

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเพื่อพัฒนาระบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบ ซึ่งผู้วิจัยได้กำหนดขั้นตอนและวิธีดำเนินการดังนี้

1. ประชากรกลุ่มเป้าหมาย
2. เครื่องมือที่ใช้ในการวิจัย
3. การเก็บรวบรวมข้อมูล
4. สถิติที่ใช้ในการวิเคราะห์ข้อมูล

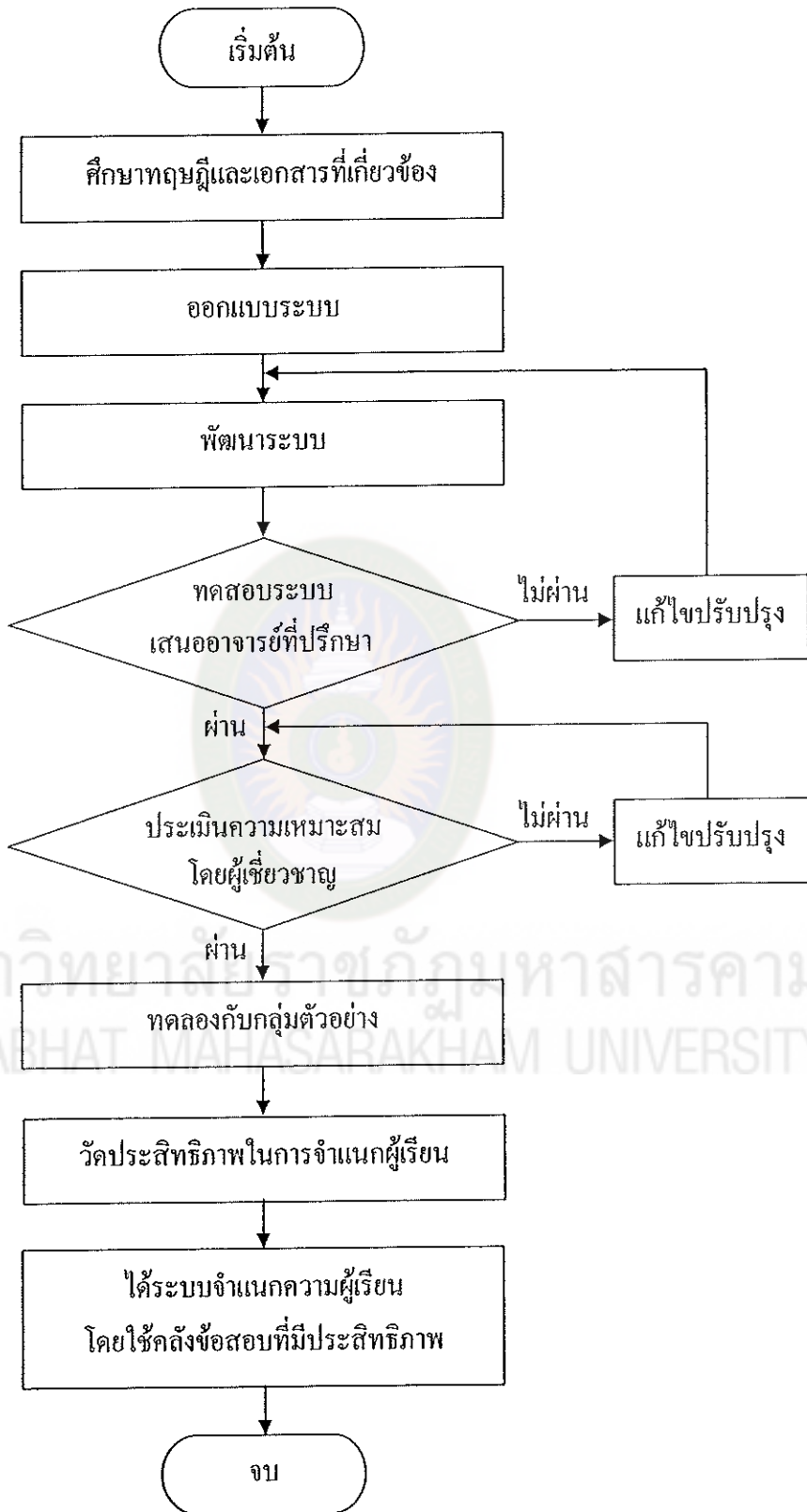
ประชากรกลุ่มเป้าหมาย

ประชากรกลุ่มเป้าหมาย ที่ใช้ในการวิจัยในครั้งนี้คือ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2552 โรงเรียนห้วยผึ้งพิทยาคม อำเภอห้วยผึ้ง จังหวัดกาฬสินธุ์ จำนวน 127 คน

เครื่องมือที่ใช้ในการวิจัย

1. ระบบจำแนกความสามารถผู้เรียน โดยใช้คลังข้อสอบ
2. แบบสอบถามประเมินความเหมาะสมของระบบจำแนกความสามารถผู้เรียน โดยใช้คลังข้อสอบ
3. แบบสอบถามประเมินความพึงพอใจในการใช้งานระบบจำแนกความสามารถผู้เรียน โดยใช้คลังข้อสอบ

ผู้วิจัย ได้พัฒนาระบบการจำแนกความสามารถผู้เรียนโดยใช้คลัง ซึ่งระบบดังกล่าวครูสามารถนำไปทดสอบกับผู้เรียนก่อนการจัดกิจกรรมการเรียนการสอน เพื่อที่ครูจะได้ทราบถึงสารสนเทศเกี่ยวกับผู้เรียน ใช้เป็นข้อมูลประกอบในการจัดกิจกรรมการเรียนการสอน