บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเพื่อประยุกต์ใช้วิธีวิเคราะห์ข้อมูลในการพยากรณ์เกรดเฉลี่ยด้วยวิธีโครงข่ายประสาทเทียมและวิธีการถดถอยพหุคูณ พร้อมทั้งตรวจสอบผลการพยากรณ์โดยใช้เกณฑ์ในการเปรียบเทียบจากการวัดความคลาดเคลื่อน 2 แบบ ได้แก่ รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error หรือ RMSE) และเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error หรือ MAPE) ผู้วิจัยนำเสนอรายละเอียดวิธีดำเนินการวิจัย ดังนี้

3.1 ประชากรและกลุ่มตัวอย่าง

3.2 เครื่องมือที่ใช้ในการวิจัย

3.3 การเก็บรวบรวมข้อมูล

3.4 การวิเคราะห์ข้อมูล

3.1 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยครั้งนี้คือ นักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ ประจำปีการศึกษา 2553 จำนวน 291 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ ประจำปีการศึกษา 2553 ผู้วิจัยสุ่มตัวอย่างแบบแบ่งชั้น (Stratified Random Sampling) โดยกำหนดขนาดกลุ่มตัวอย่างด้วยสูตรของ Krejcie and Morgan (1970 : 607 – 610) ดังสมการ (3-1) ดังนี้

$$s = \frac{\chi^2 NP(1-P)}{d^2(N-1)+\chi^2 P(1-P)} \qquad (3\text{-}1)$$

เมื่อ s คือ ขนาดกลุ่มตัวอย่างที่ต้องการ

χ^2 คือ ค่าไค-สแควร์ที่ degree of freedom เป็น 1 และระดับความเชื่อมั่นที่ต้องการ (3.841)

N คือ ขนาดของประชากร

P คือ สัดส่วนของลักษณะที่สนใจในประชากร (ประมาณการ = 0.5 ซึ่งจะทำให้ได้ขนาดกลุ่มตัวอย่างที่ใหญ่ที่สุด)

d คือ ระดับความคลาดเคลื่อนที่ยอมรับได้ในรูปของสัดส่วน (0.05)

ได้กลุ่มตัวอย่างรวม 165 คน จำแนกตามสาขาวิชาและชั้นปี แสดงดังตารางที่ 3-1

ตารางที่ 3-1 จำนวนนักศึกษาคณะวิทยาศาสตร์และเทคโนโลยี

ชั้นปี	สาขาวิชา			
	วิทยาการคอมพิวเตอร์		จุลชีววิทยาอุตสาหกรรม	
	ประชากร (N)	ตัวอย่าง (n)	ประชากร (N)	ตัวอย่าง (n)
1	44	25	71	40
2	31	18	53	30
3	22	12	21	12
4	27	15	22	13
รวม	**124**	**70**	**167**	**95**

3.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบสอบถามเกี่ยวกับเกรดเฉลี่ยและพฤติกรรมการใช้อินเทอร์เน็ตของนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ โดยมีขั้นตอนการพัฒนาเครื่องมือ ดังนี้

1. ศึกษาวิธีการสร้างแบบสอบถามจากเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องกับพฤติกรรมการใช้อินเทอร์เน็ต

2. พิจารณาพฤติกรรมการใช้อินเทอร์เน็ตที่คาดว่าจะมีความสัมพันธ์กับเกรดเฉลี่ยของนักศึกษาจากเอกสารและงานวิจัยที่เกี่ยวข้อง

3. สร้างแบบสอบถามฉบับร่างซึ่งมีทั้งหมด 2 ตอน ประกอบด้วย ตอนที่ 1 ข้อมูลสภาพทั่วไปและผลสัมฤทธิ์ทางการเรียนและตอนที่ 2 พฤติกรรมการใช้อินเทอร์เน็ต

4. นำแบบสอบถามให้ผู้เชี่ยวชาญ จำนวน 3 ท่าน (รายละเอียดดังภาคผนวก ก) ตรวจสอบดัชนีความสอดคล้องของข้อคำถามกับจุดประสงค์ (Index of Consistency: IOC) โดยใช้

สูตรของโรวิเนลลีและแฮมเบลตัน (Rowinelli & Hambleton. 1997 อ้างถึงใน บุญเชิด ภิญโญ-อนันตพงษ์. 2545 : 95) ดังสมการ (3-2) ดังนี้

$$IOC = \frac{\sum R}{N} \qquad (3\text{-}2)$$

เมื่อ IOC แทน ดัชนีความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์

$\sum R$ แทน ผลรวมคะแนนของผู้เชี่ยวชาญ

N แทน จำนวนผู้เชี่ยวชาญ

โดยใช้เกณฑ์การประเมิน ดังนี้

+1 หมายถึง แน่ใจว่าข้อคำถามวัดจุดประสงค์นั้น

0 หมายถึง ไม่แน่ใจว่าข้อคำถามวัดจุดประสงค์นั้น

-1 หมายถึง แน่ใจว่าข้อคำถามวัดไม่ตรงกับจุดประสงค์นั้น

5. นำแบบสอบถามที่ได้รับการตรวจสอบจากผู้เชี่ยวชาญ คำนวณหาค่า IOC และคัดเลือกข้อสอบที่มีค่า IOC ตั้งแต่ 0.50 ขึ้นไป

ผลการตรวจสอบแบบสอบถามโดยผู้เชี่ยวชาญได้ข้อคำถามผ่านเกณฑ์จำนวน 62 ข้อ โดยมีค่า IOC อยู่ระหว่าง 0.67 – 1.00

6. นำแบบสอบถามไปทดลองใช้กับนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติที่มีลักษณะใกล้เคียงกับกลุ่มตัวอย่าง จำนวน 50 คน

7. หาค่าความเชื่อมั่น (Reliability) ของแบบสอบถามด้วยค่าสัมประสิทธิ์ความเชื่อมั่นอัลฟา (Alpha Coefficient) ของครอนบาค (Cronbach's Alpha) โดยมีสูตรในการคำนวณดังสมการ (3-3) ดังนี้

$$\alpha = \frac{k}{k-1}\left\{1 - \frac{\sum_{i=1}^{k} S_i^2}{S_t^2}\right\} \qquad (3\text{-}3)$$

เมื่อ α แทน ค่าสัมประสิทธิ์ความเชื่อมั่นของแบบสอบถาม

k แทน จำนวนข้อคำถามในแบบสอบถาม

S_i^2 แทน ความแปรปรวนของคะแนนแต่ละข้อ

S_t^2 แทน ความแปรปรวนของคะแนนผู้ตอบแบบสอบถามทั้งหมด

ผลการหาค่าความเชื่อมั่น ได้ค่าความเชื่อมั่นเท่ากับ 0.89 ซึ่งมีความน่าเชื่อถือระดับสูง

8. จัดพิมพ์แบบสอบถามฉบับสมบูรณ์แล้วนำไปเก็บรวบรวมข้อมูลกับนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ ซึ่งเป็นกลุ่มตัวอย่าง จำนวน 165 คน

3.3 การเก็บรวบรวมข้อมูล

การวิจัยครั้งนี้ ผู้วิจัยบันทึกข้อมูลด้วยตนเอง โดยมีขั้นตอนการรวบรวมข้อมูลดังต่อไปนี้

1. การติดต่อขอความร่วมมือ ผู้วิจัยดำเนินการติดต่อขอความร่วมมือไปยังสาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาจุลชีววิทยาอุตสาหกรรมอุตสาหกรรม คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ เพื่อขอเก็บรวบรวมข้อมูลจากแบบสอบถามพฤติกรรมการใช้อินเทอร์เน็ตของนักศึกษาคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ

2. การบรรณาธิกรณ์ (Edit) ข้อมูล ผู้วิจัยตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูลที่ได้จากการรวบรวมข้อมูลมาทั้งหมด โดยใช้โปรแกรม Microsoft Excel และ SPSS for Window

3.4 การวิเคราะห์ข้อมูล

สำหรับขั้นตอนการวิเคราะห์แบ่งเป็น 5 ตอน ดังนี้ ตอนที่ 1 การวิเคราะห์ข้อมูลเบื้องต้น ตอนที่ 2 การวิเคราะห์องค์ประกอบ ตอนที่ 3 การวิเคราะห์ด้วยวิธีการถดถอยพหุคูณ ตอนที่ 4 การวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม และตอนที่ 5 การตรวจสอบผลการพยากรณ์ มีรายละเอียดดังต่อไปนี้

ตอนที่ 1 การวิเคราะห์ข้อมูลเบื้องต้น

เพื่อศึกษาสถานภาพทั่วไปของนักศึกษา ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ค่าความเบ้ ค่าความโด่ง และส่วนเบี่ยงเบนมาตรฐาน

ตอนที่ 2 การวิเคราะห์องค์ประกอบ

ผู้วิจัยดำเนินการวิเคราะห์องค์ประกอบ ซึ่งประกอบด้วย 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การตรวจสอบเมตริกซ์สหสัมพันธ์ (Correlation Matrix) เป็นขั้นตอนแรกของการวิเคราะห์องค์ประกอบเพื่อหาความสัมพันธ์ในรูปเชิงเส้นตรงโดยวิธีสหสัมพันธ์

เพียร์สัน (Pearson Correlation) ระหว่างตัวแปรทุกตัวที่ต้องการนำมาจัดกลุ่ม ซึ่งจะอยู่ในรูปของเมตริกซ์สหสัมพันธ์ (Correlation Matrix) ส่วนค่าที่ใช้ตรวจสอบว่าข้อมูลเหมาะสมที่จะวิเคราะห์องค์ประกอบหรือไม่จะพิจารณาที่ค่า Barlett's Test of Sphericity (กัลยา วานิชย์บัญชา. 2553 : 204) และค่า The Kaiser – Meyer – Olkin (KMO) หรือ Measure of Sampling Adequacy (กัลยา วานิชย์บัญชา. 2553 : 204) ดังสมการ (3-4) ดังนี้

$$KMO = \frac{\sum r_i^2}{\sum r_i^2 + \sum (partial \ \ correlation)^2} \qquad (3\text{-}4)$$

R = ค่าสัมประสิทธิ์สหสัมพันธ์ซึ่งทำให้ค่า $0 \le KMO \le 1$

ถ้าค่า KMO มีค่าน้อย (เข้าใกล้ 0) แสดงว่าการวิเคราะห์องค์ประกอบไม่เหมาะสมกับข้อมูลที่มีอยู่ แต่ถ้าค่า KMO มีค่ามาก (เข้าใกล้ 1) แสดงว่าการวิเคราะห์องค์ประกอบเหมาะสมกับข้อมูลที่มีอยู่ โดยทั่วไป ถ้าค่า KMO < 0.5 จะถือว่า ข้อมูลที่มีอยู่ไม่เหมาะสมที่จะใช้การวิเคราะห์องค์ประกอบ

ขั้นตอนที่ 2 การสกัดองค์ประกอบ (Factor Extraction) เป็นการหาองค์ประกอบเริ่มต้น (Initial Factor) เพื่อให้องค์ประกอบสามารถแทนตัวแปรทั้งหมดได้อย่างเพียงพอ วิธีการสกัดองค์ประกอบมีหลายวิธี สำหรับงานวิจัยนี้ ผู้วิจัยได้ใช้วิธีองค์ประกอบหลัก (Principal Component Analysis: PCA) ในการสกัดองค์ประกอบ เนื่องจากการสกัดองค์ประกอบด้วยวิธีดังกล่าวเป็นวิธีการที่อาศัยหลักการสัมพันธ์เชิงเส้นระหว่างตัวแปร โดยนำตัวแปรแต่ละตัวที่เหมือนกันมาสร้างเป็นสมการเส้นตรง ทำให้ได้จำนวนองค์ประกอบน้อยที่สุด และองค์ประกอบที่ได้สามารถอธิบายความผันแปรทั้งหมด ผลจากการสกัดองค์ประกอบจะได้ค่าน้ำหนักขององค์ประกอบ (Factor Loading) ซึ่งเป็นค่าที่แสดงถึงความสัมพันธ์ของตัวแปรหนึ่ง ๆ กับแต่ละองค์ประกอบที่หามาได้จากการสกัดองค์ประกอบ หลังจากได้ค่าน้ำหนักองค์ประกอบแล้ว ผู้วิจัยนำมาคำนวณหาค่าสถิติ ดังต่อไปนี้

1. Communalities เป็นค่าที่คำนวณจากการเอาค่าน้ำหนักองค์ประกอบแต่ละตัวมายกกำลังสอง ซึ่งเป็นค่าที่แสดงถึงอิทธิพลของแต่ละองค์ประกอบที่มีผลต่อตัวแปร

2. Eigenvalue เป็นค่าที่คำนวณมาจากค่า Communalities ขององค์ประกอบที่นำมารวมกันของตัวแปรซึ่งจะเป็นตัวบอกว่าจะสามารถมีองค์ประกอบได้กี่ตัว

3. ค่า Percentage of Variance คำนวณจากค่า Eigenvalue มาคิดเป็นร้อยละโดยเทียบกับยอดรวม สามารถอธิบายการเปลี่ยนแปลงของตัวแปรเป็นค่าร้อยละ

ขั้นตอนที่ 3 การหมุนแกน (Rotation) เป็นขั้นตอนดำเนินการแยกตัวแปรให้เห็นชัดเจนว่าควรจัดตัวแปรใดไว้ในกลุ่มใดหรือ Factor ใด เนื่องจากในการสกัดองค์ประกอบได้องค์ประกอบหลายตัว ซึ่งแต่ละองค์ประกอบเกิดจากการรวมตัวกันของตัวแปรแบบเชิงเส้น แต่ปัญหาที่มีคือตัวแปรบางตัวอาจเป็นสมาชิกของหลายองค์ประกอบ และการหมุนแกนก็เป็นวิธีการที่จะทำให้การเป็นสมาชิกขององค์ประกอบต่าง ๆ ชัดเจนขึ้น ซึ่งการหมุนแกนได้จากการนำค่าน้ำหนักหรือค่าสัมประสิทธิ์ของตัวแปรในสมการของแต่ละองค์ประกอบมาเขียนกราฟโดยมีจำนวนแกนเท่ากับจำนวนองค์ประกอบที่มีอยู่ สำหรับงานวิจัยนี้ผู้วิจัยใช้การหมุนแกนแบบตั้งฉาก (Orthogonal Rotation) ด้วยวิธี Varimax ซึ่งเป็นแบบที่เน้นการแปลความหมายขององค์ประกอบโดยการลดจำนวนตัวแปรในแต่ละองค์ประกอบให้เหลือน้อยที่สุด จะทำให้ได้เฉพาะที่มีค่าสัมประสิทธิ์ในการรวมตัวแปรแบบเชิงเส้นสูง

ขั้นตอนที่ 4 คะแนนองค์ประกอบ (Factor Score)

เมื่อหมุนแกนองค์ประกอบแล้ว ผู้วิจัยนำคะแนนที่ได้ไปวิเคราะห์โดยใช้สัมประสิทธิ์ของตัวแปรแต่ละตัวมาเป็นค่าน้ำหนักในการผสมตัวแปร สำหรับงานวิจัยนี้ผู้วิจัยใช้วิธีการถดถอย (Regression Factor Scores)

