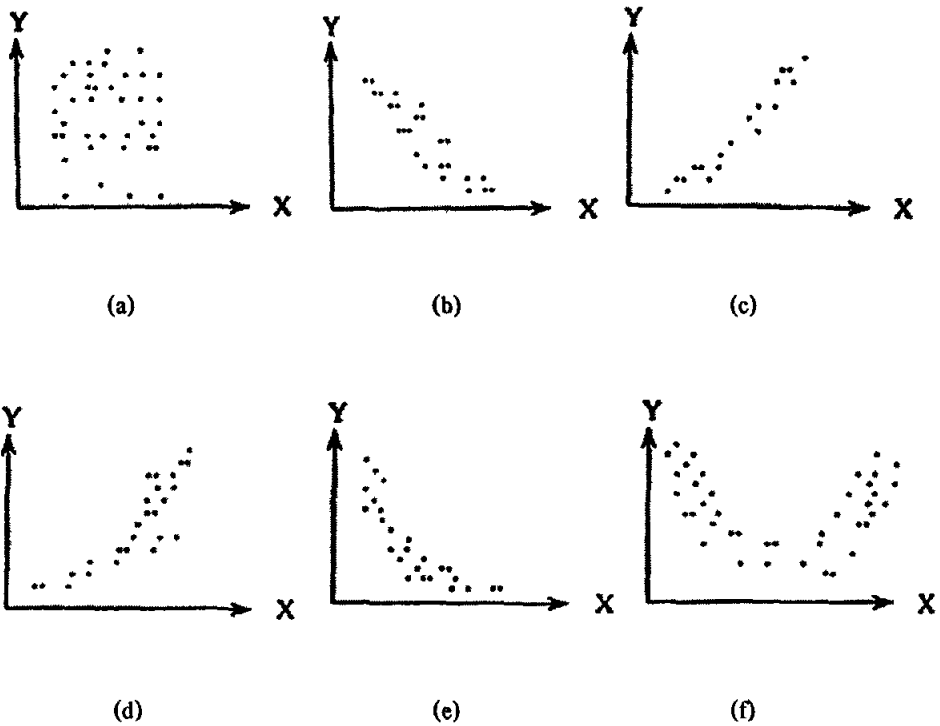
ถ้าสถิติทดสอบ DW อยู่ในช่วงที่ 1 หรือ 5 แสดงว่าค่าความคลาดเคลื่อนมีความสัมพันธ์กัน แต่ถ้า DW อยู่ในช่วงที่ 3 แสดงว่าค่าความคลาดเคลื่อนไม่มีความสัมพันธ์กัน และถ้า DW อยู่ในช่วงที่ 2 หรือ 4 แสดงว่ายังไม่สามารถสรุปได้ว่าค่าความคลาดเคลื่อนมีความสัมพันธ์กันหรือไม่

2.3.9.4 การตรวจสอบว่า e_t มีการแจกแจงแบบปกติ

การทดสอบว่า e_t มีการแจกแจงแบบปกติหรือไม่นั้น สามารถตรวจสอบได้โดยใช้กราฟ เช่น Histogram และ Normal Probability Plot หรือใช้สถิติทดสอบ เช่น Kolmogorov-Smirnov และ Lilliefort's Test

2.3.10 การตรวจสอบรูปแบบความสัมพันธ์ของตัวแปรเชิงปริมาณ 2 ตัว

ความสัมพันธ์ระหว่างตัวแปร 2 ตัวนั้น อาจจะมีหลายรูปแบบ เช่น สัมพันธ์กันในเชิงเส้น เส้นโค้ง พาราโบลา เป็นต้น ดังนั้น ก่อนที่จะคำนวณค่าสถิติเพื่อแสดงระดับความสัมพันธ์จึงควรมีการตรวจสอบลักษณะของความสัมพันธ์เสียก่อน วิธีการตรวจสอบที่ง่ายที่สุด คือ การ Plot กราฟแสดงความสัมพันธ์ของตัวแปรทั้งสอง ดังแสดงในรูปที่ 2.19 เมื่อได้กราฟที่แสดงถึงความสัมพันธ์ระหว่างตัวแปร X และ Y แล้วก็จะทราบว่าคุณสมบัติที่ได้นั้นอยู่ในรูปแบบใด จากนั้นจึงเลือกเทคนิคการวิเคราะห์ที่เหมาะสมกับรูปแบบนั้นๆ



รูปที่ 2.20 กราฟแสดงความสัมพันธ์ระหว่าง X และ Y

โดยที่ รูปที่ 2.20 (a) หมายถึง ตัวแปร X และ Y ไม่มีความสัมพันธ์กัน

รูปที่ 2.20 (b) หมายถึง ตัวแปร X และ Y มีความสัมพันธ์กันในรูปเชิงเส้นและเป็นลบ หรือมีความสัมพันธ์กันในทิศทางตรงข้าม นั่นคือ ถ้า X เพิ่มขึ้น Y จะลดลง แต่ถ้า X ลดลง Y จะเพิ่มขึ้น

รูปที่ 2.20 (c) หมายถึง ตัวแปร X และ Y มีความสัมพันธ์กันในรูปเชิงเส้นและเป็นบวก หรือมีความสัมพันธ์กันในทิศทางเดียวกัน นั่นคือ ถ้า X เพิ่ม Y จะเพิ่มขึ้น แต่ถ้า X ลดลง Y จะลดลงด้วย

รูปที่ 2.20 (d) หมายถึง ตัวแปร X และ Y มีความสัมพันธ์กันในรูป Exponential และเป็นบวก

รูปที่ 2.20 (e) หมายถึง ตัวแปร X และ Y มีความสัมพันธ์กันในรูป Exponential และเป็นลบ

รูปที่ 2.20 (f) หมายถึง ตัวแปร X และ Y มีความสัมพันธ์กันในรูปพาราโบลา

บทที่ 3

วิธีการดำเนินการวิจัย

บทนี้จะกล่าวถึงการดำเนินงานสองส่วน ส่วนแรกคือการนำโครงข่ายประสาทเทียมมาใช้ในการทำนายรูปร่างของรอยเชื่อม โดยใช้โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าและใช้กฎการเรียนรู้แบบแพร่ค่าย้อนกลับของ Levenberg-Marquardt Algorithm ในโปรแกรม MATLAB R2007a ซึ่งจะนำมาเรียนรู้ความสัมพันธ์ระหว่างปัจจัยในการเชื่อม ได้แก่ Arc current, Arc Voltage, Travel speed และ Wire speed กับรูปร่างของรอยเชื่อมที่ต้องการทำนายคือ Bead width, Penetration, Reinforcement height และ Cross section โดยมีการแบ่งชุดข้อมูลออกเป็นสองกลุ่มคือ ชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบ ส่วนที่สองคือการใช้วิธีการวิเคราะห์ความถดถอยเชิงพหุมาใช้ในการทำนายรูปร่างของรอยเชื่อมโดยทำการสร้างแบบจำลองโดยใช้โปรแกรม SPSS 15.0 for Windows

3.1 วิธีการดำเนินงานของโครงข่ายประสาทเทียม

วิธีการดำเนินงานประกอบด้วย 5 ขั้นตอน ได้แก่

1. ศึกษาทฤษฎีที่เกี่ยวข้องกับโครงข่ายประสาทเทียม
2. ศึกษาการใช้โครงข่ายประสาทเทียมในโปรแกรม MATLAB R2007a
3. นำข้อมูลที่ได้นำมาทำการ Normalization
4. นำข้อมูลที่ผ่านการ Normalization มาทำการฝึกสอนและทดสอบโครงข่ายประสาทเทียม เพื่อหาโครงสร้างของโครงข่ายประสาทเทียม ซึ่งผลจากการฝึกสอนและทดสอบจะออกมาในรูปของค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE)
5. เลือกโครงสร้างที่ให้ค่าความผิดพลาดกำลังสองเฉลี่ยน้อยที่สุดจากการทดสอบ โครงสร้าง และทำการทำนายโดยใช้ข้อมูลชุดใหม่เพื่อนำไปเปรียบเทียบกับวิธีการวิเคราะห์ความถดถอยเชิงพหุ

3.1.1 ข้อมูลจริงทั้งหมด

ข้อมูลจริงทั้งหมดที่นำมาใช้ในการฝึกสอนและทดสอบโครงข่ายประสาทเทียมมี 42 ข้อมูล ดังแสดงในตารางที่ 3.1

ตารางที่ 3.1 ข้อมูลจริงทั้งหมด 42 ข้อมูล

ลำดับที่	Arc current (A)	Arc voltage (V)	Travel speed (mm/s)	Wire speed (mm/s)	Bead width (mm)	Penetration (mm)	Reinforcement height (mm)	Cross section area (mm ²)
1	150	9.6	1.91	2.12	6.95	2.60	0.37	1.87
2	150	9.6	1.91	4.23	7.04	2.70	0.64	2.87
3	150	9.6	1.91	6.35	6.93	2.63	0.86	4.13
4	150	9.6	1.91	8.47	6.87	2.47	1.02	4.94
5	100	9.0	1.19	2.12	5.49	1.75	0.59	2.55
6	100	9.0	1.19	4.23	5.23	1.77	0.88	3.33
7	100	9.0	1.19	6.35	5.51	1.50	1.43	5.34
8	100	9.0	1.19	10.58	5.47	1.18	2.10	8.68
9	100	9.0	1.82	2.12	4.60	1.74	0.60	2.05
10	100	9.0	1.82	6.35	4.64	1.62	1.07	3.51
11	100	9.0	1.82	8.47	4.68	1.29	1.25	4.18
12	150	9.6	2.88	4.23	6.81	1.93	0.47	2.12
13	150	9.6	2.88	6.35	7.23	2.00	0.60	2.55
14	150	9.6	2.88	8.47	7.05	1.83	0.78	3.51
15	150	9.6	1.44	2.12	7.45	3.21	0.42	2.42
16	150	9.6	1.44	4.23	7.58	3.10	0.63	3.47
17	100	9.0	0.89	2.12	5.59	1.29	0.66	2.75
18	100	9.0	0.89	4.23	5.66	1.85	1.09	4.53
19	100	9.0	0.89	6.35	5.53	1.36	1.64	6.59
20	200	10.5	4.19	2.12	7.66	1.33	0.21	0.92
21	200	10.5	4.19	2.12	6.64	1.51	0.21	0.83
22	200	10.5	4.19	4.23	8.00	1.50	0.34	1.58
23	200	10.5	4.19	10.58	7.86	1.06	0.59	2.98
24	200	10.5	4.19	12.70	8.12	1.11	0.69	3.66
25	200	10.5	4.19	16.93	8.66	0.72	0.82	4.73
26	200	10.5	4.19	21.17	8.75	0.55	0.94	5.61
27	200	10.5	2.12	8.47	9.76	2.50	0.78	4.93
28	200	10.5	2.12	16.93	11.09	1.11	1.18	8.51
29	200	10.5	2.79	12.70	9.45	1.58	0.83	5.11

ตารางที่ 3.1 ข้อมูลจริงทั้งหมด 42 ข้อมูล (ต่อ)

ลำดับที่	Arc current (A)	Arc voltage (V)	Travel speed (mm/s)	Wire speed (mm/s)	Bead width (mm)	Penetration (mm)	Reinforcement height (mm)	Cross section area (mm ²)
30	200	10.5	2.79	16.93	9.58	0.89	1.05	6.52
31	200	10.5	2.79	8.47	9.31	2.28	0.63	3.64
32	100	9.0	1.19	8.47	5.35	1.43	1.89	7.27
33	100	9.0	1.82	4.23	5.70	1.73	0.69	2.64
34	150	9.6	2.88	10.58	6.06	1.94	0.97	4.23
35	150	9.6	1.44	6.35	7.55	2.86	0.95	5.06
36	100	9.0	0.89	8.47	5.78	1.44	2.13	9.44
37	200	10.5	4.19	6.35	7.80	1.58	0.44	1.85
38	200	10.5	4.19	8.47	8.09	1.80	0.53	2.59
39	200	10.5	4.19	14.82	8.27	0.87	0.75	4.00
40	200	10.5	2.12	2.12	10.08	2.16	0.42	2.55
41	200	10.5	2.12	12.70	10.24	1.76	0.97	6.48
42	200	10.5	2.79	2.12	9.35	2.05	0.30	1.74

3.1.2 การจัดเตรียมข้อมูล

ข้อมูลที่ใช้ในการฝึกสอนและทดสอบ โครงข่ายประสาทเทียม มีดังนี้

1. ชุดข้อมูลที่ใช้สำหรับการฝึกสอน

- Arc current 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Arc voltage 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Travel speed 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Wire speed 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Bead width 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Penetration 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Reinforcement height 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1
- Cross section 30 ข้อมูล คือ ข้อมูลในลำดับที่ 1 ถึง 30 ในตารางที่ 3.1

2. ชุดข้อมูลที่ใช้สำหรับการทดสอบ

- Arc current 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1
- Arc voltage 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1
- Travel speed 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1

- Wire speed 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1
- Bead width 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1
- Penetration 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1
- Reinforcement height 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1
- Cross section 12 ข้อมูล คือ ข้อมูลในลำดับที่ 31 ถึง 42 ในตารางที่ 3.1

การทำนายรูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียมในปริภูมิพารามิเตอร์นี้ โครงข่ายประสาทเทียมจะทำการพยากรณ์ตัวแปรด้านนอก คือ Bead width, Penetration, Reinforcement height และ Cross section โดยใช้ตัวแปรด้านเข้าที่มีความสัมพันธ์กับตัวแปรด้านนอก คือ Arc current, Arc voltage, Travel speed และ Wire speed ซึ่งข้อมูลทั้งหมดจะถูกทำ Normalization ให้มีค่าเฉลี่ย (Mean) เท่ากับศูนย์ และค่าเบี่ยงเบนมาตรฐาน (Variance) เท่ากับหนึ่ง และ Scaling เพื่อให้ข้อมูลอยู่ในช่วงของฟังก์ชันถ่ายโอน ก่อนที่จะทำการฝึกสอนและทดสอบตามหัวข้อ 2.1.9 ซึ่งจะทำให้กระบวนการฝึกสอนเร็วขึ้นและมีความแม่นยำสูงขึ้น หลังจากได้ผลลัพธ์แล้วเอาท์พุทที่ได้จะต้องแปลงกลับไปเป็นค่าจริงด้วยการ Unnormalization และ Unscaling

3.1.3 การทำนายรูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม

ก่อนที่จะฝึกสอนโครงข่ายประสาทเทียมเพื่อนำไปใช้ทำนายรูปร่างของรอยเชื่อม จะต้องกำหนดสถาปัตยกรรมของโครงข่ายประสาทเทียมเสียก่อน เพื่อให้โครงข่ายประสาทเทียมมีจำนวนเซลล์ประสาทของตัวแปรด้านเข้า และในชั้นตัวแปรด้านนอกสอดคล้องกับจำนวนตัวแปรด้านเข้าที่ป้อนให้แก่โครงข่าย และจำนวนตัวแปรด้านออกที่เป็นค่าผลลัพธ์ที่ต้องการ รวมถึงการกำหนดจำนวนเซลล์ประสาทในชั้นซ่อนและการเลือกใช้ฟังก์ชันถ่ายโอนที่เหมาะสม