พัฒนาผู้เรียนให้ตรงตามศักยภาพของแต่ละบุคคล ในการดำเนินการวิจัยพัฒนาระบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบนั้น ผู้วิจัยได้พัฒนาตามขั้นตอนดังนี้



ภาพที่ 8 ขั้นตอนการพัฒนาบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบ

ในการพัฒนาระบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบ มีรายละเอียดในแต่ละขั้นตอนดังนี้

1. ศึกษาทฤษฎีและเอกสารที่เกี่ยวข้อง

ในขั้นตอนนี้ผู้วิจัยได้ศึกษาทฤษฎีและเอกสารที่เกี่ยวข้อง ซึ่งประกอบด้วย

- 1.1 ทฤษฎีการตอบสนองข้อสอบ
- 1.2 คลังข้อสอบและการทดสอบแบบปรับเหมาะ
- 1.3 การจัดกลุ่มข้อมูล
- 1.4 งานวิจัยที่เกี่ยวข้อง

จากการศึกษาทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้องทำให้ผู้วิจัยเกิดแนวความคิดในการเพิ่มประสิทธิภาพในการจำแนกผู้เรียนให้สูงขึ้น ผู้วิจัยได้ดำเนินการ ออกแบบระบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบ เป็นลำดับขั้นตอนดังนี้

2. ออกแบบระบบ

ในขั้นตอนการออกแบบระบบนี้จะเป็นขั้นตอนที่รวมถึงการกำหนดปัญหาของระบบงานเดิม การวิเคราะห์ระบบ และการออกแบบระบบ

2.1 การกำหนดปัญหาของระบบงานเดิม (Problem definition)