ตอนที่ 3 การวิเคราะห์ด้วยวิธีการถดถอยพหุคูณ

ผู้วิจัยดำเนินการวิเคราะห์ข้อมูล รายละเอียด ดังนี้

ตรวจสอบข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยพหุคูณ ซึ่งในการวิเคราะห์ถดถอยพหุคูณ มีข้อตกลงเบื้องต้น ดังนี้

1. ค่าเฉลี่ยของความคลาดเคลื่อนมีค่าเป็นศูนย์
2. ความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่
3. ความคลาดเคลื่อนเป็นค่าสุ่มที่เป็นอิสระต่อกัน
4. ความคลาดเคลื่อนมีการแจกแจงปกติ

โดยผู้วิจัยดำเนินการตรวจสอบข้อตกลงเบื้องต้น ดังนี้

1. ตรวจสอบความคลาดเคลื่อนมีความเป็นอิสระต่อกันและมีความแปรปรวนคงที่จากสมการที่ได้ นำตัวแปรอิสระที่ได้รับการคัดเลือกเข้าสมการไปแทนค่าหา $\hat{Y}_i$ และ $e_i = Y_i - \hat{Y}_i$ วิเคราะห์ความผิดปกติของค่าความคลาดเคลื่อนที่ไม่เป็นอิสระโดยการเขียนกราฟ Standardized ของคู่ลำดับ $(\hat{Y}_i, e_i)$ หรือตรวจสอบจากค่า Durbin Watson Test ซึ่งถ้ามีค่าอยู่ระหว่าง 1.5 ถึง 2.5 แสดงว่าความคลาดเคลื่อนมีความเป็นอิสระต่อกัน (กัลยา วานิชย์บัญชา. 2545)

2. ตรวจสอบการแจกแจงปกติของความคลาดเคลื่อน โดยทำ Normal Probability Plot โดยการเขียนกราฟระหว่าง $P(e_i \leq \frac{i}{n})$ เมื่อ i เป็นลำดับที่ของ e_i เมื่อจัดลำดับแล้วกับ $P(Z < \frac{e_i}{\sqrt{MSE}})$ หากพบว่า จุดต่าง ๆ ที่ได้เป็นแนวทแยงมุมแสดงว่าความคลาดเคลื่อนมีการแจกแจงปกติ

3. ตรวจสอบการมีความสัมพันธ์กันเองระหว่างตัวแปรอิสระ (Multicollinearity) โดยพิจารณาค่าแฟคเตอร์ความแปรปรวนที่เพิ่มมากขึ้น (Variance Inflation Factor: VIF)

เมื่อ $(VIF)_j$ คือ Variance Inflation Factor สำหรับ b_j

$$= \frac{1}{1-R_j^2} : j = 1,2,3,..., p-1$$

R_j^2 คือ ค่าสัมประสิทธิ์ของตัวกำหนดพหุคูณของสมการถดถอยที่มี X_i เป็นตัวแปรตาม และมีตัวแปร $X_1, X_2, ..., X_{j-1}, X_{j+1}, ..., X_{p-1}$ เป็นตัวแปรอิสระ ถ้า $R_j^2 = 0$ และ $(VIF)_j = 1$ แสดงว่าไม่มีปัญหาตัวแปรอิสระมีความสัมพันธ์กันเอง และถ้า R_j^2 มีค่าเข้าใกล้ 1 $(VIF)_j$ มีค่ามากกว่า 1 แสดงว่าเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์กันเอง และถ้า $(VIF)_j > 10$ แสดงว่าเป็นปัญหารุนแรงต้องมีการแก้ไข (Marquardt. 1975) หรือตรวจสอบด้วยค่า Tolerance, Condition Index ค่าสถิติ Tolerance หมายถึง สัดส่วนความแปรปรวนในตัวแปรที่อธิบายไม่ได้ด้วยตัวแปรอื่น ๆ ถ้าค่า Tolerance มีค่าใกล้ 0 แสดงว่าตัวแปรนั้นมีความสัมพันธ์กันสูงกับตัวแปรอื่น ๆ ค่าที่ใกล้ 1 แสดงว่าตัวแปรนั้นมีความสัมพันธ์ต่ำกับตัวแปรอื่น ๆ ในการวิเคราะห์ข้อมูลค่าสถิติ VIF มีค่าเท่ากับส่วนกลับของค่า Tolerance จึงแปลความหมายตรงกันข้ามกล่าวคือ ค่า VIF ที่สูงมาก (ค่าสูงสุดเท่ากับ 10.0) แสดงว่ามีปัญหาภาวะร่วมเส้นตรงพหุสูงมากในกลุ่มตัวแปรอิสระด้วยกัน ส่วนค่าสถิติ Condition Index เป็นค่าสัดส่วนความแปรปรวนซึ่งวัดจากค่าไอเกน (Eigenvalue) เกณฑ์ที่ใช้ตรวจสอบคือ 30 ถ้าตัวแปรใดมีค่า Condition Index เกิน 30 แสดงว่าตัวแปรนั้นมีปัญหาภาวะร่วมเส้นตรงพหุ ค่าสถิติ Condition Index นี้ใช้ตรวจสอบด้วยกระบวนการ 2 ขั้นตอน ขั้นตอนแรก ตรวจดูว่าตัวแปรต้นตัวใดมีค่าเกินเกณฑ์ ขั้นตอนที่สองตรวจสอบสัดส่วนของความแปรปรวนของสัมประสิทธิ์การถดถอย เฉพาะเมื่อตัวแปรอิสระมีค่าสถิติ Condition Index สูงกว่าเกณฑ์ ถ้าค่าสัดส่วนของความแปรปรวนของสัมประสิทธิ์การถดถอยมีค่าสูงกว่า 0.90 แสดงว่าตัวแปรตัวนั้นมีความสัมพันธ์กับตัวแปรอื่นสูง

การวิจัยครั้งนี้ ผู้วิจัยดำเนินการวิเคราะห์การถดถอยพหุคูณ 2 แบบ คือ การถดถอยพหุคูณโดยวิธีการคัดเลือกตัวแปรเพื่อการพยากรณ์แบบปกติ (Enter or All Possible Multiple Regression Analysis) และการถดถอยพหุคูณโดยวิธีคัดเลือกตัวแปรเพื่อการพยากรณ์แบบขั้นตอน (Stepwise Multiple Regression Analysis)

สำหรับวิธีการถดถอยพหุคูณแบบปกติ (Enter Multiple Regression Analysis) ผู้วิจัยนำตัวแปรอิสระจำนวน 17 ตัวเข้าในสมการ เพื่อศึกษาปริมาณจำนวนตัวแปรอิสระที่มากซึ่งจะทำให้ค่าสัมประสิทธิ์การพยากรณ์มากด้วยนั้น จะมีผลต่อข้อตกลงเบื้องต้นอย่างไร และผลการพยากรณ์ที่ได้หากฝ่าฝืนข้อตกลงจะได้ผลเป็นอย่างไร ทั้งนี้เพื่อนำไปเป็นข้อมูลในการเปรียบเทียบผลการพยากรณ์กับวิธีโครงข่ายประสาทเทียม

สำหรับวิธีการถดถอยพหุคูณแบบขั้นตอน (Stepwise Multiple Regression Analysis) ผู้วิจัยนำตัวแปรอิสระจำนวน 11 ตัวเข้าในสมการเพื่อทำการประมวลผลและจะคัดเลือกตัวแปรแบบขั้นตอน ซึ่งวิธีนี้นอกจากจะได้ผลการวิเคราะห์ที่ไม่ฝ่าฝืนข้อตกลงแล้ว ตัวแปรอิสระที่เหลือจะเป็นสิ่งที่ผู้วิจัยสนใจที่จะนำไปเปรียบเทียบผลการพยากรณ์กับวิธีโครงข่ายประสาทเทียม