การนำโครงข่ายประสาทเทียมมาใช้ในการศึกษาการทำนายรูปร่างของรอยเชื่อมนั้นมียุคประกอบหลายอย่างที่มีผลต่อการเรียนรู้ของโครงข่ายประสาทเทียม การเรียนรู้ที่ดีจะส่งผลทำให้โครงข่ายประสาทเทียมมีความสามารถในการคำนวณหาคำตอบได้แม่นยำขึ้น ดังนั้นก่อนที่จะทำการฝึกสอนโครงข่ายประสาทเทียมจึงต้องมีการพิจารณาถึงองค์ประกอบที่มีผลต่อการเรียนรู้และประสิทธิภาพของโครงข่ายประสาทเทียม ดังนี้

3.1.3.1 จำนวนเซลล์ประสาทของตัวแปรด้านเข้าและชั้นตัวแปรด้านนอก

ปริภูมิพารามิเตอร์นี้ได้กำหนดจำนวนเซลล์ประสาทของตัวแปรด้านเข้า 4 เซลล์ จากปัจจัยที่มีความสัมพันธ์กับรูปร่างของรอยเชื่อม คือ Arc current, Arc voltage, Travel speed และ Wire speed ส่วนจำนวนเซลล์ประสาทในชั้นตัวแปรด้านนอก 4 เซลล์ จะกำหนดจากค่าผลลัพธ์ที่ต้องการให้โครงข่ายประสาทเทียมคำนวณออกมา คือ Bead width, Penetration, Reinforcement height และ Cross section ซึ่งสามารถสรุปได้ดังนี้

1. ข้อมูลที่นำมาใช้เป็นตัวแปรด้านเข้า
 - Arc current
 - Arc voltage
 - Travel speed
 - Wire speed
2. ข้อมูลที่นำมาใช้เป็นตัวแปรด้านนอก หรือค่าผลลัพธ์ที่ต้องการ

- Bead width
- Penetration
- Reinforcement height
- Cross section

3.1.3.2 ชนิดของฟังก์ชันถ่ายโอน

ชนิดของฟังก์ชันถ่ายโอนเป็นองค์ประกอบที่มีผลต่อการเรียนรู้ของโครงข่ายประสาทเทียมซึ่งในปริภูมิตฤษฎีการจะศึกษาการใช้ฟังก์ชันถ่ายโอนแบบเชิงเส้นและแบบไม่เชิงเส้นกับเซลล์ประสาทในชั้นซ่อนและเซลล์ประสาทในชั้นตัวแปรค่าออก คือ ฟังก์ชันถ่ายโอนชนิด Log-sigmoid ฟังก์ชันถ่ายโอนชนิด Tan-sigmoid และฟังก์ชันถ่ายโอนชนิด Linear ดังแสดงสมการที่ 3.1 ถึง 3.3

$$f(n) = \frac{1}{1 + e^{-n}} \quad (3.1)$$

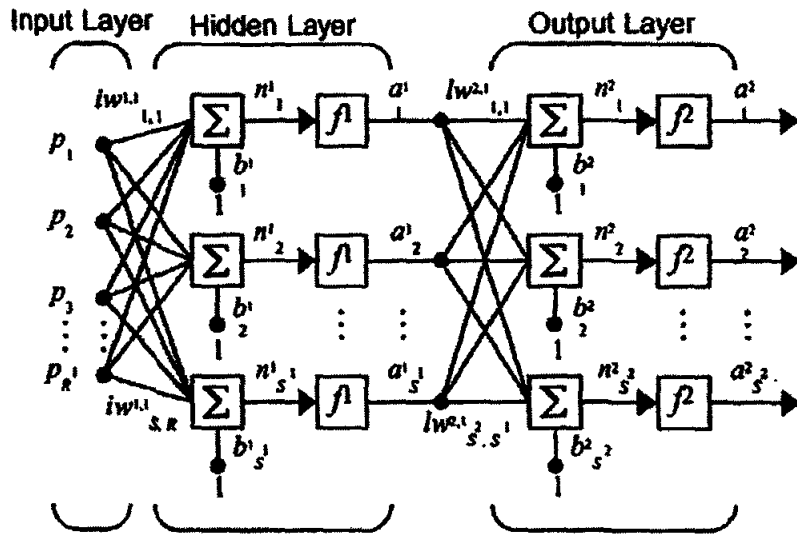
$$f(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \quad (3.2)$$

$$f(n) = n \quad (3.3)$$

เหตุผลที่นำฟังก์ชันถ่ายโอนทั้งสองชนิดนี้มาใช้ก็เพราะว่าเป็นฟังก์ชันที่สามารถหาอนุพันธ์ได้ และสร้างขอบเขตการตัดสินใจในลักษณะโค้งมนได้ โครงข่ายประสาทเทียมที่ใช้ฟังก์ชันถ่ายโอนแบบนี้ จะสามารถนำไปประยุกต์ใช้กับปัญหาที่มีความซับซ้อน และมีลักษณะไม่เชิงเส้นได้ดี

3.1.3.3 สถาปัตยกรรมของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่ใช้ในปริภูมิตฤษฎีการฉบับนี้ได้สร้างแบบจำลองจาก Neural Network Toolbox ของโปรแกรม MATLAB R2007a และใช้กฎการเรียนรู้แบบแพร่ค่าย้อนกลับของ Levenberg-Marquardt Algorithm ซึ่งจะมีการจัดเรียงเซลล์ประสาทต่อกันในทิศทางที่ส่งข้อมูลไปข้างหน้า และไม่เชื่อมต่อในทิศทางย้อนกลับ โดยจะเรียงกันเป็นชั้นๆ คือ อินพุต 1 ชั้น ชั้นซ่อน 1 ชั้น และชั้นเอาต์พุต 1 ชั้น ซึ่งเซลล์ประสาทแต่ละชั้นจะเชื่อมต่อกับเซลล์ประสาทในชั้นก่อนหน้า และชั้นถัดไปด้วยค่าน้ำหนักและค่าไบอัสที่สามารถปรับเปลี่ยนได้ในขณะที่ทำการเรียนรู้



รูปที่ 3.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมที่ใช้ในการพยากรณ์

จากรูปที่ 3.1 แสดงโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่มี 1 ชั้นซ่อนและ 1 ชั้นเอาต์พุตซึ่งค่า a^1 และ a^2 จะสามารถหาได้โดยใช้สมการที่ 3.4 และ 3.5 ดังนี้

$$a^1 = f^1(W^1 p + b^1) \quad (3.4)$$

$$a^2 = f^2(W^2 a^1 + b^2) \quad (3.5)$$

โดยที่

- p คือ ตัวแปรด้านเข้าของโครงข่ายประสาทเทียม
- w^1 คือ ค่าน้ำหนักที่เชื่อมต่อระหว่างตัวแปรด้านเข้าและชั้นซ่อน
- w^2 คือ ค่าน้ำหนักที่เชื่อมต่อระหว่างชั้นซ่อนและชั้นตัวแปรด้านออก
- b^1 คือ ค่าไบอัสของชั้นซ่อน
- b^2 คือ ค่าไบอัสของชั้นตัวแปรด้านออก
- n^1 คือ ผลรวมของข้อมูลด้านเข้าของชั้นซ่อน
- n^2 คือ ผลรวมของข้อมูลด้านเข้าของตัวแปรด้านออก
- f^1 คือ ฟังก์ชันถ่ายโอนของชั้นซ่อน
- f^2 คือ ฟังก์ชันถ่ายโอนของชั้นตัวแปรด้านออก
- a^1 คือ ตัวแปรด้านออกของชั้นซ่อน
- a^2 คือ ตัวแปรด้านออกของโครงข่ายประสาทเทียม

3.1.3.4 จำนวนเซลล์ประสาทของชั้นซ่อนที่เหมาะสม

ปฏิญานินพนธ์นี้ได้ใช้วิธีหาจำนวนเซลล์ประสาทของชั้นซ่อน ด้วยการเริ่มต้นจากจำนวน 1 เซลล์และทำการฝึกสอน ซึ่งในแต่ละครั้งค่าน้ำหนักและค่าไบอัสเริ่มต้นจะใช้เลขแบบสุ่ม จากนั้นจะทำการทดสอบและคำนวณค่าความผิดพลาดกำลังสองเฉลี่ย และทำการปรับเพิ่มจำนวนเซลล์ประสาทของชั้นซ่อนเพิ่มขึ้นครั้งละ 1 เซลล์ เพื่อหาโครงสร้างที่ให้ค่าความผิดพลาดกำลังสองเฉลี่ยต่ำที่สุด ซึ่งแต่ละโครงสร้างจะทำการฝึกสอนโดยทำการปรับค่าน้ำหนักและค่าไบอัสจำนวน 700 epochs ซึ่งพบว่าจำนวนเซลล์ประสาทของชั้นซ่อนที่เหมาะสมคือจำนวน 6 เซลล์

3.1.3.5 การปรับค่าน้ำหนักและค่าไบอัส

ปฏิญานินพนธ์นี้จะปรับค่าน้ำหนักและค่าไบอัสของโครงข่ายประสาทเทียมโดยใช้กฎการเรียนรู้แบบแพร่ค่าย้อนกลับของ Levenberg-Marquardt Algorithm โดยจะปรับค่าน้ำหนัก และค่าไบอัสเพื่อให้ค่าผิดพลาดในรอบต่อไปมีค่าลดลง ซึ่งค่าน้ำหนัก “w” และค่าไบอัส “b” ของรอบต่อไปจะถูกปรับใหม่ ตามสมการที่ 3.6 และ 3.7

$$w^{new} = w^{old} + [J^T(w) \cdot J(w) + \mu I]^{-1} J^T(w) \cdot e \quad (3.6)$$

$$b^{new} = b^{old} + [J^T(b) \cdot J(b) + \mu I]^{-1} J^T(b) \cdot e \quad (3.7)$$

โดยที่

$J(w)$	คือ	จาโคเบียนเมตริกซ์ของอนุพันธ์ค่าผิดพลาดเทียบกับค่าน้ำหนัก
$J(b)$	คือ	จาโคเบียนเมตริกซ์ของอนุพันธ์ค่าผิดพลาดเทียบกับค่าไบอัส
e	คือ	เมตริกซ์ของค่าผิดพลาดระหว่างเอาท์พุทเป้าหมายกับตัวแปรด้านออก
$J^T(w) \cdot e$	คือ	Gradient Factor ของค่าน้ำหนัก
$J^T(b) \cdot e$	คือ	Gradient Factor ของค่าไบอัส
μ	คือ	Learning Factor ซึ่งมีค่าเริ่มต้นเท่ากับ 0.001
I	คือ	Identity Matrix

ในการปรับค่าน้ำหนักและค่าไบอัสของโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าโดยใช้กฎการเรียนรู้แบบแพร่ค่าย้อนกลับของ Levenberg-Marquardt Algorithm จะมีเงื่อนไขที่เกี่ยวข้องกับการสิ้นสุดของการฝึกสอนดังนี้

1. ค่า Gradient Factor

ในการปรับค่าน้ำหนักและค่าไบอัสแต่ละรอบนั้นค่า Gradient Factor (ค่า $J^T(w) \cdot e$ หรือค่า $J^T(b) \cdot e$) จะเปลี่ยนแปลงไปโดยขึ้นอยู่กับค่าผิดพลาด (e) ใน epoch นั้นๆ ซึ่งการฝึกสอนจะสิ้นสุดลงเมื่อค่า Gradient Factor มีค่าน้อยกว่า $1e^{-10}$

2. ค่า Learning Factor

ในการปรับค่าน้ำหนักและค่าไบอัสแต่ละรอบนั้น ถ้าค่า Gradient Factor มีค่าเพิ่มขึ้นก็จะมีค่าปรับค่า μ ให้เพิ่มขึ้นจนกระทั่งค่า Gradient Factor มีแนวโน้มที่จะลดลง หลังจากนั้นค่า μ ก็จะถูกปรับลดลงเรื่อยๆ และค่า μ จะ

ถูกปรับเพิ่มขึ้นอีกครั้งเมื่อค่า Gradient Factor มีค่าเพิ่มขึ้น ซึ่งการฝึกสอนจะสิ้นสุดลงเมื่อค่า μ มีค่ามากกว่า $1e^{10}$

โดยค่า μ เริ่มต้นจะมีค่าเท่ากับ 0.001 และการปรับค่า μ ให้เพิ่มขึ้นก็จะทำโดยการคูณด้วยค่า Increased Factor ซึ่งมีค่าเท่ากับ 10 ส่วนการปรับค่า μ ให้ลดลงนั้นก็จะเป็นการคูณด้วยค่า Decreased Factor ซึ่งมีค่าเท่ากับ 0.1

3.1.3.6 การวัดประสิทธิภาพของโครงข่ายประสาทเทียม

การวัดความแม่นยำของโครงข่ายประสาทเทียมในปัญญานินทรรูปแบบนี้จะใช้ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean square error : MSE) จากค่าความคลาดเคลื่อนของค่าพยากรณ์ที่เป็นคำตอบของโครงข่ายประสาทเทียมที่ผ่านการฝึกสอนแล้วเปรียบเทียบกับคำตอบจริงดังนี้

$$MSE = \frac{1}{N} \sum_{i=1}^N (Output_{ANNs_i} - Output_{Actual_i})^2 \quad (3.8)$$