จากการศึกษาระบบงานเดิม พบว่าการพัฒนาโปรแกรมสำหรับทดสอบผู้เรียนแบบปรับเหมาะด้วยคอมพิวเตอร์ส่วนใหญ่พัฒนาด้วยยุทธวิธีหลายขั้นตอนแบบแยกทางแปรผัน (Variable Branching) โดยใช้กลวิธีความเป็นไปได้สูงสุด ผู้วิจัยพบว่าการใช้กลวิธีความเป็นไปได้สูงสุดนั้นมีข้อจำกัดในกรณีที่ผู้เรียนตอบข้อสอบถูกหรือผิดหมดทุกข้อ จะไม่สามารถประมาณค่าความสามารถของผู้เรียนคนนั้นได้ จะต้องนำผลการตอบของผู้เรียนคนนั้นออกจากระบบก่อน จึงจะสามารถคำนวณได้ และอีกประการหนึ่ง คือการพัฒนาโปรแกรมสำหรับทดสอบผู้เรียนแบบปรับเหมาะด้วยคอมพิวเตอร์จะทำการพัฒนาเป็นโปรแกรมสำเร็จรูปที่ไม่สามารถทำงานได้ในระบบเครือข่ายอินเทอร์เน็ต ซึ่งการพัฒนาโปรแกรมต่างๆ ในปัจจุบันจะนิยมพัฒนาโปรแกรมให้สามารถทำงานบนระบบเครือข่ายอินเทอร์เน็ต ทำให้ผู้ใช้งานสามารถใช้งานได้อย่างสะดวก ใช้งานได้ทุกที่ตามอรรถาศัย

ดังนั้น ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาระบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบ ให้สามารถทำงานได้บนระบบเครือข่ายอินเทอร์เน็ตและเพิ่มประสิทธิภาพการจำแนกผู้เรียนโดยใช้คลังข้อสอบให้สูงขึ้น ด้วยวิธีการนำค่าความสามารถของผู้เรียนที่ผ่านการ

ทดสอบแบบปรับเหมาะด้วยคอมพิวเตอร์ นำไปจัดกลุ่มด้วยเทคนิคและวิธีการอัลกอริทึมที่มั่นคง ซึ่งเป็นวิธีการจัดกลุ่มข้อมูลที่ได้รับการยอมรับและมีประสิทธิภาพในปัจจุบัน

2.2 การวิเคราะห์ระบบ (System analysis)

ในขั้นตอนนี้ผู้วิจัยได้ทำการวิเคราะห์ระบบ โดยผู้วิจัยทำการรวบรวมความต้องการ (Requirements) ต่างๆ เพื่อนำข้อมูลมากำหนดความต้องการของระบบใหม่ ทำการวิเคราะห์ความต้องการเพื่อสรุปเป็นข้อกำหนด สร้างแผนภาพการไหลของข้อมูล(Data Flow Diagram) เพื่อแสดงขั้นตอนการทำงานทั้งหมดที่มีอยู่ในระบบ

2.3 การออกแบบระบบ (System design)

ในขั้นตอนนี้จะเป็นการนำผลลัพธ์ที่ได้จากการวิเคราะห์ระบบมาออกแบบการทำงานของระบบ ออกแบบขั้นตอนการเข้าสอบ ออกแบบฐานข้อมูล เขียนแผนภาพภาพอีอาร์ เพื่อแสดงความสัมพันธ์ของข้อมูล ทำการแปลงแผนภาพอีอาร์ให้เป็นรีเลชันโดยปรับโครงสร้างข้อมูลให้อยู่ในบรรทัดฐานเพื่อให้โครงสร้างข้อมูลไม่มีความซ้ำซ้อนของข้อมูลที่เกิดขึ้น ทำการออกแบบโครงสร้างตารางข้อมูล ออกแบบหน้าจอติดต่อผู้ใช้งาน ออกแบบการแสดงผลข้อมูลและการรายงานผลข้อมูล เป็นต้น

3. การพัฒนาโปรแกรม (Development)

