ระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้

ณรงค์ศักดิ์ โยตมาน

ปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชานวัตกรรมเทคโนโลยีดิจิทัล
มหาวิทยาลัยแม่โจ้
พ.ศ. 2566

ระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้

ณรงค์ศักดิ์ โยตมาน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของความสมบูรณ์ของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชานวัตกรรมเทคโนโลยีดิจิทัล
สำนักบริหารและพัฒนาวิชาการ มหาวิทยาลัยแม่โจ้
พ.ศ. 2566

ลิขสิทธิ์ของมหาวิทยาลัยแม่โจ้

ระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้

ณรงค์ศักดิ์ โยตมาน

วิทยานิพนธ์นี้ได้รับการพิจารณาอนุมัติให้เป็นส่วนหนึ่งของความสมบูรณ์ของการศึกษา
ตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชานวัตกรรมเทคโนโลยีดิจิทัล

พิจารณาเห็นชอบโดย **อาจารย์ที่ปรึกษา**

อาจารย์ที่ปรึกษาหลัก .....
(ผู้ช่วยศาสตราจารย์ ดร.พาสน์ ปราโมกข์ชน)
วันที่.....เดือน.....พ.ศ. .....

อาจารย์ที่ปรึกษาร่วม .....
(ผู้ช่วยศาสตราจารย์ ดร.ปวีณ เขื่อนแก้ว)
วันที่.....เดือน.....พ.ศ. .....

อาจารย์ที่ปรึกษาร่วม .....
(อาจารย์ ดร.กิติศักดิ์ โอสถานันต์กุล)
วันที่.....เดือน.....พ.ศ. .....

ประธานอาจารย์ผู้รับผิดชอบหลักสูตร .....
(อาจารย์ ดร.กิตติกร หาญตระกูล)
วันที่.....เดือน.....พ.ศ. .....

สำนักบริหารและพัฒนาวิชาการรับรองแล้ว .....
(รองศาสตราจารย์ ดร.ญาณิน โอภาสพัฒนกิจ)
รองอธิการบดี
วันที่.....เดือน.....พ.ศ. .....

ชื่อเรื่อง ระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้
ชื่อผู้เขียน นายณรงค์ศักดิ์ โยตมาน
ชื่อปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชานวัตกรรมเทคโนโลยีดิจิทัล
อาจารย์ที่ปรึกษาหลัก ผู้ช่วยศาสตราจารย์ ดร.พาสน์ ปราโมกข์ชน

บทคัดย่อ

ระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้เป็นวิธีการช่วยอาจารย์ผู้สอนเตรียมวางแผนการสอน ปรับปรุงสื่อการสอน กลยุทธ์การสอน และวิธีการประเมินที่เหมาะสมต่อความต้องการของนักเรียน โดยใช้การประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) ของนักศึกษาปริญญาตรีชั้นปีที่ 1 - 4 และนักเรียนชั้นมัธยมศึกษาปีที่ 3 เพื่อทดสอบโมเดลการแบ่งกลุ่มอย่างเหมาะสม นำไปช่วยปรับแต่งจำนวนกลุ่มของนักเรียนในชั้นเรียนให้เหมาะสมกับสื่อการสอน กลยุทธ์การสอน และวิธีการประเมินที่เหมาะสมกับทรัพยากรด้านการศึกษาในชั้นเรียน

การวิจัยพบว่าการปรับแต่งพารามิเตอร์ของโมเดลเพื่อเพิ่มประสิทธิภาพในการแบ่งกลุ่มตามรูปแบบการเรียนรู้ และลดกระบวนการทำซ้ำของชุดข้อมูล ช่วยลดระยะเวลาในการแบ่งกลุ่มชุดข้อมูล โดยพบว่า Euclidean Distance และ Ward Linkage เป็นอัลกอริทึมและพารามิเตอร์ที่เหมาะสมที่สุด โดยแนะนำให้แบ่งกลุ่มนักเรียนตามรูปแบบการเรียนรู้ของนักเรียนเป็น 6 - 8 กลุ่ม ซึ่งให้ค่า %Change ที่เข้าเงื่อนไขและเหมาะสมกับการแบ่งกลุ่มข้อมูล จากการทดลองกับชุดข้อมูลรูปแบบการเรียนรู้ของนักเรียน 4 ชุดข้อมูล โดยมีค่า SSE น้อยกว่าพารามิเตอร์อื่น ๆ

โมเดลการแบ่งกลุ่มที่ได้รับการปรับแต่งพารามิเตอร์เพื่อเรียนรู้รูปแบบการเรียนของนักเรียน ได้ผลลัพธ์ที่มีประสิทธิภาพในการแบ่งกลุ่มชุดข้อมูลทางการศึกษา โดยอัลกอริทึมและพารามิเตอร์ที่เหมาะสมคือ Euclidean Distance และ Ward Linkage โดยจำนวนกลุ่มที่เหมาะสมสำหรับการแนะนำในการแบ่งกลุ่มคือ 6 - 8 กลุ่ม โมเดลนี้สามารถนำมาปรับใช้กับข้อมูลรูปแบบการเรียนรู้อื่น ๆ และช่วยพัฒนากระบบสนับสนุนการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ได้อย่างเหมาะสม ด้วยวิธีการหาพารามิเตอร์ที่เหมาะสมกับชุดข้อมูล และการวัดประสิทธิภาพตามข้อกำหนดที่เหมาะสมกับแต่ละรูปแบบของข้อมูลนั้น ๆ ผลการวิจัยนี้สามารถนำไปปรับใช้งานในการพัฒนาโมเดลการแบ่งกลุ่มต่าง ๆ ในอนาคตได้อย่างมีประสิทธิภาพ และเหมาะสม

คำสำคัญ : รูปแบบการเรียนรู้, การแบ่งกลุ่มนักเรียน, การทำเหมืองข้อมูล, การพัฒนาการศึกษา

Title TEACHING SUPPORT SYSTEM BY CLUSTERING STUDENTS ACCORDING TO LEARNING STYLES
Author Mr. Narongsak Yotaman
Degree Master of Science in Digital Technology Innovation
Advisory Committee Chairperson Assistant Professor Dr. Part Pramokchon

ABSTRACT

The teaching support system of grouping students according to their learning styles is a method to contribute teachers to prepare lesson plans, improve instructional media, strategies of teaching, and effective assessment. The Index Learning Style (ILS) with students studying a bachelor's degree from year 1 to 4 and student in secondary 3 was used to assess the learning style and test the appropriately stratified model, which could be used to organize the number of groups of students in the class to suit instructional media, strategies of teaching, effective assessment and educational resources.

The research found that it was possible to refine model parameters to increase segmentation efficiency by learning style and reduce the iteration process of the data set. The clustering of the dataset was to Euclidean Distance, and Ward Linkage was the optimal algorithm and parameters. To recommendation to group students according to their learning styles into 6-8 groups, which gives the %Change value that meets the conditions and is suitable for grouping data. Based on experiments with four data sets of student learning style data sets with less SSE than other parameters.