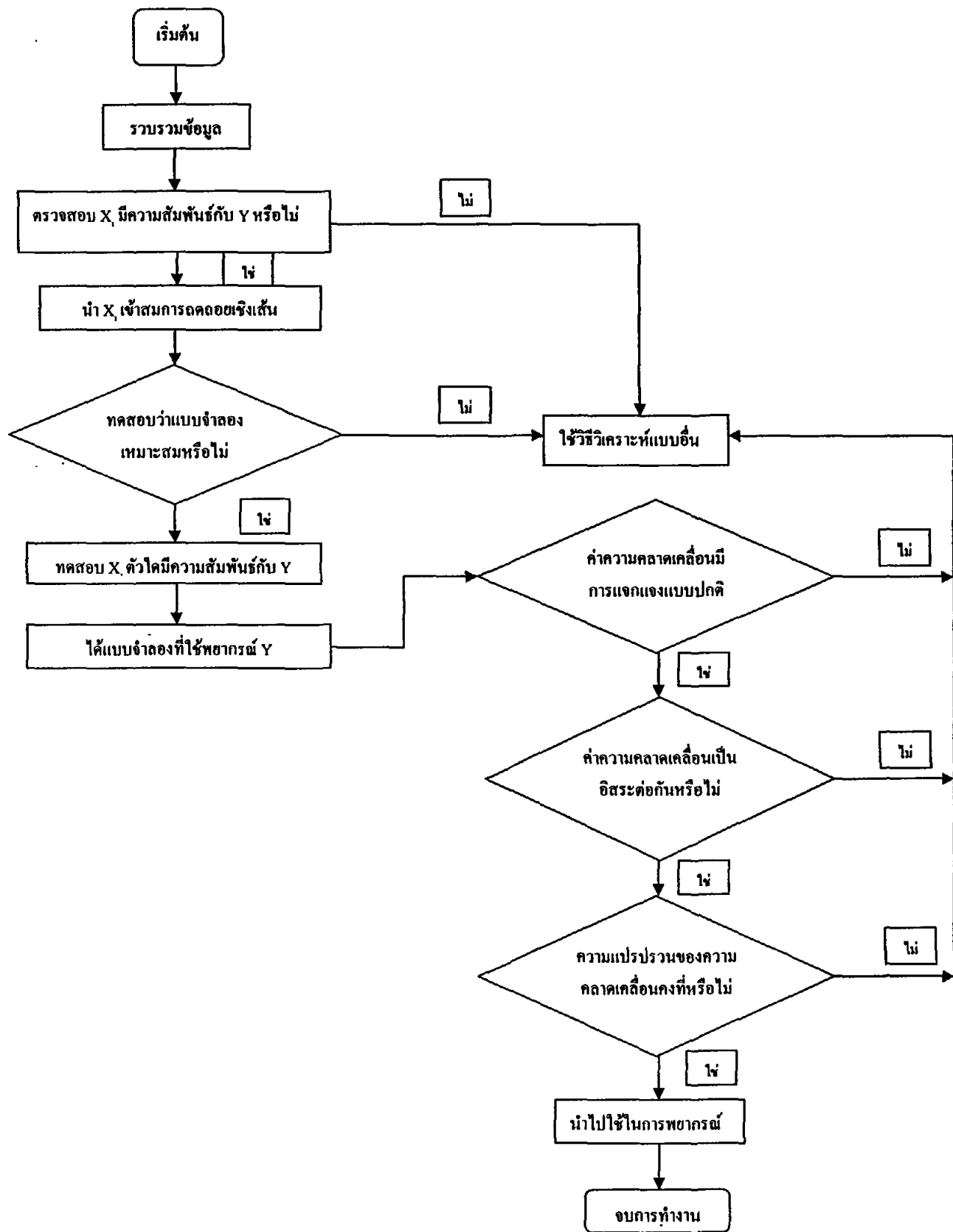
โดย

$Output_{ANNs_i}$	คือ คำตอบของโครงข่ายประสาทเทียม
$Output_{Actual_i}$	คือ คำตอบจริง
N	คือ จำนวนข้อมูลทดสอบ

3.2 การดำเนินงานของการวิเคราะห์ความถดถอยเชิงพหุ

ในขั้นตอนการดำเนินงานของการวิเคราะห์ความถดถอยเชิงพหุจะกำหนดตัวแปรที่ใช้ในการวิเคราะห์ ดังนี้

- VAR00001 = Arc current
- VAR00002 = Arc voltage
- VAR00003 = Travel speed
- VAR00004 = Wire speed
- VAR00005 = Arc current x Arc voltage
- VAR00006 = Arc current x Travel speed
- VAR00007 = Arc current x Wire speed
- VAR00008 = Arc voltage x Travel speed
- VAR00009 = Arc voltage x Wire speed
- VAR00010 = Travel speed x Wire speed
- VAR00011 = Arc current x Arc voltage x Travel speed
- VAR00012 = Arc current x Arc voltage x Wire speed
- VAR00013 = Arc current x Travel speed x Wire speed
- VAR00014 = Arc voltage x Travel speed x Wire speed
- VAR00015 = Arc current x Arc voltage x Travel speed x Wire speed



รูปที่ 3.2 ขั้นตอนการพยากรณ์รูปร่างของรอยเชื่อมโดยใช้การวิเคราะห์ความถดถอยเชิงพหุ

บทที่ 4

ผลการดำเนินงาน

บทนี้จะกล่าวถึงการนำโครงข่ายประสาทเทียมไปใช้พยากรณ์รูปร่างของรอยเชื่อม โดยจะพิจารณาผลของการปรับเปลี่ยนจำนวนเซลล์ประสาทในชั้นซ่อนและจำนวนของชั้นซ่อนที่ใช้ในการปรับสอนแตกต่างกัน และเปรียบเทียบผลการพยากรณ์ด้วยวิธีดังกล่าวกับวิธีพยากรณ์โดยใช้การวิเคราะห์ความถดถอยเชิงพหุ

4.1 การพยากรณ์รูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม

การพยากรณ์รูปร่างของรอยเชื่อม โดยใช้โครงข่ายประสาทเทียมในปริภูมิพิกัดนี้ จะศึกษาผลของจำนวนชั้นซ่อน จำนวนเซลล์ประสาทในชั้นซ่อนและชนิดของฟังก์ชันถ่ายโอนที่ใช้ว่ามีผลต่อประสิทธิภาพของโครงข่ายประสาทเทียมอย่างไร โดยทำการปรับเปลี่ยนโครงสร้างของโครงข่ายประสาทเทียมให้มีจำนวนเซลล์ประสาทในชั้นซ่อนจำนวนต่าง ๆ กัน และปรับเปลี่ยนให้มีจำนวนชั้นซ่อนเป็น 1 ชั้นและ 2 ชั้น อีกทั้งยังได้ปรับชนิดของฟังก์ชันถ่ายโอน คือ ชนิด Log-sigmoid, Tan-sigmoid และ Linear ในชั้นซ่อนและชั้นตัวแปรด้านออกอีกด้วย ซึ่งจากผลการเรียนรู้และการทดสอบโครงข่ายประสาทเทียมแต่ละโครงสร้างที่มีจำนวนเซลล์ประสาทในชั้นซ่อนแตกต่างกันนั้น สามารถแสดงได้ดังตารางที่ 4.1 และ 4.2 ซึ่งจากการดำเนินงานพบว่าฟังก์ชันถ่ายโอนที่เหมาะสมสำหรับนำไปพยากรณ์รูปร่างของรอยเชื่อมในปริภูมิพิกัดนี้ คือ การใช้ฟังก์ชันถ่ายโอนชนิด Log-sigmoid ในชั้นซ่อนและใช้ฟังก์ชันถ่ายโอนชนิด Linear ในชั้นตัวแปรด้านออก ซึ่งสอดคล้องกับทฤษฎีคือ ฟังก์ชันถ่ายโอนทั้งสองชนิดสามารถหาอนุพันธ์ได้และสร้างขอบเขตการตัดสินใจในลักษณะที่โค้งมนได้ จึงสามารถนำไปใช้กับปัญหาที่มีความซับซ้อนและมีลักษณะไม่เชิงเส้นได้ดี สำหรับจำนวนชั้นซ่อนและจำนวนเซลล์ประสาทในชั้นซ่อนที่เหมาะสมคือ โครงข่ายประสาทเทียมที่มีชั้นซ่อน 1 ชั้น และมีเซลล์ประสาทในชั้นซ่อน 6 เซลล์

จากโครงสร้างที่เหมาะสมดังกล่าวคือ โครงข่ายประสาทเทียมที่มี ชั้นซ่อน 1 ชั้น มีจำนวนเซลล์ประสาทในชั้นซ่อน 6 เซลล์ ใช้ฟังก์ชันถ่ายโอนชนิด Log-sigmoid ในชั้นซ่อน และใช้ฟังก์ชันถ่ายโอนชนิด Linear ในชั้นตัวแปรด้านออกนั้น สามารถเรียนรู้ความสัมพันธ์ระหว่างตัวแปรด้านเข้าและตัวแปรด้านออกได้ดี โดยมีค่าความผิดพลาดกำลังสองเฉลี่ยจากการฝึกสอนเท่ากับ 0.000894 และมีค่าความผิดพลาดกำลังสองเฉลี่ยจากการทดสอบเท่ากับ 0.118495

ตารางที่ 4.1 แสดงผลจากการฝึกสอนและทดสอบโครงข่ายประสาทเทียมที่มี 1 ชั้นซ่อน

ลำดับที่	จำนวนเซลล์ประสาทในชั้นซ่อน	ชนิดของฟังก์ชันถ่ายโอนในชั้นซ่อน	ชนิดของฟังก์ชันถ่ายโอนในชั้นตัวแปรค่านอก	MSE ของการฝึกสอน	MSE ของการทดสอบ
1	1	Log-sigmoid	Linear	0.031740	1.452340
2	2	Log-sigmoid	Linear	0.014200	0.442446
3	3	Log-sigmoid	Linear	0.003208	0.208646
4	4	Log-sigmoid	Linear	0.001733	0.216750
5	5	Log-sigmoid	Linear	0.001004	0.389942
6	6	Log-sigmoid	Linear	0.000894	0.118495
7	7	Log-sigmoid	Linear	0.000527	0.187808
8	8	Log-sigmoid	Linear	0.000389	1.013725
9	9	Log-sigmoid	Linear	0.000232	4.222982
10	10	Log-sigmoid	Linear	0.000157	2.585038
11	11	Log-sigmoid	Linear	0.000192	14.793610
12	12	Log-sigmoid	Linear	0.000141	10.468410
13	13	Log-sigmoid	Linear	0.000127	56.631820
14	14	Log-sigmoid	Linear	0.000123	2.240760
15	15	Log-sigmoid	Linear	0.000122	5.061126
16	16	Log-sigmoid	Linear	0.000122	10.097920
17	17	Log-sigmoid	Linear	0.000122	4.827050
18	18	Log-sigmoid	Linear	0.000122	4.362168
19	19	Log-sigmoid	Linear	0.000122	0.663727
20	20	Log-sigmoid	Linear	0.000122	13.781410

ตารางที่ 4.2 แสดงผลจากการฝึกสอนและทดสอบ โครงข่ายประสาทเทียมที่มี 2 ชั้นซ่อน

ลำดับที่	จำนวนเซลล์ประสาทในชั้นซ่อนที่ 1	จำนวนเซลล์ประสาทในชั้นซ่อนที่ 2	ชนิดของฟังก์ชันถ่ายโอนในชั้นซ่อนที่ 1	ชนิดของฟังก์ชันถ่ายโอนในชั้นซ่อนที่ 2	ชนิดของฟังก์ชันถ่ายโอนในชั้นตัวแปรด้านออก	MSE ของการฝึกสอน	MSE ของการทดสอบ
1	2	6	Log-sigmoid	Log-sigmoid	Linear	0.002459	0.295978
2	4	5	Log-sigmoid	Log-sigmoid	Linear	0.000373	0.478663
3	4	6	Log-sigmoid	Log-sigmoid	Linear	0.000325	0.240927
4	6	6	Log-sigmoid	Log-sigmoid	Linear	0.000142	0.352482
5	7	5	Log-sigmoid	Log-sigmoid	Linear	0.000178	1.076084
6	10	4	Log-sigmoid	Log-sigmoid	Linear	0.000162	7.766325
7	10	5	Log-sigmoid	Log-sigmoid	Linear	0.000128	8.913531
8	10	7	Log-sigmoid	Log-sigmoid	Linear	0.000124	0.634268

4.2 การพยากรณ์รูปร่างของรอยเชื่อมโดยใช้การวิเคราะห์ความถดถอยเชิงพหุ

4.2.1 Bead width

ตารางที่ 4.3 Variables Entered/Removed of Bead width

Model	Variables Entered	Variables Removed	Method
1	VAR00001	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	VAR00008	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	VAR00012	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a Dependent Variable: bead

ความหมายของผลลัพธ์ตารางที่ 4.3 Variables Entered/Removed of Bead width

ใน Column ของ Method ระบุว่าเป็นการเลือกตัวแปรอิสระโดยวิธี Stepwise ที่กำหนดเกณฑ์ในการเลือกตัวแปรอิสระเข้าสมการเป็น Use probability of F ที่ให้ค่า Enter <= 0.050 และกำหนดเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการเป็น Use probability of F ที่ให้ค่า Remove = 0.100

Model 1 : หมายถึง ขั้นแรกที่น่าตัวแปรอิสระ VAR00001 เข้าสมการ ทำให้สมการเป็น

$$\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + e \quad (4.1)$$

Model 2 : หมายถึง การนำตัวแปรอิสระ VAR00008 เข้าสมการ ทำให้สมการเป็น

$$\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + \beta_2 \text{VAR00008} + e \quad (4.2)$$

Model 3 : หมายถึง การนำตัวแปรอิสระ VAR00012 เข้าสมการ ทำให้สมการเป็น

$$\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + \beta_2 \text{VAR00008} + \beta_3 \text{VAR00012} + e \quad (4.3)$$

ตารางที่ 4.4 Model Summary (d) of Bead width

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.884(a)	.782	.774	.78762	
2	.958(b)	.919	.912	.48993	
3	.981(c)	.961	.957	.34322	1.952

a Predictors: (Constant), VAR00001

b Predictors: (Constant), VAR00001, VAR00008

c Predictors: (Constant), VAR00001, VAR00008, VAR00012

d Dependent Variable: bead

ความหมายของผลลัพธ์ตารางที่ 4.4 Model Summary (d) of Bead width

Model 1 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 78.2% นั่นคือตัวแปร VAR00001 สามารถอธิบายการเปลี่ยนแปลงของค่า Bead width ได้ 78.2%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.782} = .884$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ VAR00001 = R_{VAR00001}
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{adj} = .774$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{Y}) หรือคือ $SE_{\text{VAR00001}} = .78762$ มิลลิเมตร

Model 2 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 91.9% นั่นคือตัวแปร VAR00001, VAR00008 สามารถอธิบายการเปลี่ยนแปลงของค่า Bead width ได้ 91.9%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.919} = .958$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00001, VAR00008) = $R_{\text{VAR00001, VAR00008}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{adj} = .912$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{Y}) หรือคือ $SE_{\text{VAR00001, VAR00008}} = .48993$ มิลลิเมตร

Model 3 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 96.1% นั่นคือตัวแปร VAR00001, VAR00008, VAR00012 สามารถอธิบายการเปลี่ยนแปลงของค่า Bead width ได้ 96.1%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.961} = .981$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์

ของตัวแปรอิสระ (VAR00001, VAR00008, VAR00012)

$$= R_{\text{VAR00001,VAR00008,VAR00012}}$$

- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{adj} = .957$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ

$$SE_{\text{VAR00001,VAR00008,VAR00012}} = .34322 \text{ มิลลิเมตร}$$

- Durbin – Watson เป็นค่าสถิติที่ใช้ทดสอบสมมติฐาน

H_0 : ค่าคลาดเคลื่อน e_i และ e_j เป็นอิสระกัน

H_1 : ค่าคลาดเคลื่อน e_i และ e_j มีความสัมพันธ์กัน

เนื่องจากค่าสถิติทดสอบ Durbin – Watson = 1.952 ซึ่งมีค่า ≈ 2 หรืออยู่ระหว่าง 1.5 – 2.5 จึงสรุปว่ายอมรับสมมติฐาน H_0 หรือค่าคลาดเคลื่อนเป็นอิสระกัน จึงเป็นไปตามเงื่อนไขของการวิเคราะห์ความถดถอย