จากขั้นตอนการวิเคราะห์และออกแบบระบบ ทำให้ทราบถึงกระบวนการทำงานของระบบจำแนกความสามารถผู้เรียนโดยใช้คลังข้อสอบ สำหรับขั้นตอนการพัฒนา ระบบเป็นขั้นตอนการเขียนโปรแกรม โดยเบื้องต้นได้ติดตั้งโปรแกรม Appserv-win32-2.5.5 ซึ่งเป็นโปรแกรมที่รวบรวมโอเพ่นซอร์สซอฟต์แวร์หลายๆ อย่างเข้าด้วยกันไว้ในแพ็คเกจนี้ทั้งหมด ประกอบด้วยโปรแกรม Apache Web Server เวอร์ชัน 1.3.34 โปรแกรมแปลภาษาสำหรับการพัฒนา Application PHP Script Language เวอร์ชัน 5 โปรแกรมฐานข้อมูล MySQL Database เวอร์ชัน 5.0 และ โปรแกรมจัดการฐานข้อมูล phpMyAdmin Database Manger Version 2.5.7-p11 ซึ่งผู้วิจัยได้เลือกภาษา PHP ในการเขียนโปรแกรมเนื่องจากมีความยืดหยุ่นและง่ายต่อการใช้คำสั่ง และใช้ระบบจัดการฐานข้อมูลด้วย MySQL Database เวอร์ชัน 5.0.16 เนื่องจากเป็นระบบจัดการฐานข้อมูลที่แพร่หลายและสามารถใช้งานบนเครือข่ายได้ค่อนข้างสมบูรณ์ ทำการพัฒนาโปรแกรมตามขั้นตอนการทำงาน โดยองค์ประกอบที่เป็นหัวใจของการคำนวณตามวิธีการประมาณค่าความสามารถของผู้สอบนี้ใช้วิธีการของเบย์ ซึ่งนักสถิติชื่อ โรเจอร์ เจ โอเวน (Roger J. Owen) เป็นผู้เสนอขึ้น เพื่อใช้ในการประมาณค่าความสามารถของผู้สอบ ที่ปรับใหม่ เบส์เซียน อัปเดตติ้ง (Bayesian Updating) ซึ่งสามารถประมาณค่า

ความสามารถของผู้สอบได้ค่อนข้างคงที่ (Owen, 1969, 1975, pp. 351-356; Thorndike, 1982, p. 303; citing Owen, 1975, unpagged อ้างในต่ายเซ็งซ่ง, 2534 : 50-52) และพัฒนาโปรแกรม ในส่วนของการจัดกลุ่มผู้เรียนตามอัลกอริทึมเคมีน (K-means Algorithm) มีขั้นตอนดังนี้

3.1 ทำการสุ่มข้อมูล จากกลุ่มข้อมูลที่ได้จากการประมาณค่าความสามารถ ของผู้เรียน จำนวน 3 กลุ่มข้อมูล (k) เพื่อกำหนดเป็นจุดเริ่มต้นของแต่ละกลุ่มข้อมูล

3.2 แบ่งกลุ่มข้อมูลออกเป็น 3 กลุ่ม

3.3 คำนวณหาระยะทางระหว่างข้อมูลทุกตัว กับ ตัวแทนข้อมูลที่สุ่มขึ้นมา จำนวน 3 กลุ่มข้อมูล(k) ด้วยสมการ

$$d(i,j) = \sqrt{\left(|X_{i1} - X_{j1}|^2 + |X_{i2} - X_{j2}|^2 + \dots + |X_{ip} - X_{jp}|^2 \right)}$$

3.4 ตรวจสอบระยะทางระหว่างข้อมูลกับกลุ่มข้อมูล (k) ข้อมูลไหนมีระยะห่าง น้อยที่สุด นำข้อมูลนั้นไปสังกัดในกลุ่มข้อมูล(k) นั้น จะได้กลุ่มข้อมูลขึ้นมาใหม่

3.5 หาจุดศูนย์กลางของแต่ละกลุ่มข้อมูล (Centroid) ด้วยสมการ

$$M_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ik}$$

3.6 นำข้อมูลทุกตัวในแต่ละกลุ่มข้อมูล ไปคำนวณกับจุดศูนย์กลาง (Centroid) ของทุกกลุ่มข้อมูล

3.7 ตรวจสอบระยะทาง ระหว่างข้อมูลกับจุดศูนย์กลางข้อมูล (Centroid) โดย ข้อมูลไหนมีระยะทางที่สั้นที่สุด นำข้อมูลนั้นไปสังกัดกับกลุ่มข้อมูลดังกล่าว