Parameterized segmentation models to learn student learning styles, which effectively results in segmenting educational data sets. The optimal algorithm and parameters were Euclidean Distance and Ward Linkage. The optimal number of clusters for recommendation in clustering was 6 to 8 clusters. This model could be

adapted to other learning style data and would contribute to developing a system to support the grouping of students according to their learning styles appropriately by finding suitable parameters for the data set. The results of this research could be effectively and appropriately applied in the development of different clustering models in the future.

Keywords : Index learning styles, Student Grouping, Hierarchical Clustering, CRISP-DM

กิตติกรรมประกาศ

งานวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ทางผู้วิจัยต้องขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร. พาสน์ ปราโมกข์ชน อาจารย์ที่ปรึกษาวิทยานิพนธ์เป็นอย่างสูง ที่ได้ให้ความรู้ ความเข้าใจในกระบวนการ และให้คำแนะนำสำหรับการค้นคว้าทดลองตลอดจนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี และขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร. ปวีณ เขื่อนแก้ว และอาจารย์ ดร.กิติศักดิ์ โอสถานันต์กุล อาจารย์ที่ปรึกษาร่วมวิทยานิพนธ์ และเจ้าหน้าที่ประจำภาควิชา และเจ้าหน้าที่ฝ่ายบัณฑิตวิทยาลัยทุกท่านที่คอยให้คำแนะนำ ช่วยเหลือด้านเอกสาร กำหนดการ และประสานงานต่าง ๆ เป็นอย่างดี ตลอดจนคณะครูโรงเรียนชลประทานผาแตกที่ช่วยเหลือด้านการนำโมเดลไปปรับใช้กับการสอน และคุณครูรัชฎาภรณ์ ตาหล้า รวมถึงดัมโบ้ แฮมเตอร์สัตว์เลี้ยงซึ่งให้กำลังใจตลอดทำงานวิจัย

ขอขอบคุณนักศึกษาปริญญาตรี สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยแม่โจ้ และนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนชลประทานผาแตก ที่ให้ข้อมูลในการตอบแบบประเมินรูปแบบการเรียนรู้ เพื่อใช้ในงานวิทยานิพนธ์ฉบับนี้ โดยหวังว่างานวิจัยนี้จะมีประโยชน์ต่อการพัฒนาทางด้านการศึกษาในอนาคต

สุดท้ายนี้ต้องขอขอบคุณบิดา และมารดา ที่สนับสนุน และช่วยเหลือในการทำงานวิจัยด้วยการให้ทุนการศึกษาระดับปริญญาโท ตลอดจนกำลังใจในการทำงานวิจัยจนสำเร็จลุล่วง นอกจากนี้ยังมีผู้ที่ให้ความร่วมมือช่วยเหลืออีกหลายท่าน ซึ่งผู้เขียนไม่สามารถกล่าวนามในที่นี้ได้หมด จึงขอขอบคุณทุกท่านเหล่านั้นไว้ ณ โอกาสนี้ด้วย

ณรงค์ศักดิ์ โยตมาน

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ค

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ช

สารบัญ ..... ซ

สารบัญตาราง ..... ฎ

สารบัญภาพ ..... ฐ

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของการวิจัย 2

1.3 ขอบเขตของการทำงานวิจัย 2

1.3.1 ขอบเขตการวิจัย 2

1.3.2 ขอบเขตด้านเนื้อหา 2

1.3.3 ขอบเขตด้านเวลา 2

1.3.4 ขอบเขตด้านสถานที่ 3

1.4 กรอบแนวคิดการวิจัย 3

1.5 ขั้นตอนในการสร้าง และพัฒนาระบบ 3

1.6 เครื่องมือ และอุปกรณ์ 4

1.6.1 ด้านการจัดเก็บข้อมูล 4

1.6.2 ด้านเชิงระบบ และการพัฒนาซอฟต์แวร์ 4

1.7 ประโยชน์ที่คาดว่าจะได้รับ 4

บทที่ 2 ทฤษฎี และการตรวจเอกสาร 5

2.1 แนวคิด และทฤษฎีในการทำงานวิจัย 5
2.2 การทำเหมืองข้อมูล (Data Mining Techniques) 5
2.3 ขั้นตอนการทำเหมืองข้อมูล (Data Mining) 8
2.4 ประเภทของข้อมูลที่สามารถทำ Data Mining 9
2.5 ลักษณะเฉพาะของข้อมูลที่ใช้ทำเหมืองข้อมูล 10
2.6 แนวคิดในการพัฒนาด้านการศึกษา 10
2.7 การวิเคราะห์กลุ่มข้อมูลด้วยเทคนิค Clustering 10
2.8 การวัดระยะทาง (Distance Measure) 11
2.8.1 การวัดระยะทางแบบ Euclidean Distance 12
2.8.2 การวัดระยะทางแบบ Manhattan Distance 13
2.9 หลักเกณฑ์ในการรวมกลุ่ม (Linkage Methods) 13
2.9.1 หลักเกณฑ์ในการรวมกลุ่มแบบ Single Linkage Method 14
2.9.2 หลักเกณฑ์ในการรวมกลุ่มแบบ Complete Linkage Method 15
2.9.3 หลักเกณฑ์ในการรวมกลุ่มแบบ Average Linkage Method 15
2.9.4 หลักเกณฑ์ในการรวมกลุ่มแบบ Ward’s Linkage Method 17
2.9.5 หลักเกณฑ์ในการรวมกลุ่มแบบ Median Linkage Method 18
2.10 การกำหนดคลัสเตอร์ที่เหมาะสม (Determining Optimal Cluster) 19
2.11 Sum of Squared Errors (SSE) 20
2.12 อัตราการเปลี่ยนแปลงของ Sum of Squared Errors (SSE) หรือ Relative Change of SSE (%Change) 20
2.13 รูปแบบการเรียนรู้ Index of Learning Styles (ILS) 21
2.14 วรรณกรรมที่เกี่ยวข้องกับงานวิจัย 22
บทที่ 3 วิธีการดำเนินการวิจัย 29
3.1 การทำความเข้าใจกับปัญหา (Business Understanding) 29

3.2 การเตรียมข้อมูล (Data Understanding) 30
3.3 การแก้ไข และดัดแปลงข้อมูล (Data Preparation) 33
3.4 กระบวนการสร้าง และพัฒนาโมเดล (Modeling) 34
3.6 แนวคิดวิธีการทดสอบโมเดล 37
3.7 แนวคิดวิธีการวัดประสิทธิภาพข้อมูล 38
3.8 สรุปผลการวิจัย และจัดทำรายงานวิทยานิพนธ์ 39
บทที่ 4 ผลการดำเนินการวิจัย 40
4.1 การวิเคราะห์ปัญหา และการเลือกปัญหาจากข้อมูลที่มี 40
4.2 การเตรียมข้อมูล 40
4.3 การเลือกเทคนิคที่เหมาะสม 41
4.4 ผลการทดสอบความแม่นยำของตัวต้นแบบ (Evaluation) 41
4.4.1 ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 42
4.4.2 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 45
4.4.3 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2 50
4.4.4 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3 53
4.4.5 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 4 57
4.5 การวัดค่าประสิทธิภาพในการจัดกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 60
4.5.1 ขั้นตอนที่ 1 การพิจารณาค่าเฉลี่ยของพารามิเตอร์ Linkage ที่เหมาะสม 60
4.5.2 ขั้นตอนที่ 2 การพิจารณาจำนวนคลัสเตอร์ของพารามิเตอร์ Linkage ที่เหมาะสม 60
4.6 การวัดค่าประสิทธิภาพในการจัดกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 65
4.6.1 ขั้นตอนที่ 1 การพิจารณาค่าเฉลี่ยของพารามิเตอร์ Linkage ที่เหมาะสม 65
4.6.2 ขั้นตอนที่ 2 การพิจารณาจำนวนคลัสเตอร์ของพารามิเตอร์ Linkage ที่เหมาะสม 66
4.7 การนำไปใช้งาน (Deployment) 70
4.8 การประเมินความพึงพอใจจากผู้ใช้งาน 74