ตารางที่ 4.5 ANOVA (d) of Bead width

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	62.182	1	62.182	100.238	.000(a)
	Residual	17.369	28	.620		
	Total	79.551	29			
2	Regression	73.070	2	36.535	152.209	.000(b)
	Residual	6.481	27	.240		
	Total	79.551	29			
3	Regression	76.488	3	25.496	216.430	.000(c)
	Residual	3.063	26	.118		
	Total	79.551	29			

a Predictors: (Constant), VAR00001

b Predictors: (Constant), VAR00001, VAR00008

c Predictors: (Constant), VAR00001, VAR00008, VAR00012

d Dependent Variable: bead

ความหมายของผลลัพธ์ตารางที่ 4.5 ANOVA (d) of Bead width

ตารางที่ 4.5 เป็นตารางการวิเคราะห์ความแปรปรวนแบบทางเดียว

Model 1 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Bead width และ ตัวแปร VAR00001

$$\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + e$$

$$H_0 : \beta_1 = 0 \quad .\text{VS.} \quad H_1 : \beta_1 \neq 0$$

สถิติทดสอบ F = 100.238 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 : นั่นคือ ตัวแปร VAR00001 มีความสัมพันธ์กับตัวแปร Bead width

Model 2 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Bead width กับตัวแปร VAR00001 และตัวแปร VAR00008

$$\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + \beta_2 \text{VAR00008} + e$$

$$H_0 : \beta_1 = \beta_2 = 0 \quad \text{VS.} \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2$$

สถิติทดสอบ F = 152.209 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 : นั่นคือตัวแปร VAR00001 และตัวแปร VAR00008 มีความสัมพันธ์กับตัวแปร Bead width

Model 3 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Bead width กับตัวแปร VAR00001 ตัวแปร VAR00008 และตัวแปร VAR00012

$$\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + \beta_2 \text{VAR00008} + \beta_3 \text{VAR00012} + e$$

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0 \quad \text{VS.} \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2, 3$$

สถิติทดสอบ F = 216.430 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 : นั่นคือตัวแปร VAR00001 ตัวแปร VAR00008 และตัวแปร VAR00012 มีความสัมพันธ์กับตัวแปร Bead width และยอมรับว่า Model 3 มีความเหมาะสม

ตารางที่ 4.6 Coefficient (a) of Bead width

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.839	.541		3.398	.002
	VAR00001	.034	.003	.884	10.012	.000
2	(Constant)	.451	.395		1.143	.263
	VAR00001	.058	.004	1.477	14.237	.000
	VAR00008	-.089	.013	-.699	-6.735	.000
3	(Constant)	1.125	.303		3.707	.001
	VAR00001	.050	.003	1.272	15.524	.000
	VAR00008	-.087	.009	-.682	-9.376	.000
	VAR00012	3.84E-005	.000	.281	5.387	.000

a. Dependent Variable: bead

ความหมายของผลลัพธ์ตารางที่ 4.6 Coefficient (a) of Bead width

ตารางที่ 4.6 เป็นตารางแสดงค่าสัมประสิทธิ์ต่างๆ ของสมการความถดถอยของ Model 1, 2 และ 3

Model 1 สมการจะเป็น $\text{Bead width} = \beta_0 + \beta_1 \text{VAR00001} + e$

หรือ $\text{Bead width} = a + b_1 \text{VAR00001}$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 1.839$ $b_1 = 0.034$
- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.541$, $SE(b_1) = 0.003$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00001 = 0.884$
- t และ Sig. เป็นสถิติทดสอบค่า β_c และ β_2
โดยที่ Sig. ของค่าคงที่ = $0.002 < 0.05$ จึงสรุปว่า $\beta_c = 0$
Sig. ของ $VAR00001 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
สมการจึงกลายเป็น $\widehat{Bead\ width} = 1.839 + 0.034VAR00001$

Model 2

- สมการจะเป็น $Bead\ width = \beta_c + \beta_1 VAR00001 + \beta_2 VAR00008 + e$
หรือ $\widehat{Bead\ width} = a + b_1 VAR00001 + b_2 VAR00008$
- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.451$ $b_1 = 0.058$ $b_2 = -0.089$
 - Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.395$, $SE(b_1) = 0.004$, $SE(b_2) = 0.013$
 - Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00001 = 1.477$,
 $VAR00008 = -0.699$
 - t และ Sig. เป็นสถิติทดสอบค่า β_c , β_1 และ β_2
โดยที่ Sig. ของค่าคงที่ = $0.263 > 0.05$ จึงสรุปว่า $\beta_c = 0$
Sig. ของ $VAR00001 = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
Sig. ของ $VAR00008 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
สมการจึงกลายเป็น $\widehat{Bead\ width} = 0.058VAR00001 - 0.089VAR00008$

Model 3

- สมการจะเป็น $Bead\ width = \beta_c + \beta_1 VAR00001 + \beta_2 VAR00008 + \beta_3 VAR00012 + e$
หรือ $\widehat{Bead\ width} = a + b_1 VAR00001 + b_2 VAR00008 + b_3 VAR00012$
- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 1.125$ $b_1 = 0.050$ $b_2 = -0.087$ $b_3 = 3.84E-005$
 - Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.303$, $SE(b_1) = 0.003$, $SE(b_2) = 0.009$,
 $SE(b_3) = 0.000$
 - Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00001 = 1.272$,
 $VAR00008 = -0.682$, $VAR00012 = 0.281$
 - t และ Sig. เป็นสถิติทดสอบค่า β_c , β_1 , β_2 และ β_3
โดยที่ Sig. ของค่าคงที่ = $0.01 < 0.05$ จึงสรุปว่า $\beta_c = 0$
Sig. ของ $VAR00001 = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
Sig. ของ $VAR00008 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
Sig. ของ $VAR00012 = 0.000 < 0.05$ จึงสรุปว่า $\beta_3 = 0$

สมการจึงกลายเป็น

$$\widehat{Bead\ width} = 1.125 + 0.050VAR00001 - 0.087VAR00008 + 3.84E-005VAR00012$$

4.2.2 Penetration

ตารางที่ 4.7 Variables Entered/Removed (b) of Penetration

Model	Variables Entered	Variables Removed	Method
1	VAR00005, VAR00004, VAR00003, VAR00001(a)	.	Enter

a All requested variables entered.

b Dependent Variable: penetration

ความหมายของผลลัพธ์ตารางที่ 4.7 Variables Entered/Removed (b) of Penetration

ใน Column ของ Method ระบุว่าเป็นการเลือกตัวแปรอิสระ โดยวิธี Enter โดยตัวแปร Penetration เป็นตัวแปรตาม คำว่า Variables Entered หมายถึง ตัวแปรอิสระที่นำเข้ามาสมการความถดถอยเชิงเส้น คือ VAR00005, VAR00004, VAR00003, VAR00001

ตารางที่ 4.8 Model Summary (b) of Penetration

Model	R	R Square	Adjusted R Square	Std Error of the Estimate	Durbin-Watson
1	.927(a)	.859	.835	25521	2.311

a Predictors: (Constant), VAR00005, VAR00004, VAR00003, VAR00001

b Dependent Variable: penetration

ความหมายของผลลัพธ์ตารางที่ 4.8 Model Summary (b) of Penetration

Model 1 .

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 85.9% นั่นคือตัวแปร VAR00005, VAR00004, VAR00003, VAR00001 สามารถอธิบายการเปลี่ยนแปลงของค่า Penetration ได้ 85.9%
- R หมายถึง $\sqrt{R^2 \text{ Square}} = \sqrt{.859} = .927$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00005, VAR00004, VAR00003, VAR00001) = $R_{\text{VAR00005, VAR00004, VAR00003, VAR00001}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .835$
- Std Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (y) หรือคือ $SE_{\text{VAR00005, VAR00004, VAR00003, VAR00001}} = .25521$ มิลลิเมตร

- Durbin – Watson

เป็นค่าสถิติที่ใช้ทดสอบสมมติฐาน

H_0 : ค่าคลาดเคลื่อน e_i และ e_j เป็นอิสระกัน

H_1 : ค่าคลาดเคลื่อน e_i และ e_j มีความสัมพันธ์กัน

เนื่องจากค่าสถิติทดสอบ Durbin – Watson = 2.311 ซึ่งมีค่า ≈ 2 หรืออยู่ระหว่าง 1.5 – 2.5 จึงสรุปว่ายอมรับสมมติฐาน H_0 หรือค่าคลาดเคลื่อนเป็นอิสระกัน จึงเป็นไปตามเงื่อนไขของการวิเคราะห์ความถดถอย

ตารางที่ 4.9 ANOVA (b) of Penetration

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.513	4	2.378	36.515	.000(a)
	Residual	1.563	24	.065		
	Total	11.076	28			

a Predictors: (Constant), VAR00005, VAR00004, VAR00003, VAR00001

b Dependent Variable: penetration

ความหมายของผลลัพธ์ตารางที่ 4.9 ANOVA (b) of Penetration

ตารางที่ 4.9 เป็นตารางการวิเคราะห์ความแปรปรวนแบบทางเดียว ซึ่งใช้ในการทดสอบสมมติฐาน

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$.VS $H_1 : \beta_i \neq 0$ อย่างน้อย 1 ค่า; $i = 1, 2, 3, 4$

Model 1 : SSR = 9.513 SSE = 1.563

$$MSR = \frac{9.513}{4} = 2.378 \quad ; \quad MSE = \frac{1.563}{24} = 0.065$$

สถิติทดสอบ $F = 36.515$ และ Sig. of $F = .000 < 0.05$ จึงปฏิเสธ H_0 นั่นคือ มีตัวแปรอิสระอย่างน้อย 1 ตัวที่มีความสัมพันธ์กับตัวแปรตาม Penetration อย่างมีนัยสำคัญ

ตารางที่ 4.10 Coefficient (a) of Penetration

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-2.437	.608		-4.005	.001
	VAR00001	.163	.022	11.183	7.368	.000
	VAR00003	-.383	.075	-.720	-5.102	.000
	VAR00004	-.060	.011	-.498	-5.341	.000
	VAR00005	-.013	.002	-10.554	-6.821	.000

a. Dependent Variable: penetration

ความหมายของผลลัพธ์ตารางที่ 4.10 Coefficient(a) of Penetration

ตารางที่ 4.10 เป็นตารางแสดงค่าสัมประสิทธิ์ต่างๆ ของสมการความถดถอย

Model 1 : สมการจะเป็น Penetration = $\beta_0 + \beta_1 \text{VAR00001} + \beta_2 \text{VAR00003} + \beta_3 \text{VAR00004} + \beta_4 \text{VAR00005} + e$

หรือ Penetration = $a + b_1 \text{VAR00001} + b_2 \text{VAR00003} + b_3 \text{VAR00004} + b_4 \text{VAR00005}$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = -2.437$ $b_1 = 0.163$ $b_2 = -0.383$ $b_3 = -0.060$ $b_4 = -0.013$
- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.608$, $SE(b_1) = 0.022$, $SE(b_2) = 0.075$, $SE(b_3) = 0.011$, $SE(b_4) = 0.002$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $\text{VAR00001} = 11.183$, $\text{VAR00003} = -0.720$, $\text{VAR00004} = -0.498$, $\text{VAR00005} = -10.554$
- t และ Sig. เป็นสถิติทดสอบค่า $\beta_0, \beta_1, \beta_2, \beta_3$ และ β_4 และ โดยที่ Sig. ของค่าคงที่ = $0.001 < 0.05$ จึงสรุปว่า $\beta_0 = 0$
Sig ของ $\text{VAR00001} = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
Sig ของ $\text{VAR00003} = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
Sig. ของ $\text{VAR00004} = 0.000 < 0.05$ จึงสรุปว่า $\beta_3 = 0$
Sig. ของ $\text{VAR00005} = 0.000 < 0.05$ จึงสรุปว่า $\beta_4 = 0$

สมการจึงกลายเป็น Penetration = $-2.437 + 0.163\text{VAR00001} - 0.383\text{VAR00003} - 0.060\text{VAR00004} - 0.013\text{VAR00005}$

4.2.3 Reinforcement height

ตารางที่ 4 11 Variables Entered/Removed (a) of Reinforcement height

Model	Variables Entered	Variables Removed	Method
1	VAR00003	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	VAR00004	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100)
3	VAR00007	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	VAR00008	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: reinforce

ความหมายของผลลัพธ์ตารางที่ 4.11 Variables Entered/ Removed (a) of Reinforcement height

ใน Column ของ Method ระบุว่าเป็นการเลือกตัวแปรอิสระ โดยวิธี Stepwise ที่กำหนดเกณฑ์ในการเลือกตัวแปรอิสระเข้าสมการเป็น Use probability of F ที่ให้ค่า Enter <= 0.050 และกำหนดเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการเป็น Use probability of F ที่ให้ค่า Remove = 0.100

Model 1 : หมายถึง ขั้นแรกก็นำตัวแปรอิสระ VAR00003 เข้าสมการ ทำให้สมการเป็น

$$\text{Reinforcement} = \beta_0 + \beta_1 \text{VAR00003} + e \quad (4.4)$$

Model 2 : หมายถึง การนำตัวแปรอิสระ VAR00004 เข้าสมการ ทำให้สมการเป็น

$$\text{Reinforcement} = \beta_0 + \beta_1 \text{VAR00003} + \beta_2 \text{VAR00004} + e \quad (4.5)$$

Model 3 : หมายถึง การนำตัวแปรอิสระ VAR00007 เข้าสมการ ทำให้สมการเป็น

$$\text{Reinforcement} = \beta_0 + \beta_1 \text{VAR00003} + \beta_2 \text{VAR00004} + \beta_3 \text{VAR00007} + e \quad (4.6)$$

Model 4 : หมายถึง การนำตัวแปรอิสระ VAR00008 เข้าสมการ ทำให้สมการเป็น

$$\text{Reinforcement} = \beta_0 + \beta_1 \text{VAR00003} + \beta_2 \text{VAR00004} + \beta_3 \text{VAR00007} + \beta_4 \text{VAR00008} + e \quad (4.7)$$

ตารางที่ 4.12 Model Summary (e) of Reinforcement height

Model	R	R Square	Adjusted R Square	Std Error of the Estimate	Durbin-Watson
1	.476(a)	.227	.199	.37241	
2	.805(b)	.647	.621	.25611	
3	.950(c)	.903	.892	.13678	
4	.979(d)	.958	.951	.09207	1.868

a Predictors (Constant), VAR00003

b Predictors: (Constant), VAR00003, VAR00004

c Predictors: (Constant), VAR00003, VAR00004, VAR00007

d Predictors: (Constant), VAR00003, VAR00004, VAR00007, VAR00008

e Dependent Variable: reinforce

ความหมายของผลลัพธ์ตารางที่ 4.12 Model Summary (e) of Reinforcement height

Model 1

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 22.7% นั่นคือตัวแปร VAR00003 สามารถอธิบายการเปลี่ยนแปลงของค่า Reinforcement ได้ 22.7%
- R หมายถึง $\sqrt{R^2\text{-Square}} = \sqrt{.227} = .476$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ VAR00003 = R_{VAR00003}
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .199$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00003}} = .37241$ มิลลิเมตร

Model 2 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 64.7% นั่นคือตัวแปร VAR00003, VAR00004 สามารถอธิบายการเปลี่ยนแปลงของค่า Reinforcement ได้ 64.7%
- R หมายถึง $\sqrt{R^2\text{-Square}} = \sqrt{.647} = .805$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00003, VAR00004) = $R_{\text{VAR00003,VAR00004}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .621$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00003,VAR00004}} = .25611$ มิลลิเมตร

Model 3 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 90.3% นั่นคือตัวแปร VAR00003, VAR00004, VAR00007 สามารถอธิบายการเปลี่ยนแปลงของค่า