3.8 ทำการคำนวณหาจุดศูนย์กลางของข้อมูล และกำหนดข้อมูลให้กลุ่ม ข้อมูลใหม่ไปเรื่อยๆ (ทำซ้ำในข้อที่ 3.5 ถึง 3.7) โดยมีเงื่อนไขให้หยุดดังนี้

3.8.1 ข้อมูลในแต่ละกลุ่มถึงสถานะเสถียร ไม่สามารถเปลี่ยนกลุ่มได้ ดีกว่านี้ได้

3.8.2 ค่าผิดพลาดที่ต่ำสุด (minimum square error) สำหรับทุกกลุ่มข้อมูล

โดยหาค่าผิดพลาดที่ต่ำสุดได้จากสมการ $E_k^2 = \sum_{k=1}^k e_k^2$ โดยที่ $e_k^2 = \sum_{i=1}^{n_k} (x_{ik} - M_k)^2$

3.9 สิ้นสุดขั้นตอนการพัฒนาโปรแกรม

4. การทดสอบระบบ (Testing)

เป็นขั้นตอนของการทดสอบระบบก่อนนำไปใช้งานจริง ซึ่งเป็นการทดสอบความคลาดเคลื่อนในการทำงานของโปรแกรม โดยทำการเปรียบเทียบกันระหว่างโปรแกรมที่เขียนไว้กับการคำนวณด้วยมือ และให้ผู้เชี่ยวชาญทดลองใช้งาน ถ้าพบข้อผิดพลาดจะย้อนกลับไปพัฒนาโปรแกรมใหม่ ซึ่งในการทดสอบโปรแกรมนั้น จะใช้วิธีการแบบแบล็กบ็อกซ์ (Blackbox testing) เป็นกระบวนการทดสอบการทำงานของระบบโดยรวมทั้งหมดว่ามีกระบวนการทำงานถูกต้องตามวัตถุประสงค์ที่ต้องการหรือไม่ ทำการทดสอบการทำงานในแต่ละหน้าที่การทำงานทั้งหมดเพื่อหาข้อบกพร่องของโปรแกรม จากนั้นทำการปรับปรุงแก้ไขโปรแกรมให้ดีขึ้น

4.1 ประเมินความเหมาะสมของโปรแกรม โดยมีผู้เชี่ยวชาญ 3 ท่าน

ประกอบด้วย

1) นายอาทิตย์ กลีบรัง ครูประจำแผนกวิชาเทคโนโลยีสารสนเทศ วิทยาลัยอาชีวศึกษาธนบุรี จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาในระดับปริญญาโท (วท.ม.) สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น ผู้เชี่ยวชาญด้านวิเคราะห์และออกแบบระบบ

2) นายคุณเทพ ภัทรโกศล นักวิชาการคอมพิวเตอร์ สำเร็จการศึกษาในระดับปริญญาโท (วท.ม.) สาขาวิชาเทคโนโลยีคอมพิวเตอร์เพื่อการศึกษา มหาวิทยาลัยราชภัฏมหาสารคาม ผู้เชี่ยวชาญด้านประเมินระบบ

3) นายถวัลย์ ประชาชัย ครูประจำแผนกวิชาช่างยนต์ วิทยาลัยสารพัดช่างจังหวัดมหาสารคาม สำเร็จการศึกษาในระดับปริญญาโท (กศม.) สาขาวิชาวัดผลและประเมินผลการศึกษา มหาวิทยาลัยมหาสารคาม ผู้เชี่ยวชาญด้านรูปแบบ

เครื่องมือที่ใช้ในการประเมินความเหมาะสมเป็นแบบสอบถามประเมินความเหมาะสมของระบบ แบ่งออกเป็น 5 ด้าน คือ (อ้างใน รัญญรัตน์ น้อยพลกรัง, 2548 : 105-109)