บทที่ 5 สรุป และข้อเสนอแนะ 76
5.1 สรุปผลการวิจัย 76
5.2 ปัญหา และอุปสรรค 77
5.3 ข้อเสนอแนะ 78
บรรณานุกรม 79
ภาคพนวก ก การนำเสนอผลงานวิชาการระดับนานาชาติ 81
ประวัติผู้วิจัย 86

สารบัญตาราง

หน้า

ตารางที่ 1 ตารางอธิบายลักษณะของรูปแบบการเรียนรู้แบบ Felder Silverman 21

ตารางที่ 2 โครงสร้างการจัดเก็บข้อมูลพื้นฐานของนักเรียน นักศึกษา 30

ตารางที่ 3 ตัวอย่างข้อมูลดิบ (Raw Data) ข้อมูลพื้นฐานของนักเรียน และนักศึกษา 31

ตารางที่ 4 โครงสร้างการจัดเก็บข้อมูลรูปแบบการเรียนรู้ Index Learning Style (ILS) 31

ตารางที่ 5 ตัวอย่างข้อมูลดิบ (Raw Data) รูปแบบการเรียนรู้ Index Learning Style (ILS) 32

ตารางที่ 6 ตัวอย่างข้อมูลรหัสระดับการศึกษา (Education Level) ตามรหัสข้อมูลมาตรฐานกลาง (Code List) 32

ตารางที่ 7 ตัวอย่างข้อมูลรหัสชั้นปี (Grade Level) ตามรหัสข้อมูลมาตรฐานกลาง (Code List) 32

ตารางที่ 8 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่ม ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Euclidean Distance และLinkage จำนวน 2 - 10 กลุ่ม 61

ตารางที่ 9 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่ม ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Manhattan Distance และLinkage จำนวน 2-10 กลุ่ม 62

ตารางที่ 10 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Euclidean Distance และLinkage จำนวน 2 - 10 กลุ่ม 66

ตารางที่ 11 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Manhattan Distance และLinkage จำนวน 2 - 10 กลุ่ม 68

ตารางที่ 12 ตารางสรุปคะแนนจากแบบสอบถามความพึงพอใจผู้ใช้งานระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ 75

สารบัญภาพ

หน้า

ภาพที่ 1 กระบวนการเทคนิคการทำเหมืองข้อมูล 6
ภาพที่ 2 ขั้นตอนการแบ่งกลุ่มแบบ Hierarchical Agglomerative Clustering 8
ภาพที่ 3 กระบวนการเหมืองข้อมูลแบบ CRISP-DM 9
ภาพที่ 4 การวัดระยะทางแบบ Euclidean Distance 12
ภาพที่ 5 การวัดระยะทางแบบ Manhattan Distance 13
ภาพที่ 6 การรวมกลุ่มแบบ Single Linkage Method 14
ภาพที่ 7 การรวมกลุ่มแบบ Complete Linkage Method 15
ภาพที่ 8 การรวมกลุ่มแบบ Average Linkage Method 16
ภาพที่ 9 การรวมกลุ่มแบบ Ward Linkage Method 18
ภาพที่ 10 การหาจำนวนกลุ่มที่เหมาะสมด้วย Elbow Method 19
ภาพที่ 11 ลักษณะรูปแบบการเรียนรู้ Felder-Silverman Learning Preferences 22
ภาพที่ 12 การแปลงข้อมูลหลังการประเมิน ให้อยู่ในรูปแบบสำหรับการวิเคราะห์ 34
ภาพที่ 13 ตัวอย่างคำถามแบบประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) 35
ภาพที่ 14 ตารางแบบประเมินรูปแบบการเรียนรู้ ILS 36
ภาพที่ 15 ตารางสรุปผลลัพธ์ รูปแบบการเรียนรู้ Index Learning Style 37
ภาพที่ 16 กระบวนการปรับแต่งพารามิเตอร์ของโมเดลการแบ่งกลุ่ม 38
ภาพที่ 17 แผนผังสรุปกระบวนการวิจัย 39
ภาพที่ 18 ชุดข้อมูลของนักเรียนก่อนนำไปวิเคราะห์ 41
ภาพที่ 19 Dendrogram การจัดกลุ่มของนักเรียนชั้นมัธยมศึกษาปีที่ 3 44
ภาพที่ 20 ผลสมาชิกกลุ่มชุดข้อมูลชั้นมัธยมศึกษาปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม 45

ภาพที่ 21 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 1 49
ภาพที่ 22 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม 50
ภาพที่ 23 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 2 52
ภาพที่ 24 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม 53
ภาพที่ 25 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 1 56
ภาพที่ 26 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม 57
ภาพที่ 27 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 4 59
ภาพที่ 28 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 4 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม 60
ภาพที่ 29 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Euclidean Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change 64
ภาพที่ 30 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Manhattan Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change 65
ภาพที่ 31 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Euclidean Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change 69
ภาพที่ 32 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Manhattan Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change 70
ภาพที่ 33 ผังงาน FlowChart แสดงการทำงานของระบบสนับสนุนการสอน โดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ 71

ภาพที่ 34 หน้าจอการนำเข้าข้อมูลทักษะการเรียนรู้ของนักเรียน และดาวน์โหลดแบบฟอร์มการนำเข้าข้อมูล 72
ภาพที่ 35 หน้าจอแสดงการแนะนำกลุ่มที่เหมาะสมกับชุดข้อมูลระดับชั้นเรียน 72
ภาพที่ 36 หน้าจอแสดงผลการค้นหารายชื่อนักเรียน ชั้นเรียน ห้องเรียน และจำนวนกลุ่ม 73
ภาพที่ 37 หน้าจอแสดงไฟล์ดาวน์โหลด Excel File รายชื่อนักเรียนจากการแบ่งกลุ่มชั้นเรียน..... 73
ภาพที่ 38 ผู้ตอบแบบสอบถามความพึงพอใจผู้ใช้งานระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้..... 74