Reinforcement ได้ 90.3%

- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.903} = .950$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00003, VAR00004, VAR00007)
 $= R_{\text{VAR00003,VAR00004,VAR00007}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .892$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00003,VAR00004,VAR00007}} = .13678$ มิลลิเมตร

Model 4 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 95.8% นั่นคือตัวแปร VAR00003, VAR00004, VAR00007, VAR00008 สามารถอธิบายการเปลี่ยนแปลงของค่า Reinforcement ได้ 95.8%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.958} = .979$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00003, VAR00004, VAR00007, VAR00008)
 $= R_{\text{VAR00003,VAR00004,VAR00007,VAR00008}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .951$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00003,VAR00004,VAR00007,VAR00008}} = .09207$ มิลลิเมตร
- Durbin – Watson เป็นค่าสถิติที่ใช้ทดสอบสมมติฐาน
 H_0 : ค่าคลาดเคลื่อน e_t และ e_{t-1} เป็นอิสระกัน
 H_1 : ค่าคลาดเคลื่อน e_t และ e_{t-1} มีความสัมพันธ์กัน
เนื่องจากค่าสถิติทดสอบ Durbin – Watson = 1.868 ซึ่งมีค่า ≈ 2 หรืออยู่ระหว่าง 1.5 – 2.5 จึงสรุปว่ายอมรับสมมติฐาน H_0 หรือค่าคลาดเคลื่อนเป็นอิสระกัน จึงเป็นไปตามเงื่อนไขของการวิเคราะห์ความถดถอย

ตารางที่ 4.13 ANOVA (c) of Reinforcement height

Model		Sum of Squares	df	Mean Square	F	Sig
1	Regression	1.138	1	1.138	8.205	.008(a)
	Residual	3.883	28	.139		
	Total	5.021	29			
2	Regression	3.250	2	1.625	24.776	.000(b)
	Residual	1.771	27	.066		
	Total	5.021	29			
3	Regression	4.535	3	1.512	80.798	.000(c)
	Residual	.486	26	.019		
	Total	5.021	29			
4	Regression	4.809	4	1.202	141.833	.000(d)
	Residual	.212	25	.008		
	Total	5.021	29			

a Predictors: (Constant), VAR00003

b Predictors: (Constant), VAR00003, VAR00004

c Predictors: (Constant), VAR00003, VAR00004, VAR00007

d Predictors: (Constant), VAR00003, VAR00004, VAR00007, VAR00008

e Dependent Variable reinforce

ความหมายของผลลัพธ์ตารางที่ 4.13 ANOVA (c) of Reinforcement height

ตารางที่ 4.13 เป็นตารางการวิเคราะห์ความแปรปรวนแบบทางเดียว

Model 1 ใช้ทดสอบความสัมพันธ์ของตัวแปร Reinforcement และ ตัวแปร VAR00003

$$\text{Reinforcement} = \beta_0 + \beta_1 \text{VAR00003} + e$$

$$H_0 : \beta_1 = 0 \quad \text{VS.} \quad H_1 : \beta_1 \neq 0$$

สถิติทดสอบ $F = 8.205$ และ $\text{Sig. of } F = .008 < 0.05$ จึงปฏิเสธ H_0 นั่นคือ ตัวแปร VAR00003 มีความสัมพันธ์กับตัวแปร Reinforcement

Model 2 ใช้ทดสอบความสัมพันธ์ของตัวแปร Reinforcement กับตัวแปร VAR00003 และตัวแปร VAR00004

$$\text{Reinforcement} = \beta_0 + \beta_1 \text{VAR00003} + \beta_2 \text{VAR00004} + e$$

$$H_0 : \beta_1 = \beta_2 = 0 \quad \text{VS} \quad H_1 : \beta_i \neq 0 \text{ อย่างน้อย 1 ค่า, } i = 1, 2$$

สถิติทดสอบ $F = 24.776$ และ $\text{Sig. of } F = .000 < 0.05$ จึงปฏิเสธ H_0 นั่นคือตัวแปร VAR00003 และตัวแปร VAR00004 มีความสัมพันธ์กับตัวแปร Reinforcement

Model 3 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Reinforcement กับตัวแปร VAR00003 ตัวแปร VAR00004 และตัวแปร VAR00007

$$\text{Reinforcement} = \beta_1 + \beta_2 \text{VAR00003} + \beta_3 \text{VAR00004} + \beta_4 \text{VAR00007} + e$$

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0 \quad .\text{VS.} \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2, 3$$

สถิติทดสอบ F = 80.798 และ Sig of F = .000 < 0.05 จึงปฏิเสธ H_0 : นั่นคือตัวแปร VAR00003 ตัวแปร VAR00004 และตัวแปร VAR00007 มีความสัมพันธ์กับตัวแปร Reinforcement

Model 4 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Reinforcement กับตัวแปร VAR00003 ตัวแปร VAR00004 ตัวแปร VAR00007 และตัวแปร VAR00008

$$\text{Reinforcement} = \beta_1 + \beta_2 \text{VAR00003} + \beta_3 \text{VAR00004} + \beta_4 \text{VAR00007} + \beta_5 \text{VAR00008} + e$$

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0 \quad .\text{VS.} \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2, 3, 4$$

สถิติทดสอบ F = 141.833 และ Sig of F = .000 < 0.05 จึงปฏิเสธ H_0 : นั่นคือตัวแปร VAR00003 ตัวแปร VAR00004 ตัวแปร VAR00007 และตัวแปร VAR00008 มีความสัมพันธ์กับตัวแปร Reinforcement และยอมรับว่า Model 4 มีความเหมาะสม

ตารางที่ 4.14 Coefficient(a) of Reinforcement height

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std Error	Beta		
1	(Constant)	1.225	.155		7.885	.000
	VAR00003	-.169	.059	-.476	-2.864	.008
2	(Constant)	1.042	.112		9.339	.000
	VAR00003	-.267	.044	-.755	-6.070	.000
	VAR00004	.056	.010	.706	5.675	.000
3	(Constant)	.488	.090		5.448	.000
	VAR00003	-.121	.029	-.342	-4.119	.000
	VAR00004	.240	.023	.3038	10.506	.000
	VAR00007	-.001	.000	-.2574	-8.286	.000
4	(Constant)	.697	.071		9.873	.000
	VAR00003	-1.720	.282	-.4860	-6.106	.000
	VAR00004	.294	.018	.3717	16.281	.000
	VAR00007	-.001	.000	-.3443	-13.298	.000
	VAR00008	.151	.027	.4710	5.690	.000

a Dependent Variable. reinforce

ความหมายของผลลัพธ์ตารางที่ 4.14 Coefficient (a) of Reinforcement height

ตารางที่ 4.14 เป็นตารางแสดงค่าสัมประสิทธิ์ต่างๆ ของสมการความถดถอยของ Model 1, 2, 3 และ 4

Model 1 : สมการจะเป็น $Reinforcement = \beta_0 + \beta_1 VAR00003 + e$

หรือ $\widehat{Reinforcement} = a + b_1 VAR00003$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 1.225$ $b_1 = -0.169$
- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.155$, $SE(b_1) = 0.059$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00003 = -0.476$
- t และ Sig. เป็นสถิติทดสอบค่า β_0 และ β_1

โดยที่ Sig. ของค่าคงที่ = $0.000 < 0.05$ จึงสรุปว่า $\beta_0 = 0$

Sig. ของ $VAR00003 = 0.008 < 0.05$ จึงสรุปว่า $\beta_1 = 0$

สมการจึงกลายเป็น $\widehat{Reinforcement} = 1.225 - 0.169VAR00003$

Model 2 : สมการจะเป็น $Reinforcement = \beta_0 + \beta_1 VAR00003 + \beta_2 VAR00004 + e$

หรือ $\widehat{Reinforcement} = a + b_1 VAR00003 + b_2 VAR00004$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 1.042$ $b_1 = -0.267$ $b_2 = 0.056$
- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.112$, $SE(b_1) = 0.044$, $SE(b_2) = 0.010$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00003 = -0.755$,

$VAR00004 = 0.706$

- t และ Sig. เป็นสถิติทดสอบค่า β_0 , β_1 และ β_2

โดยที่ Sig. ของค่าคงที่ = $0.000 > 0.05$ จึงสรุปว่า $\beta_0 = 0$

Sig. ของ $VAR00003 = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$

Sig. ของ $VAR00004 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$

สมการจึงกลายเป็น $\widehat{Reinforcement} = 1.042 - 0.267VAR00003 - 0.056VAR00004$

Model 3 : สมการจะเป็น $Reinforcement = \beta_0 + \beta_1 VAR00003 + \beta_2 VAR00004 + \beta_3 VAR00007 + e$

หรือ $\widehat{Reinforcement} = a + b_1 VAR00003 + b_2 VAR00004 + b_3 VAR00007$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.488$ $b_1 = -0.121$ $b_2 = 0.240$ $b_3 = -0.001$
- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.090$, $SE(b_1) = 0.029$, $SE(b_2) = 0.023$,

$SE(b_3) = 0.000$

- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00003 = -0.342$,

$VAR00004 = 0.3038$, $VAR00007 = -2.574$

- t และ Sig. เป็นสถิติทดสอบค่า β_0 , β_1 , β_2 และ β_3

โดยที่ Sig. ของค่าคงที่ = $0.000 < 0.05$ จึงสรุปว่า $\beta_0 = 0$

Sig. ของ $VAR00003 = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$

Sig. ของ $VAR00004 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$

Sig. ของ $VAR00007 = 0.000 < 0.05$ จึงสรุปว่า $\beta_3 = 0$

สมการจึงกลายเป็น $\widehat{Reinforcement} = 0.488 - 0.121VAR00003 + 0.240VAR00004$

$- 0.001VAR00007$

Model 4 : สมการจะเป็น Reinforcement = $\beta_0 + \beta_1 \text{VAR00003} + \beta_2 \text{VAR00004} + \beta_3 \text{VAR00007} + \beta_4 \text{VAR00008} + e$

หรือ Reinforcement = $a + b_1 \text{VAR00003} + b_2 \text{VAR00004} + b_3 \text{VAR00007} + b_4 \text{VAR00008}$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.697$ $b_1 = -1.720$ $b_2 = 0.294$ $b_3 = -0.001$ $b_4 = -0.151$
- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.071$, $SE(b_1) = 0.282$, $SE(b_2) = 0.018$, $SE(b_3) = 0.000$, $SE(b_4) = 0.027$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $\text{VAR00003} = -4.860$, $\text{VAR00004} = 3.717$, $\text{VAR00007} = -3.443$, $\text{VAR00008} = 4.710$
- t และ Sig. เป็นสถิติทดสอบค่า $\beta_0, \beta_1, \beta_2, \beta_3$ และ β_4
 โดยที่ Sig. ของค่าคงที่ = $0.000 < 0.05$ จึงสรุปว่า $\beta_0 = 0$
 Sig. ของ $\text{VAR00004} = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
 Sig. ของ $\text{VAR00003} = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
 Sig. ของ $\text{VAR00009} = 0.004 < 0.05$ จึงสรุปว่า $\beta_3 = 0$
 Sig. ของ $\text{VAR00006} = 0.033 < 0.05$ จึงสรุปว่า $\beta_4 = 0$
 สมการจึงกลายเป็น Reinforcement = $0.697 - 1.720\text{VAR00003} + 0.294\text{VAR00004} - 0.001\text{VAR00007} - 0.001\text{VAR00008}$

4.2.4 Cross section

ตารางที่ 4.15 Variables Entered/Removed (a) of Cross section

Model	Variables Entered	Variables Removed	Method
1	VAR00004	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= 100).
2	VAR00003	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	VAR00009	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	VAR00006	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= 100).

a Dependent Variable: logcrosssec

ความหมายของผลลัพธ์ตารางที่ 4.15 Variables Entered/ Removed (a) of Cross section

ใน Column ของ Method ระบุว่าเป็นการเลือกตัวแปรอิสระโดยวิธี Stepwise ที่กำหนดเกณฑ์ในการเลือกตัวแปรอิสระเข้าสมการเป็น Use probability of F ที่ให้ค่า Enter ≤ 0.050 และกำหนดเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการเป็น Use probability of F ที่ให้ค่า Remove = 0.100

Model 1 : หมายถึง ขั้นแรกก็นำตัวแปรอิสระ VAR00004 เข้าสมการ ทำให้สมการเป็น

$$\text{Log(Cross section)} = \beta_0 + \beta_1 \text{VAR00004} + e \quad (4.8)$$

Model 2 : หมายถึง การนำตัวแปรอิสระ VAR00003 เข้าสมการ ทำให้สมการเป็น

$$\text{Log(Cross section)} = \beta_0 + \beta_1 \text{VAR00004} + \beta_2 \text{VAR00003} + e \quad (4.9)$$

Model 3 หมายถึง การนำตัวแปรอิสระ VAR00009 เข้าสมการ ทำให้สมการเป็น

$$\text{Log(Cross section)} = \beta_0 + \beta_1 \text{VAR00004} + \beta_2 \text{VAR00003} + \beta_3 \text{VAR00009} + e \quad (4.10)$$

Model 4 : หมายถึง การนำตัวแปรอิสระ VAR00006 เข้าสมการ ทำให้สมการเป็น

$$\text{Log(Cross section)} = \beta_0 + \beta_1 \text{VAR00004} + \beta_2 \text{VAR00003} + \beta_3 \text{VAR00009} + \beta_4 \text{VAR00006} + e \quad (4.11)$$

ตารางที่ 4.16 Model Summary (e) of Cross section

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.675(a)	.456	.436	.18419	
2	.963(b)	.928	.922	.06837	
3	.969(c)	.939	.931	.06423	
4	.974(d)	.949	.941	.05970	2.148

a Predictors (Constant), VAR00004

b Predictors (Constant), VAR00004, VAR00003

c Predictors (Constant), VAR00004, VAR00003, VAR00009

d Predictors (Constant), VAR00004, VAR00003, VAR00009, VAR00006

e Dependent Variable: logcrosssec

ความหมายของผลลัพธ์ตารางที่ 16 Model Summary(e) of Cross section

Model 1 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 45.6% นั่นคือตัวแปร VAR00004 สามารถอธิบายการเปลี่ยนแปลงของค่า Log(Cross section) ได้ 45.6%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.456} = .675$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ VAR00004 = R_{VAR00004}
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .436$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00003}} = .18419$ ตารางมิลลิเมตร

Model 2 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 92.8% นั่นคือตัวแปร VAR00004, VAR00003สามารถอธิบายการเปลี่ยนแปลงของค่า Log(Cross section) ได้ 92.8%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.928} = .963$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00004, VAR00003) = $R_{\text{VAR00004,VAR00003}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .922$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00004,VAR00003}} = .06837$ ตารางมิลลิเมตร

Model 3 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 93.9% นั่นคือตัวแปร VAR00004, VAR00003,VAR00009 สามารถอธิบายการเปลี่ยนแปลงของค่า Log(Cross section) ได้ 93.9%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.939} = .969$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00004, VAR00003, VAR00009) = $R_{\text{VAR00004,VAR00003,VAR00009}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .931$
- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{\text{VAR00004,VAR00003,VAR00009}} = .06423$ ตารางมิลลิเมตร

Model 4 :