4.1.1 การประเมินความเหมาะสมของหน้าจอที่ใช้ติดต่อกับผู้ใช้

4.1.2 ประเมินองค์ประกอบการทำงานโดยรวมของระบบ

4.1.3 การประเมินความสะดวกในการใช้งานของระบบ

4.1.4 การประเมินประสิทธิภาพในการทำงานของระบบ

4.1.5 ประเมินประสิทธิภาพการทำงานในแต่ละส่วนของระบบ

เกณฑ์การแปลความหมายจากการวิเคราะห์แบบสอบถามประเมินความเหมาะสมของระบบ เพื่อคำนวณหาค่าเฉลี่ยของผู้เชี่ยวชาญที่มีต่อระบบ จำนวน 3 คน ซึ่ง

ประมาณค่าเฉลี่ยโดยใช้เกณฑ์ประมาณค่าความคิดเห็นตามแนวคิดของบุญชม ศรีสะอาด (2545 : 103) ดังนี้

ค่าเฉลี่ย	แปลความว่า	ระดับความคิดเห็น
4.51 – 5.00	แปลความว่า	ดีมาก
3.51 – 4.50	แปลความว่า	ดี
2.51 – 3.50	แปลความว่า	ปานกลาง
1.51 – 2.50	แปลความว่า	พอใช้
1.00 – 1.50	แปลความว่า	ควรปรับปรุง

เกณฑ์เฉลี่ยของระดับความคิดเห็นของผู้เชี่ยวชาญในงานวิจัยนี้ ใช้ค่าเฉลี่ยของคะแนนตั้งแต่ 3.51 ขึ้นไป และค่าความเบี่ยงเบนมาตรฐานไม่เกิน 1.00

5. การติดตั้งระบบ (Implementation)

เป็นขั้นตอนหลังจากที่ได้ทำการทดลองโปรแกรม และประเมินความเหมาะสมของระบบเป็นที่เรียบร้อยแล้ว โปรแกรมห่างานได้ตามต้องการ ดำเนินการติดตั้งโปรแกรมบนระบบเครือข่ายอินเทอร์เน็ต ติดตั้งในระบบปฏิบัติการ FreeBSD 6.2 ที่ <http://202.143.136.202/irt> เพื่อทำการทดลองจริงกับประชากรกลุ่มเป้าหมายที่ได้คัดเลือกไว้

6. การประเมินผล (Evaluation)

เป็นขั้นตอนสุดท้ายของการพัฒนาโปรแกรม ซึ่งต้องมีการประเมินผลเพื่อให้ทราบถึงความพึงพอใจของผู้ใช้งาน

เครื่องมือที่ใช้ในการประเมินผล คือแบบสอบถามประเมินความพึงพอใจของผู้ใช้งาน แบ่งออกเป็น 2 ด้าน คือ (อ้างใน ภาณุพงษ์ ชัยศรีทิพย์. 2549 : 107-110)

6.1 ด้านการจัดการนำเสนอข้อสอบ

6.2 ด้านลักษณะต่างๆ ไป

เกณฑ์การแปลความหมายจากการวิเคราะห์แบบสอบถามความพึงพอใจของผู้ใช้งานที่มีต่อระบบ เพื่อคำนวณค่าเฉลี่ยของความคิดเห็นต่อระบบ ซึ่งประมาณค่าเฉลี่ยโดยใช้เกณฑ์ประมาณค่าความคิดเห็นตามแนวคิดของบุญชม ศรีสะอาด(2545 : 103) ดังนี้

ค่าเฉลี่ย	แปลความว่า	ระดับความคิดเห็น
4.51 – 5.00	แปลความว่า	พึงพอใจระดับมากที่สุด
3.51 – 4.50	แปลความว่า	พึงพอใจระดับมาก
2.51 – 3.50	แปลความว่า	พึงพอใจระดับปานกลาง

1.51 – 2.50	แปลความว่า	พึงพอใจระดับน้อย
1.00 – 1.50	แปลความว่า	พึงพอใจระดับน้อยที่สุด