ตอนที่ 4 การวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม

ผู้วิจัยดำเนินการวิเคราะห์ในแต่ละขั้นตอน ดังนี้

4.1 การวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม กรณีข้อมูลป้อนเข้า 17 นิวรอล

ผู้วิจัยใช้การวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียมโดยจำนวนข้อมูลป้อนเข้าเหมือนและเท่ากับตัวแปรอิสระที่ใช้การวิเคราะห์การถดถอยพหุคูณแบบปกติ (Enter Multiple Regression)

1. การกำหนดค่าเริ่มต้นให้ตัวแบบโครงข่ายประสาทเทียมทำงาน ดังนี้

1.1 แบบของโครงข่ายประสาทเทียม ประกอบด้วย ชั้นข้อมูลป้อนเข้า (Input Node Layer) ชั้นแอบแฝง (Hidden Node Layer) และชั้นแสดงผลลัพธ์ (Output Node Layer) แบ่งการกำหนดค่าในแต่ละชั้น ดังนี้

1.1.1 ชั้นข้อมูลป้อนเข้า (Input Node Layer) ประกอบด้วยจำนวนข้อมูลป้อนเข้า (Input) นิวรอลหรือจำนวนตัวแปรอิสระ 17 ตัวแปร

1.1.2 ชั้นแอบแฝง (Hidden Node Layer) ประกอบด้วยจำนวนนิวรอลเท่ากับ $\sqrt{MN}$ นิวรอล (Masters. 1993) เมื่อ M เป็นจำนวนนิวรอลของชั้นข้อมูลป้อนเข้า ในที่นี้คือจำนวนตัวแปรอิสระ ซึ่งมีค่าเท่ากับ 17 และ N เป็นจำนวนนิวรอลของชั้นแสดงผลลัพธ์ ซึ่งในที่นี้คือ จำนวนตัวแปรตาม ซึ่งมีค่าเท่ากับ 1 ดังนั้น ชั้นแอบแฝงจะประกอบด้วยนิวรอลจำนวน $\sqrt{MN} = \sqrt{17 \times 1} \cong 4.123$ ซึ่งในที่นี้กำหนดให้มีจำนวน 4 นิวรอล ซึ่งสูตรนี้เป็นการประมาณขนาดชั้นแอบแฝง แต่การกำหนดจำนวนนิวรอลในชั้นแอบแฝง อาจกำหนดได้โดย

- ถ้ามีจำนวนนิวรอลนำเข้าและนิวรอลแสดงผลลัพธ์น้อย ๆ รวมทั้งปัญหามีความซับซ้อน นิวรอลในชั้นแอบแฝงอาจต่ำกว่าความต้องการ

- ถ้าปัญหาเป็นปัญหาที่ง่าย มีจำนวนนิวรอลนำเข้าและนิวรอลแสดงผลลัพธ์มาก จำนวนนิวรอลในชั้นแอบแฝงน้อย ๆ ก็เพียงพอ

- กฎนี้ไม่ควรนำไปใช้กับโครงข่ายที่มีจำนวนนิวรอลนำเข้าเท่ากับจำนวนนิวรอลแสดงผลลัพธ์ (Autoassociative Networks)

1.1.3 ชั้นแสดงผลลัพธ์ (Output Node Layer) ประกอบด้วยจำนวนตัวแปรตาม นั่นคือ 1 ตัว (1 นิวรอล) ซึ่งแสดงถึง เกรดเฉลี่ยของนักศึกษาคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ

1.2 โมเดลที่ใช้ในการวิจัยคือโครงข่ายประสาทเทียมแบบที่มีการเรียนรู้แบบแพร่ย้อนกลับ (Backpropagation Neural Network)

2. ปรับค่าของตัวแปรให้อยู่ในสเกลเดียวกัน ทั้งในส่วนของตัวแปรอินพุทและตัวแปรเอาท์พุทให้มีค่าอยู่ระหว่าง 0 ถึง 1 โดยใช้สมการ (3-5) ดังนี้

$$x_i' = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}} \qquad (3\text{-}5)$$

เมื่อ $x_{\max}$ คือ ค่าที่คาดว่าจะเป็นค่าสูงสุดของข้อมูล

$x_{\min}$ คือ ค่าที่คาดว่าจะเป็นค่าต่ำสุดของข้อมูล

3. กำหนดค่าน้ำหนักเริ่มต้นให้กับเส้นเชื่อมโยงระหว่างโหนด โดยการสุ่มหรือการกำหนดขึ้นเอง ให้มีค่าอยู่ระหว่าง -1 ถึง 1

4. ป้อนข้อมูลที่ปรับค่าแล้วเข้าสู่ชั้นข้อมูลป้อนเข้า

5. ส่งข้อมูลจากชั้นข้อมูลป้อนเข้าไปยังโหนดในชั้นแอบแฝง โดยไม่ผ่านขั้นตอนการคำนวณใด ๆ

6. คำนวณผลรวมของสัญญาณอินพุทของแต่ละโหนดในชั้นแอบแฝง ได้จากสมการ (3-6) ดังนี้

$$I_{o,j,o} = \sum_{i=1}^{n} O_{i,o,o} w_{i,j} \qquad (3\text{-}6)$$

เมื่อ $I_{o,j,o}$ คือ ข้อมูลอินพุทของโหนดที่ j ในชั้นแอบแฝง

$O_{i,o,o}$ คือ ข้อมูลเอาท์พุทของโหนดที่ i ในชั้นข้อมูลป้อนเข้า

$w_{i,j}$ คือ ค่าน้ำหนักบนเส้นเชื่อมโยงจากโหนดที่ i ไปยังโหนดที่ j

7. คำนวณค่าเอาท์พุทของโหนดที่ j ในชั้นแอบแฝงโดยใช้ฟังก์ชันการแปลง (Transfer Function) คำนวณผลลัพธ์จากสมการ (3-7) ดังนี้

$$O_{o,j,o} = \frac{1}{1+e^{-I_{o,j,o}}} \qquad (3\text{-}7)$$

เมื่อ $O_{o,j,o}$ คือ ข้อมูลเอาท์พุทของโหนดที่ j ในชั้นแอบแฝง

$I_{o,j,o}$ คือ ข้อมูลอินพุทของโหนดที่ j ในชั้นแอบแฝง

8. คำนวณผลรวมสัญญาณอินพุทของแต่ละโหนดในชั้นแสดงผลลัพธ์ จากสมการ (3-8) ดังนี้

$$I_{o,o,k} = \sum_{i=1}^{m} O_{o,j,o} w_{j,k} \qquad (3\text{-}8)$$

เมื่อ $I_{o,o,k}$ คือ ข้อมูลอินพุทของโหนดที่ k ในชั้นแสดงผลลัพธ์

$O_{o,j,o}$ คือ ข้อมูลเอาท์พุทของโหนดที่ j ในชั้นแอบแฝง

$w_{j,k}$ คือ ค่าน้ำหนักบนเส้นเชื่อมโยงจากโหนดที่ j ไปยังโหนดที่ k

9. คำนวณค่าเอาท์พุทของโหนด k ในชั้นแสดงผลลัพธ์โดยใช้ฟังก์ชันซิกมอยด์ (Sigmoid Function) คำนวณได้จากสมการ (3-9) ดังนี้