- R Square หมายถึง สัมประสิทธิ์การตัดสินใจ = 94.9% นั่นคือตัวแปร VAR00004, VAR00003,VAR00009, VAR00006 สามารถอธิบายการเปลี่ยนแปลงของค่า Log(Cross section) ได้ 94.9%
- R หมายถึง $\sqrt{R^2 \text{Square}} = \sqrt{.949} = .974$ ซึ่งเป็นค่าที่แสดงความสัมพันธ์ของตัวแปรอิสระ (VAR00004, VAR00003, VAR00009, VAR00006) = $R_{\text{VAR00004,VAR00003,VAR00009,VAR00006}}$
- Adjusted R Square หมายถึง เป็นค่า R^2 ที่ได้ปรับแล้ว ในที่นี้ $R^2_{\text{adj}} = .941$

- Std. Error of the estimate หมายถึง ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า (\hat{y}) หรือ คือ $SE_{VAR00004, VAR00001, VAR00009, VAR00006} = .05970$ ตารางมิลลิเมตร

- Durbin – Watson เป็นค่าสถิติที่ใช้ทดสอบสมมติฐาน

H_0 : ค่าคลาดเคลื่อน e_t และ e_{t-1} เป็นอิสระกัน

H_1 : ค่าคลาดเคลื่อน e_t และ e_{t-1} มีความสัมพันธ์กัน

เนื่องจากค่าสถิติทดสอบ Durbin – Watson = 2.148 ซึ่งมีค่า ≈ 2 หรืออยู่ระหว่าง 1.5 – 2.5 จึงสรุปว่ายอมรับสมมติฐาน H_0 หรือค่าคลาดเคลื่อนเป็นอิสระกัน จึงเป็นไปตามเงื่อนไขของการวิเคราะห์ความถดถอย

ตารางที่ 4 17 ANOVA(e) of Cross section

Model		Sum of Squares	df	Mean Square	F	Sig
1	Regression	.795	1	.795	23.442	.000(a)
	Residual	.950	28	.034		
	Total	1.745	29			
2	Regression	1.619	2	.810	173.210	.000(b)
	Residual	.126	27	.005		
	Total	1.745	29			
3	Regression	1.638	3	.546	132.345	.000(c)
	Residual	.107	26	.004		
	Total	1.745	29			
4	Regression	1.656	4	.414	116.163	.000(d)
	Residual	.089	25	.004		
	Total	1.745	29			

a Predictors. (Constant), VAR00004

b Predictors: (Constant), VAR00004, VAR00003

c Predictors: (Constant), VAR00004, VAR00003, VAR00009

d Predictors. (Constant), VAR00004, VAR00003, VAR00009, VAR00006

e Dependent Variable logcrosssec

ความหมายของผลลัพธ์ตารางที่ 4.16 ANOVA (e) of Cross section

ตารางที่ 4.16 เป็นตารางการวิเคราะห์ความแปรปรวนแบบทางเดียว

Model 1 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Log(Cross section) และ ตัวแปร VAR00004

$$\text{Log(Cross section)} = \beta_0 + \beta_1 \text{VAR00004} + e$$

$$H_0 : \beta_2 = 0 \quad .VS. \quad H_1 : \beta_2 = 0$$

สถิติทดสอบ F = 23.442 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 นั่นคือ ตัวแปร VAR00004 มีความสัมพันธ์กับตัวแปร Log(Cross section)

Model 2 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Log(Cross section) กับตัวแปร VAR00004 และตัวแปร VAR00003

$$\text{Log(Cross section)} = \beta_c + \beta_2 \text{VAR00004} + \beta_3 \text{VAR00003} + e$$

$$H_0 : \beta_2 = \beta_3 = 0 \quad .VS. \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2$$

สถิติทดสอบ F = 173.210 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 นั่นคือตัวแปร VAR00004 และตัวแปร VAR00003 มีความสัมพันธ์กับตัวแปร Log(Cross section)

Model 3 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Log(Cross section) กับตัวแปร VAR00004 ตัวแปร VAR00003 และตัวแปร VAR00009

$$\text{Log(Cross section)} = \beta_c + \beta_2 \text{VAR00004} + \beta_3 \text{VAR00003} + \beta_4 \text{VAR00009} + e$$

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0 \quad VS \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2, 3$$

สถิติทดสอบ F = 132.345 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 นั่นคือตัวแปร VAR00004 ตัวแปร VAR00003 และตัวแปร VAR00009 มีความสัมพันธ์กับตัวแปร Log(Cross section)

Model 4 : ใช้ทดสอบความสัมพันธ์ของตัวแปร Log(Cross section) กับตัวแปร VAR00004 ตัวแปร VAR00003 ตัวแปร VAR00009 และตัวแปร VAR00006

$$\text{Log(Cross section)} = \beta_c + \beta_2 \text{VAR00004} + \beta_3 \text{VAR00003} + \beta_4 \text{VAR00009} + \beta_5 \text{VAR00006} + e$$

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0 \quad .VS. \quad H_1 : \beta_i = 0 \text{ อย่างน้อย 1 ค่า ; } i = 1, 2, 3, 4$$

สถิติทดสอบ F = 116.163 และ Sig. of F = .000 < 0.05 จึงปฏิเสธ H_0 นั่นคือตัวแปร VAR00004 ตัวแปร VAR00003 ตัวแปร VAR00009 และตัวแปร VAR00006 มีความสัมพันธ์กับตัวแปร Log(Cross section) และยอมรับว่า Model 4 มีความเหมาะสม

ตารางที่ 4.18 Coefficient (a) of Cross section

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	295	.059		4.998	.000
	VAR00004	.031	.006	.675	4.842	.000
2	(Constant)	.563	.030		18.898	.000
	VAR00004	.045	.003	.971	17.231	.000
	VAR00003	-.156	.012	-.748	-13.276	.000
3	(Constant)	.494	.042		11.650	.000
	VAR00004	.123	.037	2.648	3.374	.002
	VAR00003	-.139	.014	-.664	-10.108	.000
	VAR00009	-.007	.003	-1.715	-2.142	.042
4	(Constant)	.550	.046		11.841	.000
	VAR00004	.190	.045	4.088	4.218	.000
	VAR00003	-.283	.065	-1.354	-4.345	.000
	VAR00009	-.014	.004	-3.237	-3.223	.004
	VAR00006	.001	.000	.796	2.257	.033

a Dependent Variable: logcrosssec

ความหมายของผลลัพธ์ตารางที่ 4.16 Coefficient (a) of Cross section

ตารางที่ 4.16 เป็นตารางแสดงค่าสัมประสิทธิ์ต่างๆ ของสมการความถดถอยของ Model 1, 2, 3 และ 4

Model 1 . สมการจะเป็น $\text{Log}(\text{Cross section}) = \beta_0 + \beta_1 \text{VAR00004} + e$

หรือ $\text{Log}(\text{Cross section}) = a + b_1 \text{VAR00004}$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.295$ $b_1 = 0.031$
- Std Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.059$, $SE(b_1) = 0.006$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $\text{VAR00004} = 0.675$
- และ Sig. เป็นสถิติทดสอบค่า t_0 และ t_1

โดยที่ Sig. ของค่าคงที่ $= 0.000 < 0.05$ จึงสรุปว่า $t_0 = 0$

Sig ของ $\text{VAR00004} = 0.000 < 0.05$ จึงสรุปว่า $t_1 = 0$

สมการจึงกลายเป็น $\text{Log}(\text{Cross section}) = 0.295 - 0.031 \text{VAR00004}$

Model 2 : สมการจะเป็น $\text{Log}(\text{Cross section}) = \beta_0 + \beta_1 \text{VAR00004} + \beta_2 \text{VAR00003} + e$

หรือ $\text{Log}(\text{Cross section}) = a + b_1 \text{VAR00004} + b_2 \text{VAR00003}$

- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.563$ $b_1 = 0.045$ $b_2 = -0.156$

- Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.030$, $SE(b_1) = 0.003$, $SE(b_2) = 0.012$
- Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00004 = -0.971$,
 $VAR00003 = -0.748$
- t และ Sig. เป็นสถิติทดสอบค่า β_0 , β_1 และ β_2
โดยที่ Sig. ของค่าคงที่ = 0.000 < 0.05 จึงสรุปว่า $\beta_0 = 0$
Sig. ของ $VAR00003 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
Sig. ของ $VAR00004 = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
สมการจึงกลายเป็น $Log(\widehat{Cross\ section}) = 0.563 + 0.045VAR00004 - 0.156VAR00003$

- Model 3 : สมการจะเป็น $Log(Cross\ section) = \beta_0 + \beta_1 VAR00004 + \beta_2 VAR00003 + \beta_3 VAR00009 + e$
หรือ $= a + b_1 VAR00004 + b_2 VAR00003 + b_3 VAR00009$
- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.494$ $b_1 = 0.123$ $b_2 = -0.139$ $b_3 = -0.007$
 - Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.042$, $SE(b_1) = 0.037$, $SE(b_2) = 0.014$,
 $SE(b_3) = 0.003$
 - Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00004 = 2.648$,
 $VAR00003 = -0.664$, $VAR00009 = -1.715$
 - t และ Sig. เป็นสถิติทดสอบค่า β_0 , β_1 , β_2 และ β_3
โดยที่ Sig. ของค่าคงที่ = 0.000 < 0.05 จึงสรุปว่า $\beta_0 = 0$
Sig. ของ $VAR00004 = 0.002 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
Sig. ของ $VAR00003 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$
Sig. ของ $VAR00009 = 0.042 < 0.05$ จึงสรุปว่า $\beta_3 = 0$
สมการจึงกลายเป็น $Log(\widehat{Cross\ section}) = 0.494 + 0.123VAR00004 - 0.139VAR00003$
 $- 0.007VAR00009$

- Model 4 : สมการจะเป็น $Log(Cross\ section) = \beta_0 + \beta_1 VAR00004 + \beta_2 VAR00003 + \beta_3 VAR00009$
 $+ \beta_4 VAR00006 + e$
หรือ $Log(\widehat{Cross\ section}) = a + b_1 VAR00004 + b_2 VAR00003 + b_3 VAR00009$
 $+ b_4 VAR00006$
- B เป็นค่าสัมประสิทธิ์ถดถอยที่มี $a = 0.550$ $b_1 = 0.190$ $b_2 = -0.283$ $b_3 = -0.014$ $b_4 = 0.001$
 - Std. Error หมายถึง ค่าคลาดเคลื่อนมาตรฐาน นั่นคือ $SE(a) = 0.046$, $SE(b_1) = 0.045$, $SE(b_2) = 0.065$,
 $SE(b_3) = 0.004$, $SE(b_4) = 0.000$
 - Beta หมายถึง ค่าสัมประสิทธิ์ความถดถอยมาตรฐานของ $VAR00004 = 4.088$,
 $VAR00003 = -1.354$, $VAR00009 = -3.237$, $VAR00006 = 0.796$
 - t และ Sig. เป็นสถิติทดสอบค่า β_0 , β_1 , β_2 , β_3 และ β_4
โดยที่ Sig. ของค่าคงที่ = 0.000 < 0.05 จึงสรุปว่า $\beta_0 = 0$
Sig. ของ $VAR00004 = 0.000 < 0.05$ จึงสรุปว่า $\beta_1 = 0$
Sig. ของ $VAR00003 = 0.000 < 0.05$ จึงสรุปว่า $\beta_2 = 0$

Sig. ของ VAR00009 = 0.004 < 0.05 จึงสรุปว่า $\beta_3 = 0$

Sig. ของ VAR00006 = 0.033 < 0.05 จึงสรุปว่า $\beta_4 = 0$

$$\text{สมการจึงกลายเป็น } \text{Log}(\text{Cross section}) = 0.550 + 0.190\text{VAR00004} - 0.283\text{VAR00003} \\ - 0.014\text{VAR00009} + 0.001\text{VAR00006}$$

4.3 สรุปผลการพยากรณ์รูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ

ผลการพยากรณ์รูปร่างของรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม และการวิเคราะห์ความถดถอยเชิงพหุ สามารถแสดงได้ดังตารางที่ 4.19, 4.20 และ 4.21 ซึ่งจากตารางพบว่าวิธีการพยากรณ์โดยใช้โครงข่ายประสาทเทียมให้ค่าพยากรณ์ที่มีความถูกต้องสูงกว่าวิธีการวิเคราะห์ความถดถอยเชิงพหุ โดยสังเกตได้จากค่าความผิดพลาดกำลังสองเฉลี่ยในตารางที่ 4.21 ซึ่งจากตารางพบว่ารูปร่างของรอยเชื่อม คือ Penetration, Reinforcement height และ Cross section ที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียมมีค่าความผิดพลาดกำลังสองเฉลี่ยน้อยกว่าค่าที่ได้จากวิธีการวิเคราะห์ความถดถอยเชิงพหุ ยกเว้น Bead width เพียงค่าเดียวที่วิธีการวิเคราะห์ความถดถอยเชิงพหุให้ค่าความผิดพลาดกำลังสองเฉลี่ยน้อยกว่า จึงสรุปได้ว่าโดยเฉลี่ยแล้ววิธีโครงข่ายประสาทเทียมสามารถพยากรณ์รูปร่างของรอยเชื่อมได้ดีกว่าวิธีการวิเคราะห์ความถดถอยเชิงพหุ

นอกจากจะแสดงผลการพยากรณ์ในรูปของค่าความผิดพลาดกำลังสองเฉลี่ยแล้ว ยังได้แสดงผลการพยากรณ์ในรูปแบบของกราฟเส้นดังแสดงในรูปที่ 4.1 ถึง 4.4 เพื่อแสดงให้เห็นถึงแนวโน้มของข้อมูล และกราฟการกระจายดังแสดงในรูปที่ 4.5 ถึง 4.12 เพื่อแสดงให้เห็นถึงการกระจายตัวของข้อมูล

ตารางที่ 4.19 แสดงค่าที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียม

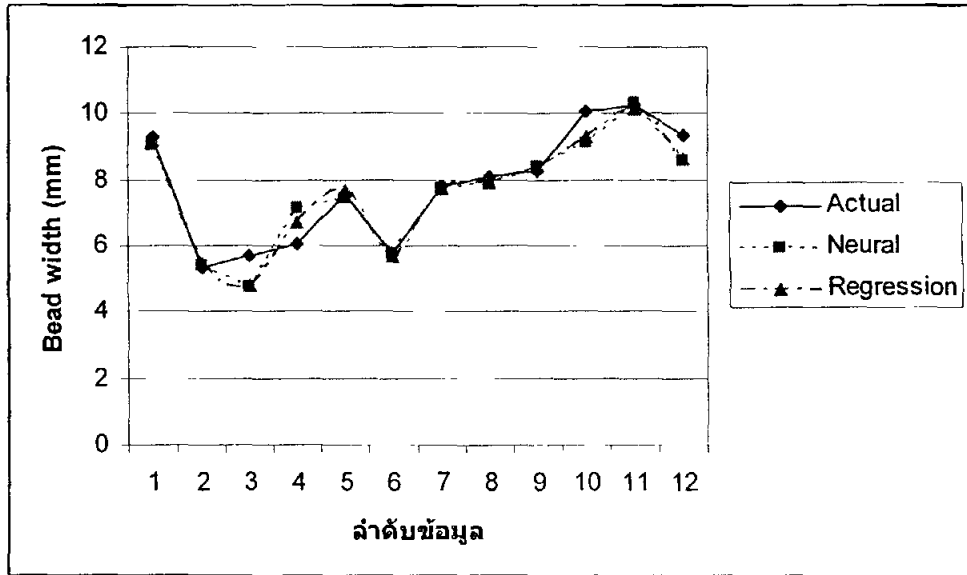
ลำดับที่	Bead width (mm)	Penetration (mm)	Reinforcement height (mm)	Cross section (mm ²)
1	9.055385	2.155992	0.662544	4.157765
2	5.378151	1.455788	1.731984	6.882830
3	4.723310	1.720184	0.859920	2.553722
4	7.145378	1.800982	0.894288	4.193927
5	7.411468	2.904116	0.913872	4.414343
6	5.733154	1.336812	2.067792	8.813192
7	7.745703	1.454808	0.412368	2.055203
8	7.945595	1.649580	0.501648	2.580413
9	8.414822	0.897130	0.747792	4.096634
10	9.139755	2.754108	0.253584	2.443514
11	10.341054	1.863242	0.976656	6.472133
12	8.551112	2.135392	0.176784	1.894196