การจำแนกระดับความสามารถของผู้เรียนที่ทำการสอบว่าอยู่ระดับ เก่ง ปานกลาง หรืออ่อน นั้นมีเกณฑ์การตัดสิน ดังนี้ (โสฬส สุขานนท์สวัสดิ์. 2545 : 60)

ค่าความสามารถ	+1.50	ถึง	+3.00	หมายถึง	เก่ง
ค่าความสามารถ	-1.50	ถึง	+1.49	หมายถึง	ปานกลาง
ค่าความสามารถ	-3.00	ถึง	-1.49	หมายถึง	อ่อน

การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลตามรายละเอียด ดังนี้

1. นำหนังสือขออนุญาตเข้าทำการทดลองกับกลุ่มประชากรที่เป็นนักเรียนในระดับชั้นมัธยมศึกษา ปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2552 โรงเรียนห้วยผึ้งพิทยาคม อำเภอห้วยผึ้ง จังหวัดกาฬสินธุ์ จำนวน 127 คน

2. ทำการชี้แจงเกี่ยวกับโปรแกรม ขั้นตอนการใช้งาน วิธีการใช้งานต่างๆ ในภาพรวม จนเข้าใจเป็นอย่างดีแล้ว จึงเริ่มทดสอบด้วยแบบทดสอบวัดความสามารถในรายวิชาคณิตศาสตร์ (ค102) ชั้นมัธยมศึกษาปีที่ 1 เรื่องสมการ จุดประสงค์การเรียนรู้ข้อที่ 1 นักเรียนสามารถแก้สมการและตรวจคำตอบได้ จำนวน 107 ข้อ ซึ่งเป็นข้อสอบมาตรฐานที่ผ่านการวิเคราะห์ค่าตัวแปรเสริม ตามทฤษฎีตอบสนองข้อสอบ คลังข้อสอบ ผลการตอบของผู้เรียนที่ทำการตอบแต่ละคนถูกบันทึกไว้ในระบบ ที่พัฒนาขึ้น โดยอัคร โนมัตติ

3. เมื่อทำการทดสอบเป็นที่เรียบร้อยแล้ว ผู้วิจัยได้ทำการแจกแบบสอบถามประเมินความพึงพอใจในการใช้งานระบบจำแนกความสามารถผู้เรียน โดยใช้คลังข้อสอบ

4. รวบรวมข้อมูลที่ได้จากแบบสอบถาม นำมาวิเคราะห์ข้อมูล ตามค่าสถิติที่ใช้ในการวิเคราะห์ข้อมูล

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. ค่าเฉลี่ย (Arithmetic Mean)

โดยใช้สูตรในการคำนวณ ดังนี้ (ธานินทร์ ศิลป์จารุ. 2549 : 153)

$$\bar{X} = \frac{\sum X}{N}$$

เมื่อ	\bar{X}	คือ	ค่าเฉลี่ย
	$\sum X$	คือ	ผลรวมของข้อมูลทั้งหมด
	N	คือ	จำนวนข้อมูลทั้งหมด

2. ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

โดยใช้สูตรในการคำนวณดังนี้ (ธานินทร์ ศิลป์จารุ. 2549 : 167)

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}}$$

เมื่อ	σ	คือ	ค่าเบี่ยงเบนมาตรฐาน
	X	คือ	ข้อมูลแต่ละจำนวน
	μ	คือ	ค่าเฉลี่ย (Mean) ของข้อมูลในชุดนั้น
	N	คือ	จำนวนข้อมูลจากประชากรทั้งหมด

3. ค่าความแปรปรวน (Variance)

ค่าความแปรปรวน หมายถึง การพิจารณาค่าความแปรปรวนในการจำแนกเพื่อจัดกลุ่มผู้เรียนด้วยวิธีการจำแนกผู้เรียน โดยใช้เกณฑ์ และวิธีการจำแนกผู้เรียน โดยใช้อัลกอริทึมเคมีน ในการจำแนกผู้เรียนตามระดับความสามารถใช้สูตรในการคำนวณดังนี้