$$O_{o,o,k} = \frac{1}{1+e^{-I_{o,o,k}}} \qquad (3\text{-}9)$$

เมื่อ $O_{o,o,k}$ คือ ข้อมูลเอาท์พุทของโหนดที่ k ในชั้นแอบแฝง

$I_{o,o,k}$ คือ ข้อมูลอินพุทของโหนดที่ k ในชั้นแสดงผลลัพธ์

10. คำนวณความคลาดเคลื่อนโดยการนำค่าเอาท์พุทของโหนดที่ k ในชั้นแสดงผลลัพธ์มาเปรียบเทียบค่าเป้าหมาย ซึ่งหาได้จากสมการ (3-10) ดังนี้

$$e_k = D_k - O_{o,o,k} \qquad (3\text{-}10)$$

เมื่อ e_k คือ ความคลาดเคลื่อนของโหนดที่ k

D_k คือ ข้อมูลจริงหรือค่าเป้าหมายของโหนดที่ k ในชั้นแสดงผลลัพธ์

$O_{o,o,k}$ คือ ข้อมูลเอาท์พุทของโหนดที่ k ในชั้นแอบแฝง

11. ปรับค่าน้ำหนักของเส้นเชื่อมโยงโดยการแพร่แบบย้อนกลับของค่าความคลาดเคลื่อน (Backpropagation Error to Adjust Weights) สำหรับการวิจัยครั้งนี้ได้เลือกใช้การเรียนรู้แบบมีผู้สอน ซึ่งการปรับค่าน้ำหนักบนเส้นเชื่อมโยง มีจุดมุ่งหมายเพื่อให้ค่าผลรวมกำลังสองของค่าความคลาดเคลื่อน (SSE) ของชุดข้อมูลที่ใช้ในการฝึกสอนมีค่าต่ำสุด มีรายละเอียดดังสมการ (3-11) ดังนี้

$$I_{o,o,k} = \sum O_{o,o,k} w_{jk} \tag{3-11}$$

จากสมการ (3-9)

เมื่อ $I_{o,o,k}$ คือ ข้อมูลอินพุทของโหนดที่ k ในชั้นแสดงผลลัพธ์

$O_{o,o,k}$ คือ ข้อมูลเอาท์พุทของโหนดที่ k ในชั้นแอบแฝง

$w_{j,k}$ คือ ค่าน้ำหนักบนเส้นเชื่อมโยงจากโหนดที่ j ไปยังโหนดที่ k

การหาอนุพันธ์ของการแพร่แบบย้อนกลับใช้กฎลูกโซ่ (Chain Rule) มีรูปแบบ ดังสมการ (3-12) ดังนี้

$$\delta_{jk} = \frac{\partial SSE}{\partial w_{jk}} \left(\frac{\partial SSE}{\partial O_{o,o,k}} \right) \left(\frac{\partial O_{o,o,k}}{\partial I_{o,o,k}} \right) \left(\frac{\partial I_{o,o,k}}{\partial w_{jk}} \right) \tag{3-12}$$

จากสมการ (3-10)

และ $$SSE = \sum \left(D_k - O_{o,o,k} \right)^2 \tag{3-13}$$

ดังนั้น $$\left(\frac{\partial SSE}{\partial O_{o,o,k}} \right) = -2 \sum \left(D_k - O_{o,o,k} \right) \tag{3-14}$$

จากสมการ (3-9)

ดังนั้น $$\left(\frac{\partial O_{o,o,k}}{\partial I_{o,o,k}} \right) = O_{o,o,k} \left(1 - O_{o,o,k} \right) \tag{3-15}$$

จากสมการ (3-8)

ดังนั้น $$\left(\frac{\partial I_{o,o,k}}{\partial w_{jk}}\right) = O_{o,j,o} \quad (3\text{-}16)$$

จากสมการ (3-12)

สามารถเขียนรูปสมการใหม่ ได้ดังสมการ (3-17) ดังนี้

$$\delta_{jk} = 2e_k O_{o,o,k}\left(1 - O_{o,o,k}\right)O_{o,j,o} \quad (3\text{-}17)$$

จากสมการข้างต้น จะได้สมการที่ใช้ในการปรับค่าน้ำหนัก ดังสมการ (3-18) และ (3-19) ดังนี้

$$\Delta w_{jk}^{(new)} = \eta \delta_{jk} O_{o,o,k} + \alpha \Delta w_{jk}^{(old)} \quad (3\text{-}18)$$

และ $$w_{jk}^{(new)} = w_{jk}^{(old)} + \Delta w_{jk}^{(new)} \quad (3\text{-}19)$$

เมื่อ α คือ ค่าโมเมนตัม เป็นค่าคงที่ที่ใช้ในการปรับค่าน้ำหนัก มีค่าอยู่ระหว่าง 0 ถึง 1 โมเมนตัมจะแสดงถึงสัดส่วนของน้ำหนักครั้งก่อนที่นำมาใช้ในการปรับน้ำหนักครั้งปัจจุบัน

η คือ สัมประสิทธิ์การเรียนรู้ มีค่าอยู่ระหว่าง 0 ถึง 1

สำหรับการปรับน้ำหนักบนเส้นเชื่อมโยงระหว่างชั้นข้อมูลป้อนเข้ากับชั้นแอบแฝง ค่าความคลาดเคลื่อนของค่าเอาท์พุทในชั้นแอบแฝง หาได้จากความสัมพันธ์ของความคลาดเคลื่อนในชั้นแสดงผลลัพธ์ แสดงดังสมการ (3-20) ดังนี้

$$e_i = \sum w_{jk} e_k \quad (3\text{-}20)$$

เมื่อคำนวณค่าความคลาดเคลื่อนของเอาท์พุทในชั้นแอบแฝงได้แล้ว จะสามารถหาค่าปรับน้ำหนักและคำนวณค่าน้ำหนักใหม่ได้จากสมการ (3-21) ดังนี้

$$\delta_{ij} = 2e_j O_{o,j,o}\left(1 - O_{o,j,o}\right)O_{o,j,o} \quad (3\text{-}21)$$

สมการที่ใช้ในการปรับค่าน้ำหนัก คือ

$$\Delta w_{ij}^{(new)} = \eta \delta_{ij} O_{o,j,o} + \alpha \Delta w_{ij}^{(new)} \quad (3\text{-}22)$$

และ

$$w_{ij}^{(new)} = w_{ij}^{(old)} + \Delta w_{ij}^{(new)} \quad (3\text{-}23)$$

เมื่อ α คือ ค่าโมเมนตัม เป็นค่าคงที่ที่ใช้ในการปรับค่าน้ำหนัก มีค่าอยู่ระหว่าง 0 ถึง 1 โมเมนตัมจะแสดงถึงสัดส่วนของน้ำหนักครั้งก่อนที่นำมาใช้ในการปรับน้ำหนักครั้งปัจจุบัน

η คือ สัมประสิทธิ์การเรียนรู้ มีค่าอยู่ระหว่าง 0 ถึง 1

12. ทำซ้ำขั้นตอนที่ 4 ถึง 11 จนกว่าข้อมูลจะถูกป้อนเข้าสู่โครงข่ายประสาทเทียมจนหมด ซึ่งจะเรียกการคำนวณที่ใช้ข้อมูลทั้งหมดจนครบรอบว่า “Epoch”

13. คำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : *MSE*) ได้ดังสมการ (3-24) ดังนี้

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n} \quad (3\text{-}24)$$

MSE คือ ค่าประมาณความคลาดเคลื่อน

e_i คือ ค่าความคลาดเคลื่อนของค่าพยากรณ์กับค่าจริง

n คือ จำนวนข้อมูล

กำหนดค่าความคลาดเคลื่อนของค่าผลลัพธ์ (Output) ที่ยอมรับได้

หากค่า *MSE* มีค่าต่ำที่สุดหรืออยู่ในเกณฑ์ที่ตั้งไว้ จะหยุดทำงาน แต่ถ้า *MSE* มากกว่าเกณฑ์ที่กำหนด (Terminate Criteria) ให้ย้อนกลับไปทำขั้นตอนที่ 4 ถึง 13 อีกครั้ง