ตารางที่ 4.20 แสดงค่าที่ได้จากการพยากรณ์โดยใช้การวิเคราะห์ความถดถอยเชิงพหุ

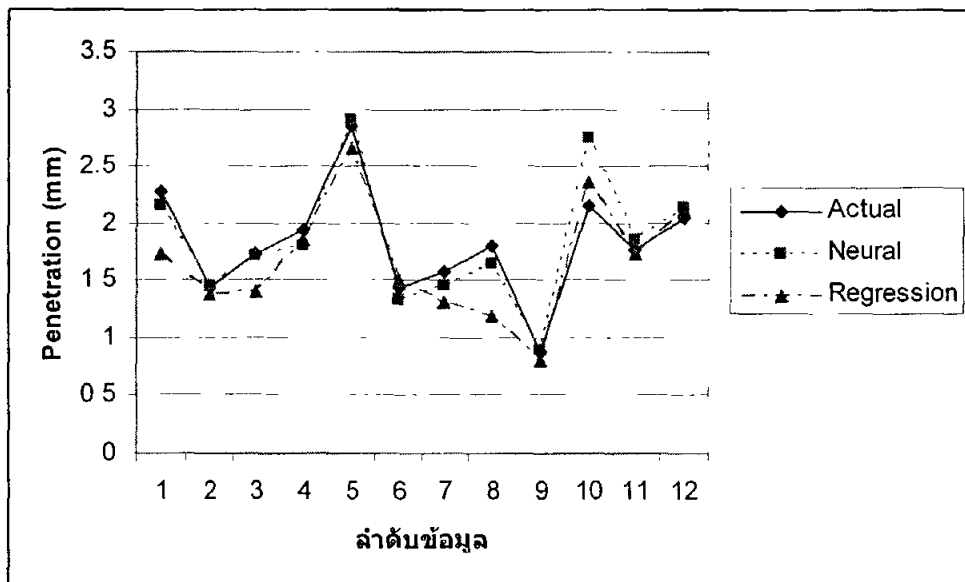
ลำดับที่	Bead width (mm)	Penetration (mm)	Reinforcement height (mm)	Cross section (mm ²)
1	9.172240	1.726409	0.673390	3.221696
2	5.442580	1.383478	1.687429	6.806207
3	4.802281	1.394928	0.749482	2.693783
4	6.739053	1.853333	1.024668	4.149590
5	7.708430	2.657017	0.971765	4.384972
6	5.677639	1.498380	1.795689	7.883579
7	7.721353	1.316568	0.397652	1.654977
8	7.892474	1.190195	0.485420	2.032964
9	8.405027	0.811672	0.748312	3.764625
10	9.272146	2.361548	0.500455	2.168628
11	10.126133	1.730875	0.938471	6.053846
12	8.659686	2.104932	0.410498	1.739772

ตารางที่ 4.21 แสดงค่าความผิดพลาดกำลังสองเฉลี่ยของรูปร่างรอยเชื่อมที่ได้จากการพยากรณ์โดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุ

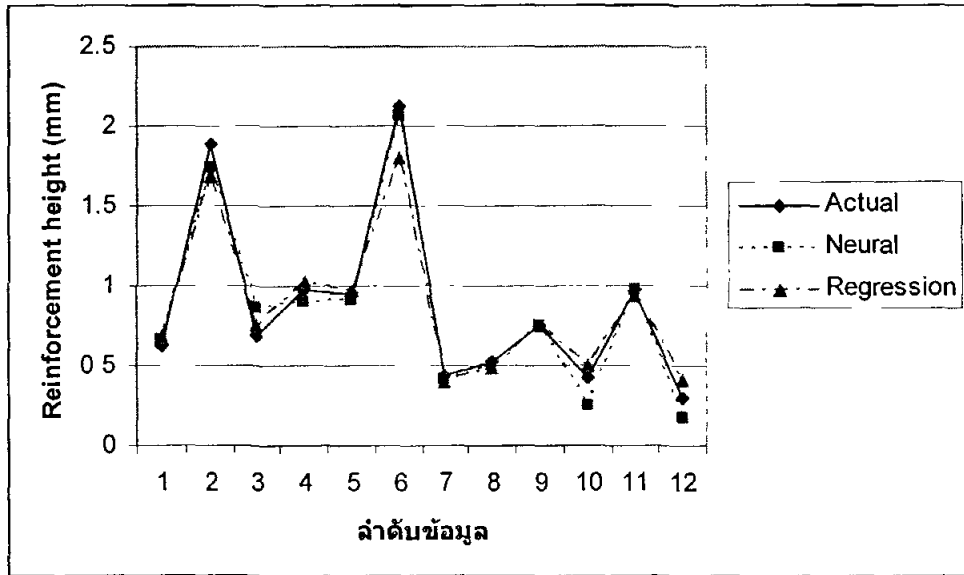
	Bead width	Penetration	Reinforcement height	Cross section	เฉลี่ย
โครงข่ายประสาทเทียม	0.316354	0.038168	0.009192	0.110267	0.118495
การวิเคราะห์ความถดถอยเชิงพหุ	0.211308	0.080181	0.015429	0.334027	0.160236



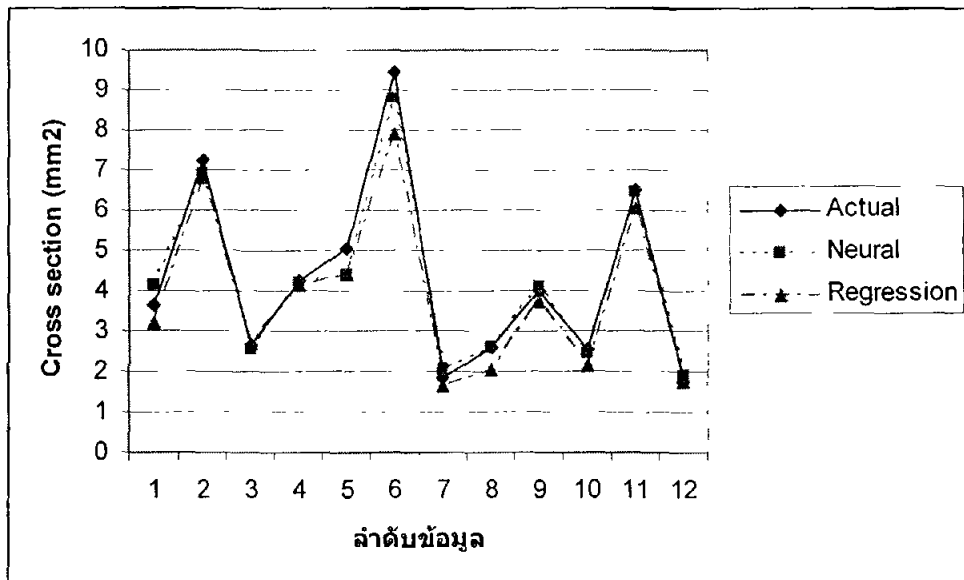
รูปที่ 4 1 Bead width ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง



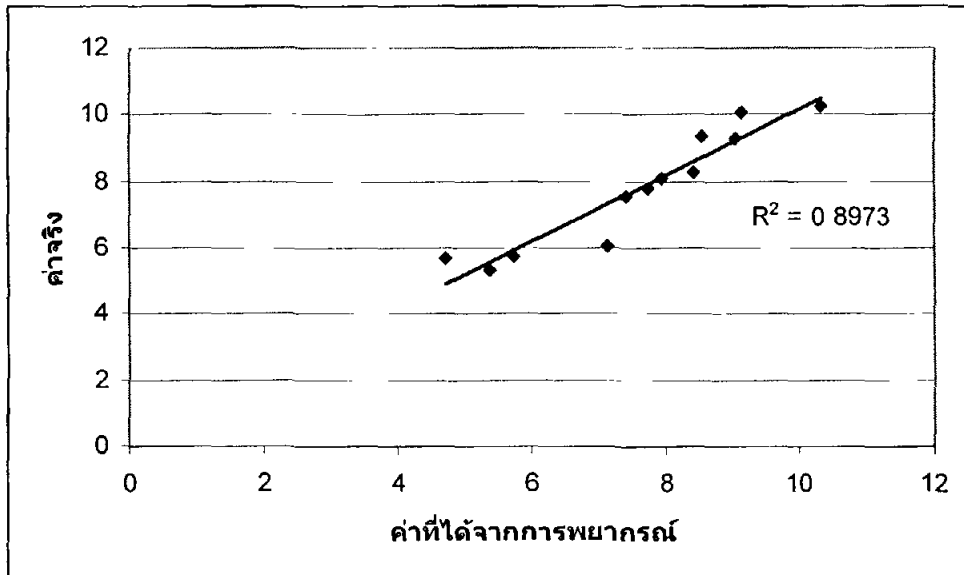
รูปที่ 4 2 Penetration ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง



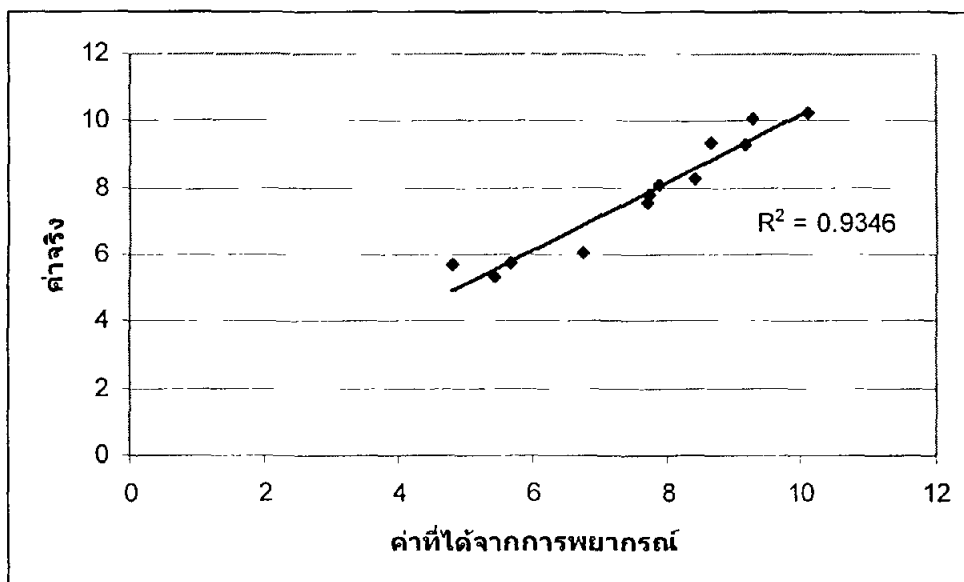
รูปที่ 4.3 Reinforcement height ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง



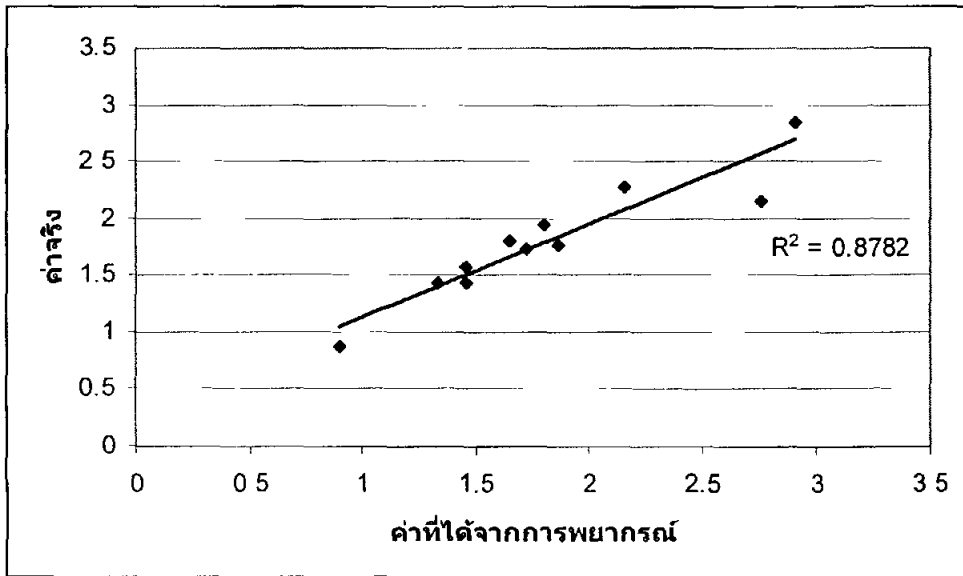
รูปที่ 4.4 Cross section ที่ได้จากการพยากรณ์เทียบกับข้อมูลจริง



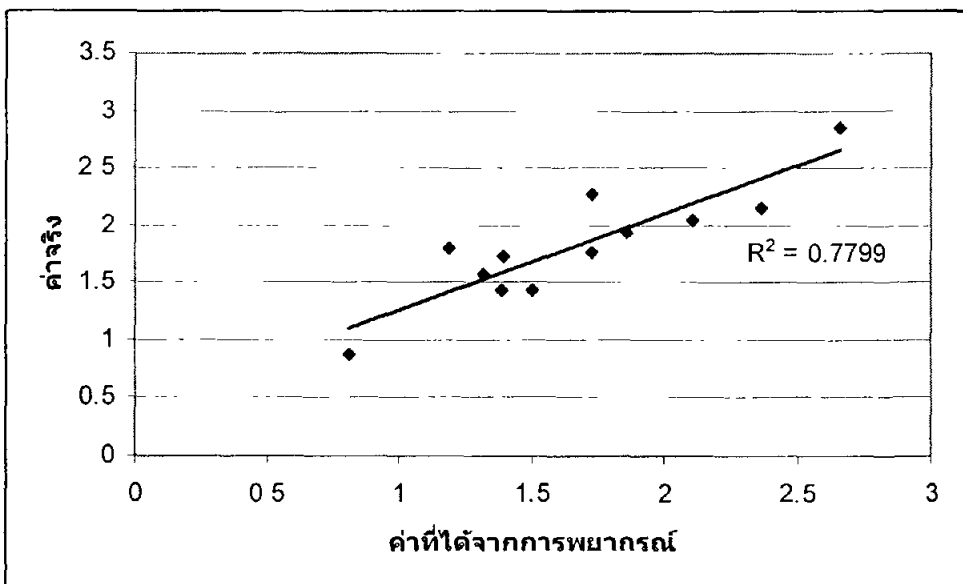
รูปที่ 4.5 การกระจายของ Bead width ที่ได้จากการพยากรณ์
โดยใช้โครงข่ายประสาทเทียม