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การศึกษารูปแบบการเรียนรู้ที่หลากหลายของนักเรียน และนักศึกษาในชั้นเรียนแต่ละรายวิชานั้น ครูและอาจารย์ผู้สอนมักจะมีแนวทางหรือรูปแบบในการสอนของตัวเอง ซึ่งจะปรับใช้กับนักเรียนในแต่ละชั้นปี เช่นในภาคการศึกษาที่ 1 มีการสอนแบบใช้สื่อวิดิโอหรือรูปภาพในการสอน และภาคการศึกษาที่ 2 ปรับให้มีการเรียนการสอนแบบบรรยาย ซึ่งต้องให้ระยะเวลาในการวางแผนการสอน และการทำความเข้าใจนักเรียนในชั้นเรียนที่ค่อนข้างมาก ทำให้เกิดความไม่ตรงกันของรูปแบบการเรียนรู้ของนักเรียน และรูปแบบการสอนแบบดั้งเดิมของอาจารย์ซึ่งเกิดประสิทธิภาพในการเรียนในชั้นเรียนที่น้อย นอกจากนี้ยังมีปัญหาเกี่ยวกับรูปแบบการเรียนรู้ของนักเรียน และรูปแบบการสอนของอาจารย์ที่ไม่ตรงกันในหลายด้าน การทำงานกลุ่มของนักเรียนภายในชั้นเรียนจะพบกับความหลายหลายของความต้องการการเรียนรู้ที่อยู่ในกลุ่ม และการมอบหมายงานหรือรูปแบบการสอนแบบดั้งเดิมจะตรงกับนักเรียนบางคนภายในกลุ่มเท่านั้น ทำให้การสอนรูปแบบดั้งเดิมจะเกิดประสิทธิภาพกับแค่นักเรียนบางคนในกลุ่มที่มีความเข้ากันได้ของรูปแบบการสอนดั้งเดิม การวางแผนการสอนที่เหมาะสมกับนักเรียนแต่ละคนตามรูปการเรียนรู้นั้นจะเพิ่มประสิทธิภาพในการเรียนให้กับกับนักเรียนอย่างสูงสุด แต่จะเพิ่มภาระการวางแผนการสอนให้กับอาจารย์อย่างมากเพราะจะต้องวางแผนการสอนให้เท่ากับจำนวนนักเรียนแต่ละคนในชั้นเรียน ซึ่งการลดภาระของอาจารย์ในการวางแผนการสอน การทำความเข้าใจรูปแบบการเรียนรู้ของนักเรียนในชั้นเรียน การแบ่งกลุ่มให้เหมาะสมโดยที่ภายในกลุ่มมีความต้องการการรูปแบบเรียนรู้ที่คล้ายคลึงกันลดการวางแผนการสอน และภาระการสอนกับอาจารย์ผู้สอน

การจัดกลุ่มนักเรียนในห้องเรียนเป็นกลยุทธ์การสอนทั่วไปที่เกี่ยวข้องกับการแบ่งนักเรียนออกเป็นกลุ่มเล็ก ๆ เพื่อจุดประสงค์หรือกิจกรรมเฉพาะ เป้าหมายคือเพื่อให้นักเรียนมีโอกาสทำงานร่วมกัน สื่อสาร และเรียนรู้จากกันและกัน โดยการจัดกลุ่มทำได้หลายวิธี เช่น สุ่ม ตามความสามารถหรือตามความสนใจ วิธีการนี้สามารถเพิ่มการมีส่วนร่วมของนักเรียน ส่งเสริมทักษะทางสังคม และสนับสนุนรูปแบบการเรียนรู้ที่แตกต่างกัน การทำงานกลุ่มที่มีประสิทธิภาพยังต้องการความคาดหวังที่ชัดเจน งานที่ชัดเจน และการสนับสนุนจากอาจารย์อย่างต่อเนื่อง เพิ่มประสิทธิภาพการเรียนให้กับนักเรียน และลดภาระของอาจารย์ในการวางแผนการสอน ความเข้าใจรูปแบบการเรียนรู้ของนักเรียนแต่ละคนในชั้นเรียน โดยใช้รูปแบบการเรียนรู้ Index Learning Style (ILS) ในการแทนรูปแบบการเรียนรู้ของนักเรียน และลักษณะการเรียนรู้ที่เหมาะสม

ระบบสนับสนุนการตัดสินใจการสอนของอาจารย์ผู้สอน โดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ และการหาวิธีการในการแบ่งกลุ่มนักเรียนในชั้นเรียนที่เหมาะสม และมีประสิทธิภาพ

1.2 วัตถุประสงค์ของการวิจัย

เพื่อการพัฒนาระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ซึ่งจะช่วยให้อาจารย์ผู้สอนเตรียมวางแผนการสอน การปรับปรุงสื่อการสอน กลยุทธ์การสอน และวิธีการประเมินให้เหมาะสมกับความต้องการของนักเรียนที่เหมาะสมกับกลุ่มรูปแบบการเรียนรู้ของนักเรียนในชั้นเรียน และเพิ่มประสิทธิผลของการศึกษาโดยปรับแต่งวิธีการสอนให้เหมาะสมกับความชอบในการเรียนรู้ของนักเรียนแต่ละคนด้วยการกำหนดรูปแบบการเรียนรู้ที่โดดเด่นของนักเรียนแต่ละคนจากการประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) และอาจารย์ผู้สอนสามารถปรับแต่งจำนวนกลุ่มของนักเรียนในชั้นเรียนให้เหมาะสมกับสื่อการสอน กลยุทธ์การสอน และวิธีการประเมินให้เหมาะสมกับทรัพยากรด้านการศึกษา

1.3 ขอบเขตของการทำงานวิจัย

1.3.1 ขอบเขตการวิจัย

งานวิจัยผู้วิจัยได้ทำการศึกษาวิธีการแบ่งกลุ่มในชั้นเรียน และจำนวนกลุ่มที่เหมาะสมกับชั้นเรียนโดยใช้รูปแบบการเรียนรู้ของนักเรียน ซึ่งการเก็บรวบรวมข้อมูลจริงจากชั้นเรียนนักศึกษาระดับปริญญาตรี ชั้นปีที่ 1 - 4 สาขาวิชาวิทยาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยแม่โจ้ และชั้นเรียนนักเรียนระดับมัธยมศึกษาปีที่ 3 โรงเรียนชลประทานผาแตก โดยการเก็บข้อมูลจากการตอบแบบสอบถามรูปแบบการเรียนรู้ และเก็บข้อมูลเพื่อใช้ในการแบ่งกลุ่มชุดข้อมูล

1.3.2 ขอบเขตด้านเนื้อหา

การทบทวนงานวิจัยที่มีอยู่อย่างครอบคลุมเกี่ยวกับรูปแบบการเรียนรู้ ดัชนีรูปแบบการเรียนรู้ และการประยุกต์ใช้ในการศึกษาศึกษา ค้นคว้าทฤษฎี และเอกสารที่เกี่ยวข้องกับรูปแบบการเรียนรู้ของนักเรียน Index Learning Style (ILS) จากนั้นนำชุดข้อมูลรูปแบบการเรียนรู้ของนักเรียนมาเข้าสู่กระบวนการวิเคราะห์ข้อมูล เพื่อหาวิธีการแบ่งกลุ่มชุดข้อมูลที่เหมาะสมกับชุดข้อมูลรูปแบบการเรียนรู้ของนักเรียนในชั้นเรียน รวมถึงการเผยแพร่ผลการวิจัยและข้อเสนอแนะผ่านการนำเสนอ การตีพิมพ์ในวารสารวิชาการ และช่องทางอื่น ๆ