4.2 การวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม กรณีข้อมูลป้อนเข้า 11 นิวรอล

ผู้วิจัยใช้การวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียมโดยจำนวนข้อมูลป้อนเข้าเหมือนและเท่ากับตัวแปรอิสระที่ใช้การวิเคราะห์การถดถอยพหุคูณแบบขั้นตอน โดยวิธีการและขั้นตอน

เหมือนกับขั้นตอนที่ 4.1 แต่เปลี่ยนจำนวนข้อมูลป้อนเข้าให้เป็นตัวแปรอิสระที่ได้จากการวิเคราะห์ถดถอยพหุคูณแบบขั้นตอน จำนวน 11 โหนด ดังนี้

1. การกำหนดค่าเริ่มต้นให้ตัวแบบโครงข่ายประสาทเทียมทำงาน ดังนี้

1.1 แบบของโครงข่ายประสาทเทียม ประกอบด้วย ชั้นข้อมูลป้อนเข้า (Input Node Layer) ชั้นแอบแฝง (Hidden Node Layer) และชั้นแสดงผลลัพธ์ (Output Node Layer) แบ่งการกำหนดค่าในแต่ละชั้น ดังนี้

1.1.1 ชั้นข้อมูลป้อนเข้า (Input Node Layer) ประกอบด้วยจำนวนข้อมูลป้อนเข้า (Input) นิวรอลหรือจำนวนตัวแปรอิสระ 11 ตัวแปร

1.1.2 ชั้นแอบแฝง (Hidden Node Layer) ประกอบด้วยจำนวนนิวรอลเท่ากับ $\sqrt{MN}$ นิวรอล (Masters, 1993) เมื่อ M เป็นจำนวนนิวรอลของชั้นข้อมูลป้อนเข้า ในที่นี้คือ จำนวนตัวแปรอิสระ ซึ่งมีค่าเท่ากับ 11 และ N เป็นจำนวนนิวรอลของชั้นแสดงผลลัพธ์ ซึ่งในที่นี้คือ จำนวนตัวแปรตาม ซึ่งมีค่าเท่ากับ 1 ดังนั้น ชั้นแอบแฝงจะประกอบด้วยนิวรอลจำนวน $\sqrt{MN} = \sqrt{11 \times 1} \cong 3.317$ ซึ่งในที่นี้กำหนดให้มีจำนวน 3 นิวรอล ซึ่งสูตรนี้เป็นการประมาณขนาดชั้นแอบแฝง แต่การกำหนดจำนวนนิวรอลในชั้นแอบแฝง อาจกำหนดได้โดย

- ถ้ามีจำนวนนิวรอลนำเข้าและนิวรอลแสดงผลลัพธ์น้อย ๆ รวมทั้งปัญหามีความซับซ้อน นิวรอลในชั้นแอบแฝงอาจต่ำกว่าความต้องการ
- ถ้าปัญหาเป็นปัญหาที่ง่าย มีจำนวนนิวรอลนำเข้าและนิวรอลแสดงผลลัพธ์มาก จำนวนนิวรอลในชั้นแอบแฝงน้อย ๆ ก็เพียงพอ
- กฎนี้ไม่ควรนำไปใช้กับโครงข่ายที่มีจำนวนนิวรอลนำเข้าเท่ากับจำนวนนิวรอลแสดงผลลัพธ์ (Autoassociative Networks)

1.1.3 ชั้นแสดงผลลัพธ์ (Output Node Layer) ประกอบด้วยจำนวนตัวแปรตามนั่นคือ 1 ตัว (1 นิวรอล) ซึ่งแสดงถึง เกรดเฉลี่ยของนักศึกษาคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ

1.2 โมเดลที่ใช้ในการวิจัย คือ โครงข่ายประสาทเทียมแบบที่มีการเรียนรู้แบบแพร่ย้อนกลับ (Backpropagation Neural Network)

2. ปรับค่าของตัวแปรให้อยู่ในสเกลเดียวกัน ทั้งในส่วนของตัวแปรอินพุทและตัวแปรเอาท์พุทให้มีค่าอยู่ระหว่าง 0 ถึง 1 โดยใช้สมการ (3-5) ดังนี้

$$x_i' = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}}$$

เมื่อ $x_{\max}$ คือ ค่าที่คาดว่าจะเป็นค่าสูงสุดของข้อมูล

$x_{\min}$ คือ ค่าที่คาดว่าจะเป็นค่าต่ำสุดของข้อมูล

3. กำหนดค่าน้ำหนักเริ่มต้นให้กับเส้นเชื่อมโยงระหว่างโหนด โดยการสุ่มหรือการกำหนดขึ้นเอง ให้มีค่าอยู่ระหว่าง -1 ถึง 1

4. ป้อนข้อมูลที่ปรับค่าแล้วเข้าสู่ชั้นข้อมูลป้อนเข้า

5. ส่งข้อมูลจากชั้นข้อมูลป้อนเข้าไปยังโหนดในชั้นแอบแฝง โดยไม่ผ่านขั้นตอนการคำนวณใด ๆ

6. คำนวณผลรวมของสัญญาณอินพุทของแต่ละโหนดในชั้นแอบแฝง ได้จากสมการ (3-6) ดังนี้

$$I_{o,j,o} = \sum_{i=1}^{n} O_{i,o,o} w_{i,j}$$

เมื่อ $I_{o,j,o}$ คือ ข้อมูลอินพุทของโหนดที่ j ในชั้นแอบแฝง

$O_{i,o,o}$ คือ ข้อมูลเอาท์พุทของโหนดที่ i ในชั้นข้อมูลป้อนเข้า

$w_{i,j}$ คือ ค่าน้ำหนักบนเส้นเชื่อมโยงจากโหนดที่ i ไปยังโหนดที่ j

7. คำนวณค่าเอาท์พุทของโหนดที่ j ในชั้นแอบแฝงโดยใช้ฟังก์ชันการแปลง (Transfer Function) คำนวณผลลัพธ์จากสมการ (3-7) ดังนี้

$$O_{o,j,o} = \frac{1}{1 + e^{-I_{o,j,o}}}$$

เมื่อ $O_{o,j,o}$ คือ ข้อมูลเอาท์พุทของโหนดที่ j ในชั้นแอบแฝง

$I_{o,j,o}$ คือ ข้อมูลอินพุทของโหนดที่ j ในชั้นแอบแฝง

8. คำนวณผลรวมสัญญาณอินพุทของแต่ละโหนดในชั้นแสดงผลลัพธ์จากสมการ (3-8) ดังนี้

$$I_{o,o,k} = \sum_{i=1}^{m} O_{o,j,o} w_{j,k}$$

เมื่อ $I_{o,o,k}$ คือ ข้อมูลอินพุทของโหนดที่ k ในชั้นแสดงผลลัพธ์
$O_{o,j,o}$ คือ ข้อมูลเอาท์พุทของโหนดที่ j ในชั้นแอบแฝง
$w_{j,k}$ คือ ค่าน้ำหนักบนเส้นเชื่อมโยงจากโหนดที่ j ไปยังโหนดที่ k