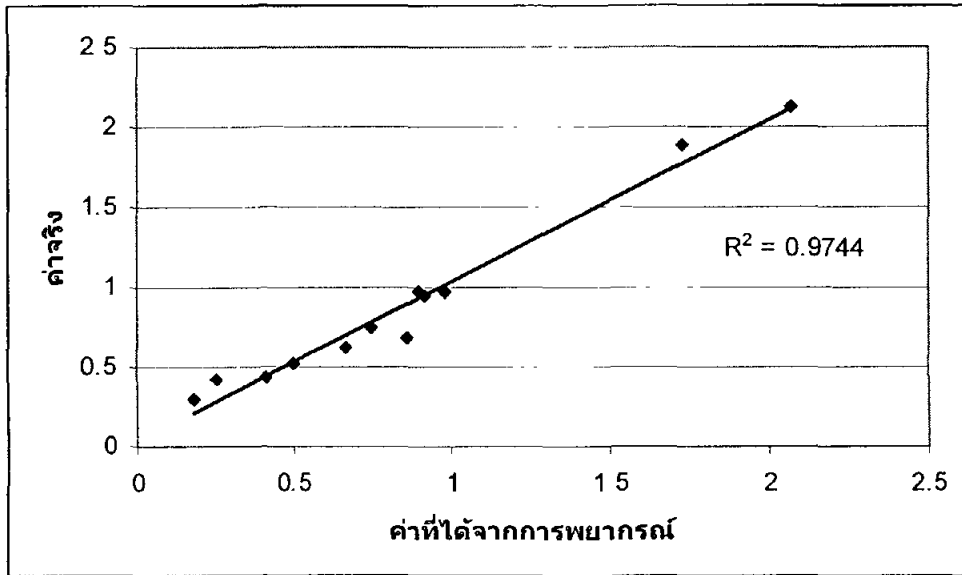
รูปที่ 4.6 การกระจายของ Bead width ที่ได้จากการพยากรณ์
โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ



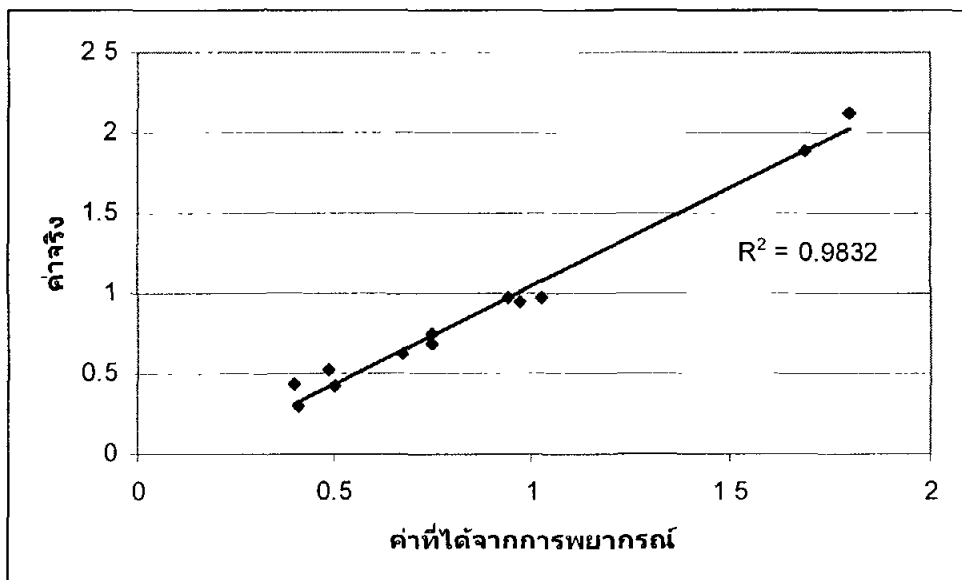
รูปที่ 4 7 การกระจายของ Penetration ที่ได้จากการพยากรณ์
โดยใช้โครงข่ายประสาทเทียม



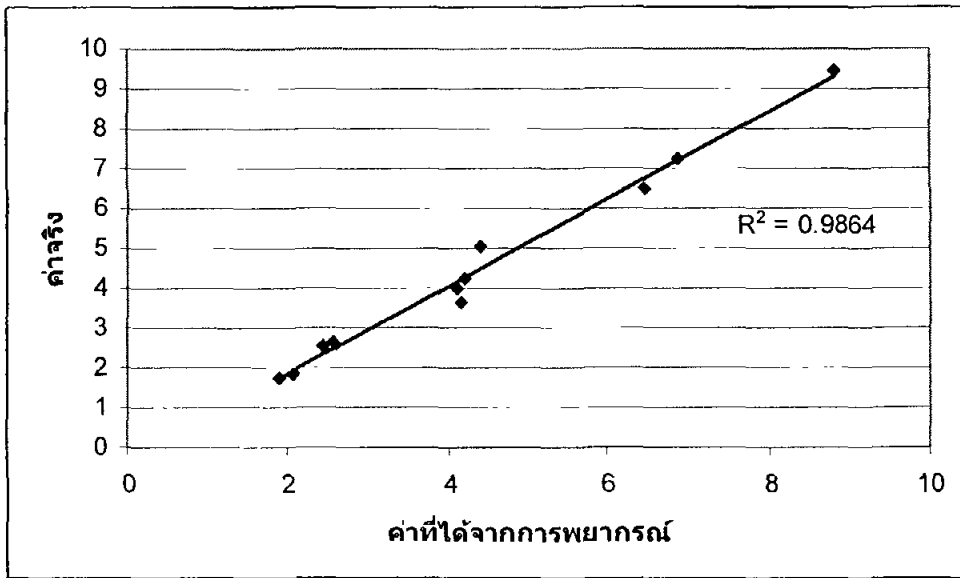
รูปที่ 4 8 การกระจายของ Penetration ที่ได้จากการพยากรณ์
โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ



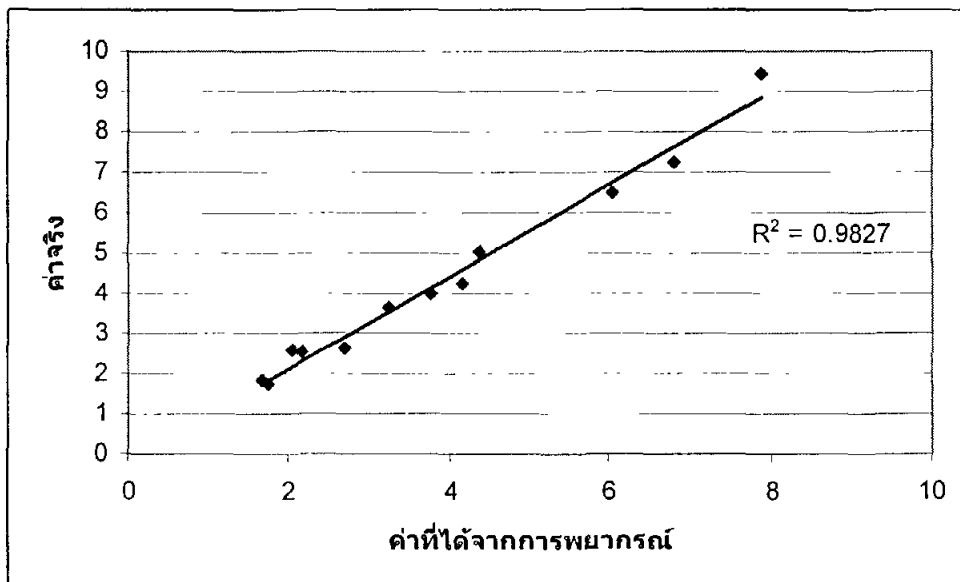
รูปที่ 4 9 การกระจายของ Reinforcement height ที่ได้จากการพยากรณ์
โดยใช้โครงข่ายประสาทเทียม



รูปที่ 4 10 การกระจายของ Reinforcement height ที่ได้จากการพยากรณ์
โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ



รูปที่ 4.11 การกระจายของ Cross section ที่ได้จากการพยากรณ์
โดยใช้โครงข่ายประสาทเทียม



รูปที่ 4.12 การกระจายของ Cross section ที่ได้จากการพยากรณ์
โดยวิธีการวิเคราะห์ความถดถอยเชิงพหุ

บทที่ 5

สรุปผลและข้อเสนอแนะ

ปริญญานิพนธ์นี้ได้ศึกษาการพยากรณ์รูปร่างของรอยเชื่อม โดยใช้โครงข่ายประสาทเทียมและการวิเคราะห์ความถดถอยเชิงพหุและเปรียบเทียบผลของการพยากรณ์ทั้งสองวิธีในรูปของค่าความผิดพลาดกำลังสองเฉลี่ย โดยปัจจัยที่ใช้ในการพยากรณ์ ได้แก่ Arc current, Arc Voltage, Travel speed และ Wire speed และรูปร่างของรอยเชื่อมที่ต้องการพยากรณ์คือ Bead width, Penetration, Reinforcement height และ Cross section ซึ่งหลังจากการพยากรณ์จะแสดงค่าผลลัพธ์ที่ได้ในรูปของกราฟเส้นและกราฟการกระจายเพื่อแสดงให้เห็นถึงแนวโน้มและการกระจายตัวของค่าที่ได้จากการพยากรณ์

5.1 สรุปผลการวิจัย

ปริญญานิพนธ์ฉบับนี้ ได้แบ่งข้อมูลที่จะนำไปใช้ในการพยากรณ์ออกเป็นสองกลุ่ม กลุ่มแรก คือ ชุดข้อมูลที่ใช้สำหรับการฝึกสอนในโครงข่ายประสาทเทียมและการดำเนินการของการวิเคราะห์ความถดถอยเชิงพหุจำนวน 30 ข้อมูล กลุ่มที่สอง คือ ชุดข้อมูลที่ใช้สำหรับการทดสอบ จำนวน 12 ข้อมูล ซึ่งจากการพยากรณ์ทั้งสองวิธีสามารถสรุปได้ดังนี้

5.1.1 การพยากรณ์รูปร่างรอยเชื่อมโดยใช้โครงข่ายประสาทเทียม

การพยากรณ์รูปร่างรอยเชื่อมโดยใช้โครงข่ายประสาทเทียมในปริญญานิพนธ์ฉบับนี้ ได้เลือกใช้โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า และใช้กฎการเรียนรู้แบบแพร่ค่าย้อนกลับของ Levenberg-Marquardt Algorithm ในโปรแกรม MATLAB R2007a ซึ่งจากการศึกษาผลของจำนวนชั้นซ่อน จำนวนเซลล์ประสาทในชั้นซ่อนและชนิดของฟังก์ชันถ่ายโอนที่มีผลต่อการฝึกสอนและการพยากรณ์ พบว่า โครงสร้างของโครงข่ายประสาทเทียมที่เหมาะสมคือ โครงข่ายประสาทเทียมที่มี 1 ชั้นซ่อน จำนวนเซลล์ประสาทในชั้นซ่อน 6 เซลล์ ใช้ฟังก์ชันถ่ายโอนชนิด Log-sigmoid ในชั้นซ่อนและใช้ฟังก์ชันถ่ายโอนชนิด Linear ในชั้นตัวแปรด้านออก จากผลการพยากรณ์ โดยใช้โครงสร้างดังกล่าว ได้ค่าความผิดพลาดกำลังสองเฉลี่ยเท่ากับ 0.118459 ซึ่งมีค่าต่ำที่สุด

5.1.2 การพยากรณ์รูปร่างของรอยเชื่อมโดยใช้การวิเคราะห์ความถดถอยเชิงพหุ

จากการใช้การวิเคราะห์ความถดถอยเชิงพหุในโปรแกรม SPSS 15.0 For Windows เพื่อใช้พยากรณ์รูปร่างรอยเชื่อม นั้น พบว่า จากผลการพยากรณ์ได้ค่าความผิดพลาดกำลังสองเฉลี่ยน้อยที่สุดเท่ากับ 0.160236 ซึ่งมีค่ามากกว่าการใช้โครงข่ายประสาทเทียม

จากค่าความผิดพลาดกำลังสองเฉลี่ย แสดงให้เห็นว่าโครงข่ายประสาทเทียมสามารถพยากรณ์รูปร่างของรอยเชื่อมได้ดีกว่าการวิเคราะห์ความถดถอยเชิงพหุ

5.2 สิ่งที่ต้องพัฒนาต่อไป

1. ควรนำโครงข่ายประสาทเทียมชนิดอื่นๆ เช่น Recurrent Networks มาประยุกต์ใช้กับการพยากรณ์รูปร่างรอยเชื่อม เพื่อทำการเปรียบเทียบข้อดีข้อเสียกับโครงข่ายประสาทเทียมชนิดที่ใช้อยู่
2. พัฒนาระบบขั้นตอนการปรับสอนให้มีความรวดเร็วมากขึ้น เนื่องจากโครงข่ายประสาทเทียมเมื่อใช้กับตัวแปรด้านเข้าจำนวนมาก จะทำให้โครงข่ายประสาทเทียมใช้เวลาในการคำนวณนานขึ้น

หนังสืออ้างอิง

- กัลยา วานิชย์บัญชา, รศ.ดร., 2548 การใช้ SPSS for Window ในการวิเคราะห์ข้อมูล. พิมพ์ครั้งที่ 7. กรุงเทพฯ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- กัลยา วานิชย์บัญชา, รศ.ดร., 2549 การวิเคราะห์สถิติขั้นสูงด้วย SPSS. พิมพ์ครั้งที่ 5. กรุงเทพฯ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย
- ฉัตรชัย สานติสุขรัตน์, 2548. การประเมินความเชื่อถือได้ของระบบผลิตกำลังไฟฟ้า โดยการใช้โครงข่ายประสาทเทียม. วิทยานิพนธ์ปริญญาโท สาขาวิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- ภูริสันธ์ ลักขร, 2548 การทำนายการจ่ายโหลดอย่างประหยัดโดยวิธีโครงข่ายประสาทเทียม. วิทยานิพนธ์ปริญญาโท สาขาวิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- อนันต์ วิริยะกิจชัย, 2549. การประยุกต์ใช้โครงข่ายประสาทเทียม สำหรับการพยากรณ์ค่ากำลังรับแรงอัดของตัวอย่างคอนกรีตทรงกระบอก. วิทยานิพนธ์ปริญญาโท สาขาวิชาวิศวกรรมก่อสร้างและการจัดการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- โอฬาร ชรรมานนท์, 2548 การพยากรณ์ความต้องการกำลังไฟฟ้าสูงสุดระยะยาวของระบบ กฟผ โดยใช้โครงข่ายประสาทเทียม. วิทยานิพนธ์ปริญญาโท สาขาวิชาวิศวกรรมไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- Hong Tae Kang and Jung Me Park, 2006. Prediction of fatigue life for spot weld using back-propagation neural networks University of Michigan-Dearborn. United States.