1.3.3 ขอบเขตด้านเวลา

ระยะเวลาในการศึกษาวิจัยรวบรวมข้อมูลในการเรียนรู้ และทดสอบกับชุดข้อมูล เพื่อหาวิธีการแบ่งกลุ่มที่เหมาะสม และมีประสิทธิภาพของระบบแนะนำการแบ่งกลุ่มนักเรียนตามรูปแบบการเรียนในครั้งนี้ ระหว่างเดือน กันยายน 2562 ถึงเดือน กันยายน 2563

1.3.4 ขอบเขตด้านสถานที่

สำหรับการทดลองชุดข้อมูลของนักศึกษามหาวิทยาลัย จัดเก็บข้อมูลชั้นเรียนนักศึกษาระดับปริญญาตรี ชั้นปีที่ 1 - 4 สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์มหาวิทยาลัยแม่โจ้ อำเภอสันทราย จังหวัดเชียงใหม่ ปีการศึกษา 2563 จำนวน 162 คน และสำหรับการทดลองชุดข้อมูลนักเรียน จัดเก็บข้อมูลชั้นเรียนระดับมัธยมศึกษาปีที่ 3 โรงเรียนชลประทานผาแตก อำเภอดอยสะเก็ด จังหวัดเชียงใหม่ ปีการศึกษา 2564 จำนวน 35 คน ซึ่งโดยรวมแล้วการวิจัยมีจุดมุ่งหมายเพื่อแนะนำการจัดกลุ่มตามรูปแบบการเรียนรู้ Index Learning Style (ILS) อย่างมีประสิทธิภาพ ส่งเสริมสนับสนุนการเรียน การสอนสำหรับนักเรียน และอาจารย์ผู้สอน

1.4 กรอบแนวคิดการวิจัย

การทำวิจัยเป็นกระบวนการที่ต้องใช้แนวคิดหลายอย่างเพื่อให้เกิดผลลัพธ์ที่น่าเชื่อถือ และมีคุณค่าต่อการพัฒนาองค์ความรู้โดยกรอบแนวคิดของการวิจัยครั้งนี้จะประกอบไปด้วยองค์ประกอบหลายอย่าง ดังนี้

1. การสำรวจวรรณกรรม (Literature review) เป็นการศึกษา และสรุปผลการวิจัยที่เคยมีอยู่เกี่ยวกับเรื่องที่ต้องการวิจัยในด้านการศึกษา เรื่องรูปแบบการเรียนรู้ และเทคนิคการแบ่งกลุ่ม เพื่อให้เข้าใจแนวคิด และผลลัพธ์ เปรียบเทียบผลการวิจัยที่มีอยู่กับผลการวิจัยที่จะทำ

2. การวางแผนวิจัย (Research design) เพื่อกำหนดวัตถุประสงค์ของการวิจัย ออกแบบแผนการวิจัย กำหนดกลุ่มตัวอย่างที่จะเป็นเป้าหมายของการวิจัย รวมถึงการกำหนดวิธีการเก็บข้อมูลและวิเคราะห์ข้อมูล

3. การเก็บข้อมูล (Data collection) เป็นกระบวนการที่ต้องพยายามรวบรวมข้อมูลที่เป็นประโยชน์ และเป็นอิสระจากผู้ที่เป็นเป้าหมายของการวิจัย การเก็บข้อมูลจากการประเมินรูปแบบการเรียนรู้ของนักเรียนผ่านแบบสอบถามรูปแบบการเรียนรู้ Index Learning Style (ILS)

4. การวิเคราะห์ข้อมูล (Data analysis) เป็นกระบวนการที่ใช้เครื่องมือวิเคราะห์ข้อมูลที่จัดเก็บ โดยใช้เทคนิคเหมืองข้อมูลที่เหมาะสมกับชุดข้อมูลรูปแบบการเรียนรู้ของนักเรียน

5. การประเมินผล (Evaluation) เป็นกระบวนการสรุปผลลัพธ์ของการวิจัย เทคนิคเหมืองข้อมูลที่เหมาะสมกับการใช้กับชุดข้อมูล โดยการนำผลลัพธ์จากการแบ่งกลุ่มแบบต่าง ๆ มาเปรียบเทียบ และประเมินประสิทธิภาพเพื่อจัดทำระบบแนะนำการแบ่งกลุ่มรูปแบบการเรียนรู้

1.5 ขั้นตอนในการสร้าง และพัฒนาระบบ

1. ศึกษาข้อมูลรูปแบบการเรียนรู้ของนักเรียน ทฤษฎี งานวิจัย และกระบวนการที่เกี่ยวข้องกับการแบ่งกลุ่มนักเรียน

2. ศึกษาวิธีการทำเหมืองข้อมูลโดยใช้เทคนิคต่าง ๆ ที่เกี่ยวข้องกับงานวิจัย

3. ตรวจสอบความสมบูรณ์ข้อมูลดิบ จากการประเมินรูปแบบการเรียนรู้ของนักเรียน

4. นำเทคนิคในการแบ่งกลุ่ม สร้างรูปแบบการแบ่งกลุ่มต่าง ๆ เพื่อทดลองหาวิธีการแบ่งกลุ่มที่เหมาะสมกับชุดข้อมูลรูปแบบการเรียนรู้ของนักเรียนในชั้นเรียน

5. ทดสอบประสิทธิภาพของกระบวนการแบ่งกลุ่มที่เหมาะสมกับชุดข้อมูล

6. สรุปผลการวิจัย และจัดทำรายงานการวิจัย

1.6 เครื่องมือ และอุปกรณ์

1.6.1 ด้านการจัดเก็บข้อมูล

1.1 แบบสอบถาม (Questionnaire) เป็นชุดคำถามเพื่อใช้ในการจัดรูปแบบการเรียนรู้ของนักเรียนตามรูปแบบของ Index Learning Style (ILS) โดยออกมาเป็นรูปแบบของตัวเลข และคะแนนสำหรับการวิเคราะห์

1.6.2 ด้านเชิงระบบ และการพัฒนาซอฟต์แวร์

2.1 ภาษาที่ใช้ในการพัฒนาระบบ ภาษา PHP และ HTML

2.2 ฐานข้อมูล และการจัดเก็บข้อมูล PhpMyAdmin และ MySQL

2.3 เครื่องมือที่ใช้ในการพัฒนาระบบ Visual Studio Code และ SourceTree

2.4 เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล Minitab

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1. จำนวนกลุ่มที่เหมาะสมกับนักเรียนในชั้นเรียน เพื่อให้อาจารย์ทราบถึงจำนวนกลุ่มที่เหมาะสมกับนักเรียนในชั้นเรียนตามรูปแบบการเรียนรู้ของนักเรียน และระบบสามารถช่วยให้อาจารย์จัดการห้องเรียนได้อย่างมีประสิทธิภาพมากขึ้นโดยการให้ข้อมูลเชิงลึกเกี่ยวกับความชอบและความท้าทายในการเรียนรู้ของนักเรียนแต่ละคน

2. ขั้นตอนวิธีการ (Algorithm) ในการแบ่งกลุ่มของนักเรียนตามรูปแบบการเรียนรู้ เพื่อหาวิธีการที่เหมาะสม และมีประสิทธิภาพในการแบ่งกลุ่มชุดข้อมูล (Cluster) ให้ข้อเสนอแนะรายบุคคลแก่นักเรียน ช่วยให้พวกเขาเข้าใจจุดแข็งจุดอ่อนของตนเอง และกำหนดเป้าหมายสำหรับการปรับปรุง