9. คำนวณค่าเอาท์พุทของโหนด k ในชั้นแสดงผลลัพธ์โดยใช้ฟังก์ชันซิกมอยด์ (Sigmoid Function) คำนวณได้จากสมการ (3-9) ดังนี้

$$O_{o,o,k} = \frac{1}{1+e^{-I_{o,o,k}}}$$

เมื่อ $O_{o,o,k}$ คือ ข้อมูลเอาท์พุทของโหนดที่ k ในชั้นแอบแฝง
$I_{o,o,k}$ คือ ข้อมูลอินพุทของโหนดที่ k ในชั้นแสดงผลลัพธ์

10. คำนวณความคลาดเคลื่อนโดยการนำค่าเอาท์พุทของโหนดที่ k ในชั้นแสดงผลลัพธ์มาเปรียบเทียบค่าเป้าหมาย ซึ่งหาได้จากสมการ (3-10) ดังนี้

$$e_k = D_k - O_{o,o,k}$$

เมื่อ e_k คือ ความคลาดเคลื่อนของโหนดที่ k
D_k คือ ข้อมูลจริงหรือค่าเป้าหมายของโหนดที่ k ในชั้นแสดงผลลัพธ์
$O_{o,o,k}$ คือ ข้อมูลเอาท์พุทของโหนดที่ k ในชั้นแอบแฝง

11. ปรับค่าน้ำหนักของเส้นเชื่อมโยงโดยการแพร่แบบย้อนกลับของค่าความคลาดเคลื่อน (Backpropagation Error to Adjust Weights) สำหรับการวิจัยครั้งนี้ได้เลือกใช้การเรียนรู้แบบมีผู้สอน ซึ่งการปรับค่าน้ำหนักบนเส้นเชื่อมโยง มีจุดมุ่งหมายเพื่อให้ค่าผลรวมกำลังสองของค่าความคลาดเคลื่อน (SSE) ของชุดข้อมูลที่ใช้ในการฝึกสอนมีค่าต่ำสุด มีรายละเอียดดังสมการ (3-11) ดังนี้

$$I_{o,o,k} = \sum O_{o,o,k} w_{jk}$$

จากสมการ (3-9)

เมื่อ $I_{o,o,k}$ คือ ข้อมูลอินพุทของโหนดที่ k ในชั้นแสดงผลลัพธ์

$O_{o,o,k}$ คือ ข้อมูลเอาท์พุทของโหนดที่ k ในชั้นแอบแฝง

$w_{j,k}$ คือ ค่าน้ำหนักบนเส้นเชื่อมโยงจากโหนดที่ j ไปยังโหนดที่ k

การหาอนุพันธ์ของการแพร่แบบย้อนกลับใช้กฎลูกโซ่ (Chain Rule) มีรูปแบบดังสมการ (3-12) ดังนี้

$$\delta_{jk} = \frac{\partial SSE}{\partial w_{jk}} \left(\frac{\partial SSE}{\partial O_{o,o,k}} \right) \left(\frac{\partial O_{o,o,k}}{\partial I_{o,o,k}} \right) \left(\frac{\partial I_{o,o,k}}{\partial w_{jk}} \right)$$

จากสมการ (3-10) และสมการ (3-13)

จะได้ดังสมการ (3-14) ดังนี้

$$\left(\frac{\partial SSE}{\partial O_{o,o,k}} \right) = -2 \sum \left(D_k - O_{o,o,k} \right)$$

จากสมการ (3-9)

ดังนั้น จะได้ดังสมการ (3-15) ดังนี้

$$\left(\frac{\partial O_{o,o,k}}{\partial I_{o,o,k}} \right) = O_{o,o,k} \left(1 - O_{o,o,k} \right)$$

จากสมการ (3-8)

ดังนั้น จะได้ดังสมการ (3-16) ดังนี้

$$\left(\frac{\partial I_{o,o,k}}{\partial w_{jk}} \right) = O_{o,j,o}$$

จากสมการ (3-32)

สามารถเขียนรูปสมการใหม่ ได้ดังสมการ (3-17) ดังนี้

$$\delta_{jk} = 2 e_k O_{o,o,k} \left(1 - O_{o,o,k} \right) O_{o,j,o}$$

จากสมการข้างต้น จะได้สมการที่ใช้ในการปรับค่าน้ำหนัก ดังสมการ (3-18) และ (3-19) ดังนี้

$$\Delta w_{jk}^{(new)} = \eta \delta_{jk} O_{o,o,k} + \alpha \Delta w_{jk}^{(old)}$$

และ $$w_{jk}^{(new)} = w_{jk}^{(old)} + \Delta w_{jk}^{(new)}$$

เมื่อ α คือ ค่าโมเมนตัม เป็นค่าคงที่ที่ใช้ในการปรับค่าน้ำหนัก มีค่าอยู่ระหว่าง 0 ถึง 1 โมเมนตัมจะแสดงถึงสัดส่วนของน้ำหนักครั้งก่อนที่นำมาใช้ในการปรับน้ำหนักครั้งปัจจุบัน

η คือ สัมประสิทธิ์การเรียนรู้ มีค่าอยู่ระหว่าง 0 ถึง 1

สำหรับการปรับน้ำหนักบนเส้นเชื่อมโยงระหว่างชั้นข้อมูลป้อนเข้ากับชั้นแอบแฝง ค่าความคลาดเคลื่อนของค่าเอาท์พุทในชั้นแอบแฝง หาได้จากความสัมพันธ์ของความคลาดเคลื่อนในชั้นแสดงผลลัพธ์ แสดงดังสมการ (3-20)

$$e_i = \sum w_{jk} e_k$$

เมื่อคำนวณค่าความคลาดเคลื่อนของเอาท์พุทในชั้นแอบแฝงได้แล้ว จะสามารถหาค่าปรับน้ำหนักและคำนวณค่าน้ำหนักใหม่ได้จากสมการ (3-21) ดังนี้

$$\delta_{ij} = 2e_j O_{o,j,o}\left(1 - O_{o,j,o}\right) O_{o,j,o}$$

สมการที่ใช้ในการปรับค่าน้ำหนัก แสดงดังสมการ (3-22) และ (3-23) ดังนี้

$$\Delta w_{ij}^{(new)} = \eta \delta_{ij} O_{o,j,o} + \alpha \Delta w_{ij}^{(new)}$$

และ

$$w_{ij}^{(new)} = w_{ij}^{(old)} + \Delta w_{ij}^{(new)}$$

เมื่อ α คือ ค่าโมเมนตัม เป็นค่าคงที่ที่ใช้ในการปรับค่าน้ำหนัก มีค่าอยู่ระหว่าง 0 ถึง 1 โมเมนตัมจะแสดงถึงสัดส่วนของน้ำหนักครั้งก่อนที่นำมาใช้ในการปรับน้ำหนักครั้งปัจจุบัน

η คือ สัมประสิทธิ์การเรียนรู้ มีค่าอยู่ระหว่าง 0 ถึง 1

12. ทำซ้ำขั้นตอนตอนที่ 4 ถึง 11 จนกว่าข้อมูลจะถูกป้อนเข้าสู่โครงข่ายประสาทเทียมจนหมด ซึ่งจะเรียกการคำนวณที่ใช้ข้อมูลทั้งหมดจนครบรอบว่า “Epoch”

13. คำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) ได้จากสมการ (3-24) ดังนี้