(วามิณี นิยาภาศ. 2549 : 29)

$$\sigma^2 = \frac{\sum f_i(x_i - \mu)^2}{N} \quad \text{หรือ} \quad \sigma^2 = \left[\frac{\sum f_i x_i^2 - (\sum f_i x_i)^2}{N} \right] \frac{1}{N}$$

4. การวัดประสิทธิภาพการแบ่งกลุ่มข้อมูล (Root Mean Square Standard Deviation : RMSSTD)

การวัดค่าความแตกต่างของข้อมูลภายในกลุ่ม โดยค่าความแตกต่างของข้อมูลภายในกลุ่มนั้นแสดงให้เห็นถึงประสิทธิภาพการจัดกลุ่มว่ามีมากน้อยเพียงใด ซึ่งหากค่าความแตกต่างภายในกลุ่มน้อย นั้นย่อมหมายถึงการแบ่งกลุ่มที่ดีใกล้เคียงกันมาก สูตรการคำนวณค่าความแตกต่างภายในกลุ่ม มีดังนี้ (Jain et al., 1999 อ้างอิงใน วามิณี นิยาการศ. 2549 : 31)

$$\text{RMSSTD} = \sqrt{\frac{\sum_{j=1..d}^{i=1..c} \sum_{k=1}^{n_{ij}} (x_k - \bar{x}_j)^2}{\sum_{j=1..d}^{i=1..c} (n_{ij} - 1)}}$$

โดยที่

- c คือ จำนวนกลุ่มที่แบ่งได้ทั้งหมด
- d คือ จำนวนคอลัมน์ทั้งหมดภายในชุดข้อมูล
- \bar{x}_j คือ ค่าเฉลี่ยของข้อมูลคอลัมน์ที่ j
- n_{ij} คือ จำนวนข้อมูลในกลุ่มที่ i คอลัมน์ที่ j

5. R-Squared (RS)

การวัดค่าความแตกต่างของข้อมูลระหว่างกลุ่ม ค่าความแตกต่างของข้อมูลระหว่างกลุ่ม เป็นค่าซึ่งแสดงให้เห็นถึงประสิทธิภาพการจัดกลุ่มข้อมูล ซึ่งหากค่าความแตกต่างระหว่างกลุ่มมาก แสดงถึงการแบ่งกลุ่มที่ดี กลุ่มแต่ละกลุ่มมีลักษณะที่แตกต่างกัน โดยที่ค่า RS นี้จะอยู่ระหว่าง 0 ถึง 1 สูตรการคำนวณค่าความแตกต่างระหว่างกลุ่ม มีดังนี้ (Jain et al., 1999 อ้างอิงใน วามิณี นิยาการศ. 2549 : 31-32)

$$\text{RS} = \frac{\text{SS}_t - \text{SS}_w}{\text{SS}_t}$$

$$\text{SS}_t = \sum_{j=1}^d \sum_{k=1}^{n_j} (x_k - \bar{x}_j)^2 \quad \text{และ} \quad \text{SS}_w = \sum_{j=1..d}^{i=1..c} \sum_{k=1}^{n_{ij}} (x_k - \bar{x}_j)^2$$

โดยที่

- SS_t คือ ผลรวมของผลต่างกำลังสองของทุกข้อมูลแต่ละตัวแปร
- SS_w คือ ผลรวมของผลต่างกำลังสองทุกข้อมูลภายในกลุ่ม
- c คือ จำนวนกลุ่มที่แบ่งได้ทั้งหมด
- d คือ จำนวนตัวแปรทั้งหมด
- \bar{x}_j คือ ค่าเฉลี่ยของข้อมูลคอลัมน์ที่ j



มหาวิทยาลัยราชภัฏมหาสารคาม
RAJABHAT MAHASARAKHAM UNIVERSITY