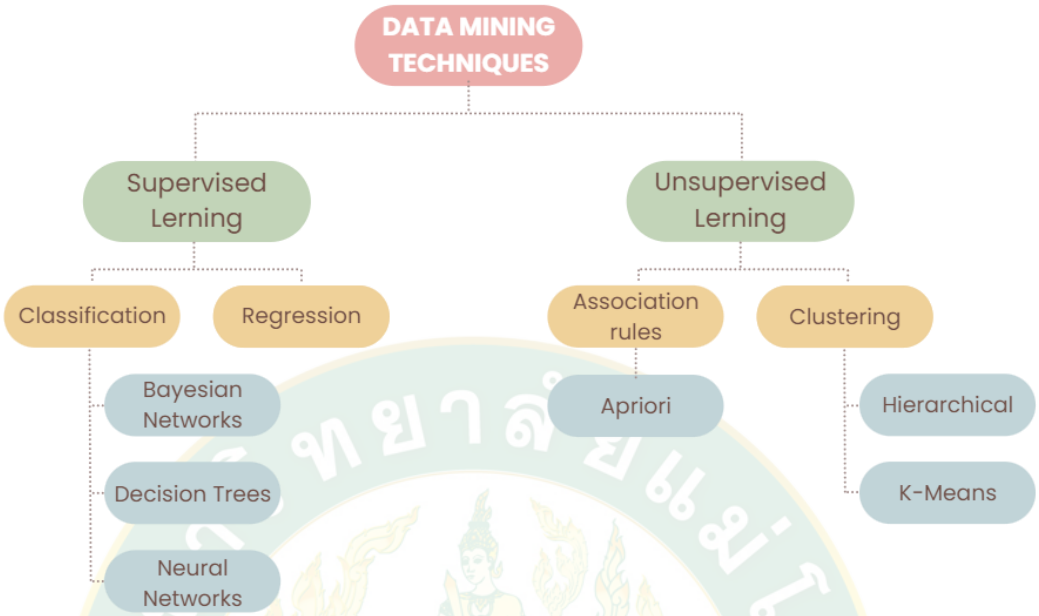
บทที่ 2
ทฤษฎี และการตรวจเอกสาร

2.1 แนวคิด และทฤษฎีในการทำงานวิจัย

งานวิจัยดังกล่าวเป็นงานวิจัยเพื่อสนับสนุนทางด้านการศึกษา โดยการวิเคราะห์ข้อมูลดัชนีการเรียนรู้ และพัฒนาขั้นตอนการแบ่งกลุ่มตามรูปแบบการเรียนรู้ของนักเรียนในชั้นเรียน เพื่อการทำความเข้าใจรูปแบบการเรียนรู้ของนักเรียน และภาพรวมของกลุ่มรูปแบบการเรียนรู้ในชั้นเรียน รวมถึงประสิทธิภาพในการแบ่งกลุ่ม ตามรูปแบบการเรียนรู้ให้เหมาะสมกับนักเรียนชั้นเรียนโดยมีพื้นฐานมาจากแนวคิดที่ว่าบุคคลมีรูปแบบการเรียนรู้และความชอบเฉพาะตัว การประเมินรูปแบบการเรียนรู้ของนักเรียนมีวัตถุประสงค์เพื่อประเมินวิธีการรับและประมวลผลข้อมูลที่ต้องการของแต่ละคน เช่น ภาพ การได้ยิน การเคลื่อนไหวร่างกาย และรูปแบบอื่น ๆ ผลลัพธ์ของการประเมินสามารถใช้เพื่อปรับแต่งประสบการณ์การเรียนรู้ และปรับปรุงประสิทธิภาพของระบบสนับสนุนการสอน เป้าหมายคือการสร้างระบบที่ปรับให้เข้ากับความต้องการ และความพึงพอใจเฉพาะของผู้เรียนแต่ละคน โดยมอบประสบการณ์การเรียนรู้ที่น่ามีส่วนร่วม และมีประสิทธิภาพมากขึ้น ซึ่งการทดลองงานวิจัยได้มีการเก็บข้อมูล การวิเคราะห์ข้อมูล และการทดลองแบ่งกลุ่มในชั้นเรียนของ กลุ่มนักศึกษาปริญญาตรี ชั้นปีที่ 1-4 สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้ จังหวัดเชียงใหม่ และกลุ่มนักเรียน ชั้นมัธยมศึกษาปีที่ 3 โรงเรียนชลประทานผาแตก อำเภอดอยสะเก็ด จังหวัดเชียงใหม่ โดยใช้แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ดังต่อไปนี้

2.2 การทำเหมืองข้อมูล (Data Mining Techniques)

การทำเหมืองข้อมูล (Data Mining Techniques) เป็นกระบวนการค้นหารูปแบบความสัมพันธ์ และข้อมูลเชิงลึกจากข้อมูลจำนวนมากโดยใช้เทคนิค และอัลกอริทึมต่าง ๆ ซึ่งเกี่ยวข้องกับการสกัดข้อมูลที่ซ่อนอยู่ และมีประโยชน์จากฐานข้อมูล คลังข้อมูล และการแปลงข้อมูลให้เป็นรูปแบบที่เข้าใจ เป้าหมายของการทำเหมืองข้อมูลคือ การระบุรูปแบบความสัมพันธ์ภายในข้อมูลที่สามารถนำมาใช้ในการตัดสินใจการสร้างแบบจำลองเชิงทำนายช่วยให้เข้าใจกลุ่มเป้าหมาย การดำเนินงาน แนวโน้มของกลุ่มเป้าหมาย ดังภาพที่ 1



ภาพที่ 1 กระบวนการเทคนิคการทำเหมืองข้อมูล

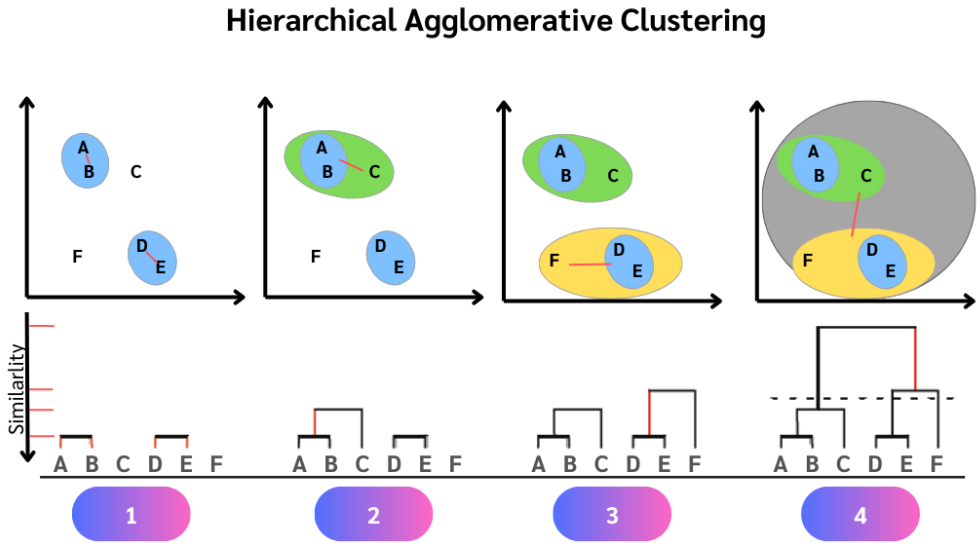
Supervised Learning เป็นการเรียนรู้จากข้อมูลที่มีการตอบสนองต่อผลลัพธ์ (Output) ที่ต้องการได้แก่แต่ละตัวอย่างของข้อมูล โดยมีข้อมูล Input และ Output ที่เป็นข้อมูลสอน (Training Data) เพื่อสร้างโมเดล (Model) ที่สามารถใช้ทำนายผลลัพธ์จากข้อมูล Input ใหม่ ๆ ที่ไม่เคยเห็นมาก่อนได้ โดยมักนำมาใช้ในงานที่เรียนรู้การจำแนก และการคาดการณ์อย่างเช่น การจำแนกสิ่งของออกเป็นกลุ่ม หรือการคาดการณ์ราคาของหุ้น การอ่านอีเมล์แล้วจำแนกว่าเป็นอีเมล์สแปมหรือไม่เป็นต้น