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n}$$

MSE คือ ค่าประมาณความคลาดเคลื่อน

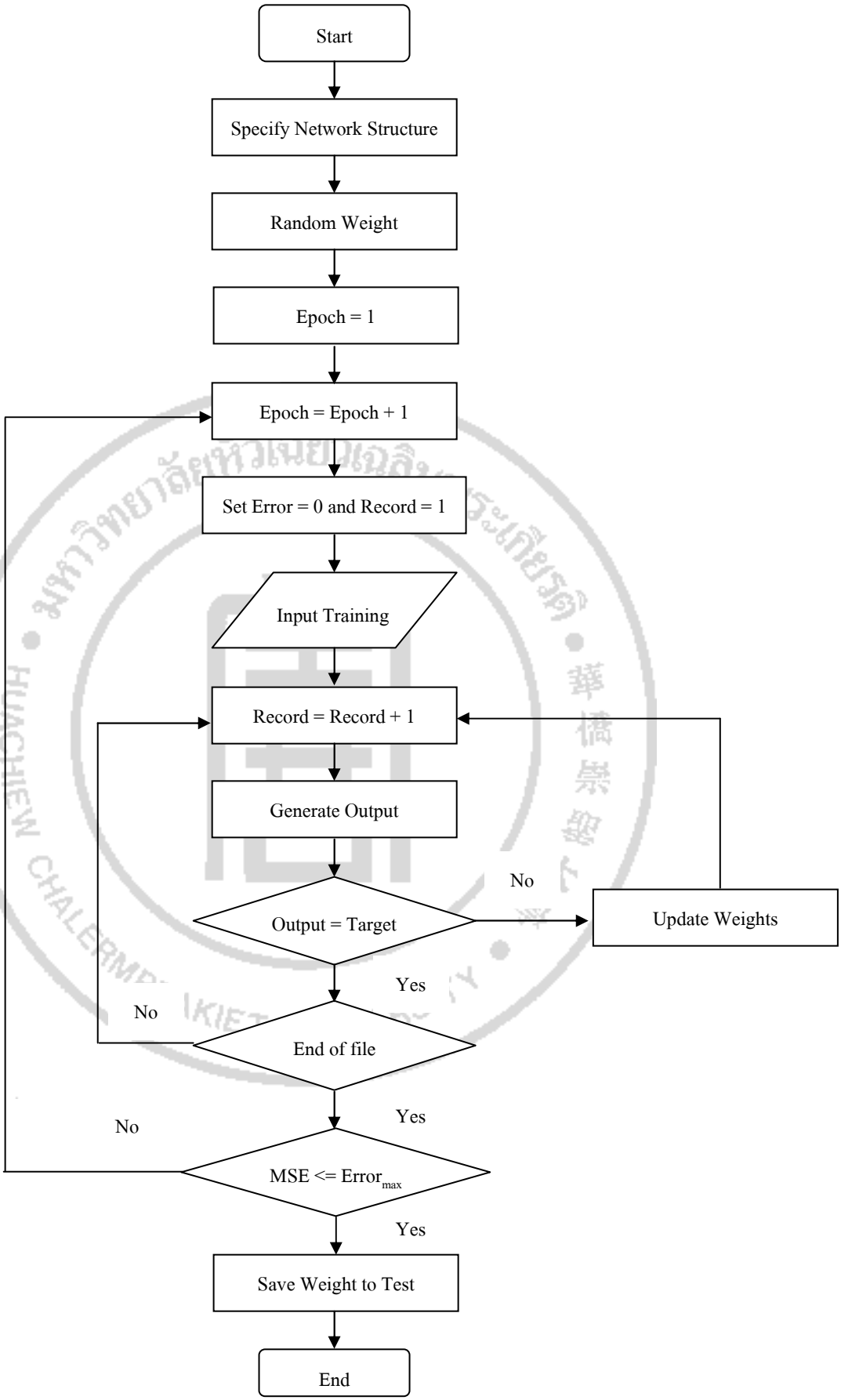
e_i คือ ค่าความคลาดเคลื่อนของค่าพยากรณ์กับค่าจริง

n คือ จำนวนข้อมูล

กำหนดค่าความคลาดเคลื่อนของค่าผลลัพธ์ (Output) ที่ยอมรับได้

หากค่า MSE มีค่าต่ำที่สุดหรืออยู่ในเกณฑ์ที่ตั้งไว้ จะหยุดทำงาน แต่ถ้า MSE มากกว่าเกณฑ์ที่กำหนด (Terminate Criteria) ให้ย้อนกลับไปทำขั้นตอนที่ 4 ถึง 13 อีกครั้ง

สำหรับขั้นตอนการทำงานของโครงข่ายประสาทเทียม สามารถแสดงได้ดังภาพที่ 3-1



ภาพที่ 3-1 ขั้นตอนการทำงานของโครงข่ายประสาทเทียม

ตอนที่ 5 การตรวจสอบผลการพยากรณ์

การวิเคราะห์ในตอนนี้เป็นการนำผลจากการพยากรณ์ด้วยวิธีโครงข่ายประสาทเทียม มาเปรียบเทียบกับผลการพยากรณ์ด้วยวิธีการถดถอยพหุคูณ โดยคำนวณความคลาดเคลื่อนในการพยากรณ์ในแต่ละวิธี ด้วยค่าวัดความคลาดเคลื่อน 2 แบบ ได้แก่ รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error หรือ RMSE) และเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error หรือ MAPE) โดยมีขั้นตอน ดังนี้

ขั้นตอนที่ 1 คำนวณความคลาดเคลื่อนในการพยากรณ์แต่ละวิธีด้วยความคลาดเคลื่อน 2 แบบ ดังนี้

1. รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error หรือ RMSE) หมายถึง ค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดจากค่าความคลาดเคลื่อนกำลังสองระหว่างค่าพยากรณ์และค่าจริง ค่าวัดความคลาดเคลื่อนนี้จะไวต่อความคลาดเคลื่อนที่มีขนาดใหญ่ โดยมีเกณฑ์ที่ใช้ในการพิจารณาความเหมาะสม คือ ค่าความคลาดเคลื่อนที่คำนวณได้ยิ่งต่ำ แสดงว่าวิธีการพยากรณ์นั้นมีความคลาดเคลื่อนน้อย ซึ่งมีสูตรในการคำนวณดังสมการ (1-1) ดังนี้

$$RMSE = \left(\frac{\sum_{n=1}^{n} \left(\hat{Y}_i - Y_i \right)^2}{n} \right)^{\frac{1}{2}}$$

$RMSE$	คือ	รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย
$\hat{Y}_i$	คือ	ค่าพยากรณ์
Y_i	คือ	ค่าจริง
n	คือ	ขนาดของข้อมูล

2. เปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error หรือ MAPE) หมายถึง ค่าวัดความคลาดเคลื่อนของการพยากรณ์ที่วัดความคลาดเคลื่อนของค่าพยากรณ์เทียบกับค่าจริงโดยมีเกณฑ์ที่ใช้ในการพิจารณาความเหมาะสม คือ ค่าความคลาดเคลื่อนที่คำนวณได้ยิ่งต่ำ แสดงว่าวิธีการพยากรณ์นั้นมีความคลาดเคลื่อนน้อย ซึ่งมีสูตรในการคำนวณดังสมการ (1-2) ดังนี้

$$MAPE = \frac{\sum_{n=1}^{n} \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right|}{n} \times 100$$

$MAPE$	คือ	เปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์เฉลี่ย
$\hat{Y}_i$	คือ	ค่าพยากรณ์
Y_i	คือ	ค่าจริง
n	คือ	ขนาดของข้อมูล

ขั้นตอนที่ 2 ภายหลังจากการคำนวณความคลาดเคลื่อนในการพยากรณ์แต่ละวิธีด้วยความคลาดเคลื่อน 2 แบบแล้ว จากนั้นเปรียบเทียบผลการพยากรณ์แต่ละวิธี ถ้าพบว่าวิธีใดมีค่า RMSE และ MAPE น้อยที่สุด แสดงว่า วิธีนั้นเป็นวิธีที่มีความคลาดเคลื่อนในการพยากรณ์น้อยที่สุด

มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ HUACHIEW CHALERMPRAKIET UNIVERSITY 華僑崇聖大學