Unsupervised Learning เป็นกระบวนการเรียนรู้ของเครื่องจักร (Machine Learning) ที่ไม่ได้รับการสอนด้วยข้อมูลที่มีคำตอบหรือผลเฉลย โดยการเรียนรู้นี้จะใช้ข้อมูลที่ไม่มีการติดตั้งหมวดหมู่หรือไม่มีการตอบกลับเป็นผลลัพธ์ (Unlabeled Data) เพื่อหาความสัมพันธ์หรือโครงสร้างภายในข้อมูล โดยเรียนรู้แบบ Unsupervised จะทำการแบ่งกลุ่มหรือจัดกลุ่มข้อมูลโดยอัตโนมัติ โดยไม่ต้องมีการระบุว่าแต่ละกลุ่มมีคุณลักษณะอะไรบ้าง หรือกำหนดลักษณะของแต่ละกลุ่มไว้ล่วงหน้า โดยจะให้โมเดลหรือเครื่องจักรหากลุ่ม และลักษณะของแต่ละกลุ่มเองจากข้อมูลที่ได้รับมา และใช้เป็นข้อมูลเพื่อทำนายหรือค้นหาความสัมพันธ์ในข้อมูลต่อไป

Clustering คือกระบวนการที่ใช้เพื่อแบ่งข้อมูลที่มีความหลากหลายออกเป็นกลุ่มหรือหมวดหมู่ที่มีลักษณะคล้ายกัน โดยจะใช้ข้อมูลเชิงลักษณะ (feature) หรือคุณสมบัติต่างๆ ของข้อมูลเพื่อค้นหาลักษณะที่เหมือนกันระหว่างข้อมูลที่มีอยู่ ซึ่งมักนิยมใช้ในการจัดกลุ่มข้อมูลแบบไม่มีการระบุชื่อหรือป้ายกำกับ (Unsupervised Learning) และสามารถนำไปใช้ในการค้นหาความสัมพันธ์ในข้อมูล (Data Mining) หรือในการวิเคราะห์ข้อมูลเพื่อหาข้อมูลที่มีลักษณะคล้ายกันเพื่อสร้างสถิติหรือวิเคราะห์ข้อมูลต่อไป อย่างไรก็ตาม การแบ่งกลุ่มข้อมูลไม่ได้รับการกำหนดเป็นชุดเอง แต่ต้องอาศัยการเรียนรู้จากการจัดกลุ่มแบบตัวอย่าง (example-based learning) โดยใช้แนวคิดของหลักการคลัสเตอร์ (clustering principles) เพื่อค้นหาความคล้ายคลึงระหว่างข้อมูลที่ต้องการจัดกลุ่มกัน ซึ่งวิธีการจัดกลุ่มแบบนี้จะช่วยให้สามารถสกัดข้อมูลที่ซับซ้อนออกมาเป็นสาระสำคัญ และช่วยประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ

K-means เป็นวิธีการแบ่งกลุ่ม (clustering) ของข้อมูลแบบ unsupervised learning ซึ่งใช้วิธีการหาจุดศูนย์กลาง (centroids) ของกลุ่มแล้วนำข้อมูลทั้งหมดมาจัดกลุ่มโดยใช้ระยะทางของแต่ละจุดกับจุดศูนย์กลางของกลุ่มเพื่อกำหนดว่าจะจัดให้เป็นกลุ่มไหน วิธีการนี้เหมาะสำหรับการแบ่งกลุ่มข้อมูลที่มีลักษณะเป็นวงกลม หรือกลุ่มที่มีขนาดแตกต่างกันได้ โดยการเลือกจำนวนกลุ่มที่ต้องการจะแบ่งอยู่ที่ผู้ใช้งานต้องกำหนด และการแบ่งกลุ่มนั้นจะทำซ้ำไปเรื่อย ๆ จนกว่าจุดศูนย์กลางจะไม่เปลี่ยนแปลง หรือมีการเปลี่ยนแปลงในระดับที่ต่ำกว่าที่ผู้ใช้งานกำหนดไว้

Hierarchical Clustering เป็นวิธีการแบ่งกลุ่มข้อมูลในการทำ Clustering โดยที่ข้อมูลจะถูกแบ่งออกเป็นกลุ่มย่อย ๆ ตามความคล้ายคลึงกันของข้อมูล โดยวิธีนี้จะมีการสร้าง Dendrogram (Diagram แบบต้นไม้) ซึ่งแสดงถึงความสัมพันธ์ของข้อมูลในรูปแบบของต้นไม้ ดังภาพที่ 2 โดย Dendrogram จะแสดงข้อมูลแต่ละอันเป็นลำดับเช่นเดียวกับต้นไม้ และการแบ่งกลุ่มจะเริ่มต้นจากจุดที่มีความคล้ายคลึงกันมากที่สุด และสร้างกลุ่มที่ใหญ่ขึ้นไปเรื่อย ๆ จนกว่าจะได้จำนวนกลุ่มที่ต้องการหรือไม่สามารถแบ่งกลุ่มได้อีกต่อไป โดยวิธีการหากลุ่มที่เหมาะสมจะใช้ Linkage Method ซึ่งเป็นวิธีการหาความสัมพันธ์ของข้อมูลในการสร้าง Dendrogram โดยมี Linkage Method 5 แบบ ได้แก่ Single Linkage, Complete Linkage, Average Linkage, Ward Linkage และ Median Linkage



ภาพที่ 2 ขั้นตอนการแบ่งกลุ่มแบบ Hierarchical Agglomerative Clustering

2.3 ขั้นตอนการทำเหมืองข้อมูล (Data Mining)

สำหรับขั้นตอนการทำเหมืองข้อมูล (Data Mining) เป็นกระบวนการที่ค้นหารูปแบบ และความสัมพันธ์ภายในชุดข้อมูล ซึ่งในกระบวนการทำเหมืองข้อมูลแบบที่ใช้ในงานวิจัยนี้ใช้แบบ CRISP-DM (Cross-Industry Standard Process For Data Mining) กระบวนการมาตรฐานที่ใช้สำหรับการทำเหมืองข้อมูล เพื่อทำการวิเคราะห์ และนำไปใช้ประโยชน์ทางธุรกิจ ซึ่งมีลักษณะของกระบวนการ ดังภาพที่ 3 โดยประกอบด้วยขั้นตอนต่อไปนี้

- Business Understanding (เข้าใจปัญหาทางธุรกิจ): รู้จักปัญหาทางธุรกิจ และกำหนดวัตถุประสงค์ของโครงการ เพื่อให้การทำเหมืองข้อมูลสามารถแก้ไขปัญหาได้อย่างมีประสิทธิภาพ

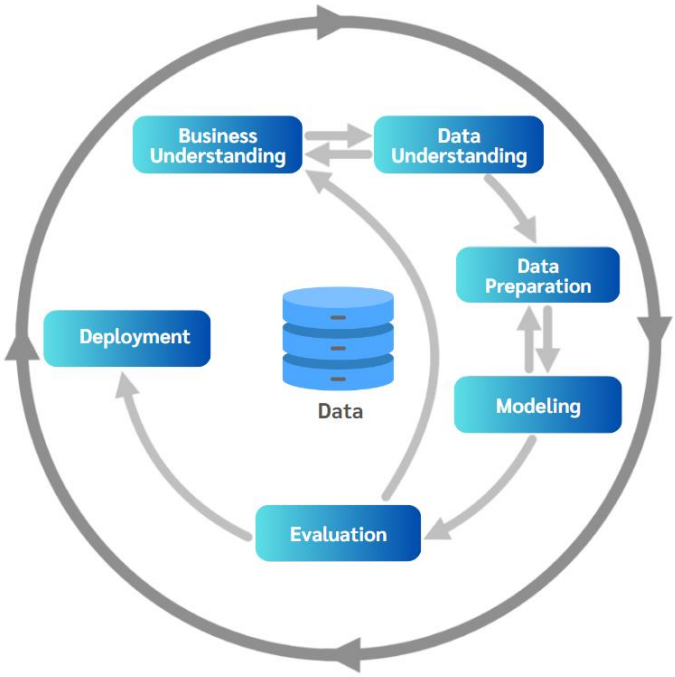
- Data Understanding (เข้าใจข้อมูล): ศึกษาและเข้าใจข้อมูลที่เป็นพื้นฐานสำหรับการทำเหมืองข้อมูล เช่น การเก็บรวบรวมข้อมูล คุณภาพของข้อมูล และความสมบูรณ์ของข้อมูล

- Data Preparation (เตรียมข้อมูล): การเตรียมข้อมูลเพื่อให้สามารถนำมาวิเคราะห์ได้ ซึ่งอาจประกอบด้วยการทำความสะอาดข้อมูล การแปลงรูปแบบของข้อมูล การลบข้อมูลที่ไม่เกี่ยวข้อง เป็นต้น

- Modeling (การสร้างโมเดล): ใช้เทคนิค และอัลกอริทึมต่างๆ เพื่อสร้างโมเดลที่เหมาะสมกับข้อมูล เช่น การวิเคราะห์หลักการความสัมพันธ์ การวิเคราะห์ปัจจัยที่ส่งผลต่อผลลัพธ์ การสร้างโมเดลทางคณิตศาสตร์ เป็นต้น

- Evaluation (การประเมินผล): การประเมินผลของโมเดลที่สร้างขึ้น โดยใช้ข้อมูลการทดสอบเพื่อตรวจสอบว่าโมเดลที่ได้สามารถแก้ไขปัญหาได้อย่างมีประสิทธิภาพหรือไม่

- Deployment (การนำโมเดลไปใช้งาน): การนำโมเดลที่สร้างขึ้นที่เหมาะสมที่สุด ไปใช้งานงานจริง เพื่อวิเคราะห์ และแก้ปัญหาที่ต้องการ



ภาพที่ 3 กระบวนการเหมืองข้อมูลแบบ CRISP-DM

2.4 ประเภทของข้อมูลที่สามารถทำ Data Mining

Data Mining สามารถทำกับประเภทข้อมูลต่อไปนี้ได้:

- Numeric Data: ข้อมูลที่เป็นตัวเลข เช่น อุณหภูมิ, ราคา, ความเร็ว เป็นต้น
- Categorical Data: ข้อมูลที่เป็นหมวดหมู่ เช่น สี, สัญชาติ, ประเภทของสินค้า เป็นต้น
- Ordinal Data: ข้อมูลที่เป็นความสัมพันธ์ของหมวดหมู่ เช่น ระดับการศึกษา, อายุ, ระดับความพึงพอใจ เป็นต้น
- Text Data: ข้อมูลที่เป็นข้อความ เช่น บทความ, ข้อความจากโซเชียลมีเดีย เป็นต้น
- Image Data: ข้อมูลที่เป็นภาพ เช่น รูปถ่าย, ภาพถ่ายดาวเคราะห์ เป็นต้น
- Time-Series Data: ข้อมูลที่เป็นลำดับของข้อมูลตามเวลา เช่น ราคาหุ้น, อุณหภูมิตามช่วงเวลา เป็นต้น
- Spatial Data: ข้อมูลที่เกี่ยวกับพิกัดเช่น พิกัดที่ตั้งของร้านค้า, พิกัด GPS เป็นต้น

2.5 ลักษณะเฉพาะของข้อมูลที่ใช้ทำเหมืองข้อมูล

- มีปริมาณมาก: ข้อมูลที่ใช้ทำเหมืองข้อมูลมักจะมีปริมาณมาก ซึ่งอาจเป็นรูปแบบตารางหรือไฟล์ข้อมูลอื่น ๆ เช่น ภาพถ่ายหรือไฟล์เสียง
- ความหลากหลาย: ข้อมูลที่ใช้ทำเหมืองข้อมูลมักจะมีความหลากหลายในแง่ต่าง ๆ เช่น ความสูง น้ำหนัก อายุ เพศ สถานที่ หรือ เวลา
- มีความไม่แน่นอน: ข้อมูลที่ใช้ทำเหมืองข้อมูลมักมีความไม่แน่นอน เนื่องจากข้อมูลนั้นมักมาจากแหล่งต่าง ๆ ที่ไม่สามารถควบคุมได้ เช่น ข้อมูลจากโซเชียลมีเดีย ข้อมูลการจองโรงแรมออนไลน์ เป็นต้น
- มีความซับซ้อน: ข้อมูลที่ใช้ทำเหมืองข้อมูลมักมีความซับซ้อน เนื่องจากมีข้อมูลหลายรูปแบบและข้อมูลที่เกี่ยวข้องซับซ้อนกัน เช่น ข้อมูลจากฐานข้อมูลระบบการจัดการคลังสินค้า ข้อมูลจากระบบการจัดการโรงงาน และอื่น ๆ

2.6 แนวคิดในการพัฒนาด้านการศึกษา

ระบบสนับสนุนการสอนโดยใช้ Index of Learning Style (ILS) เป็นเครื่องมือวัดลักษณะการเรียนรู้ของนักเรียน และนักศึกษาที่ได้รับความนิยมมากในการศึกษาแนวทางสนับสนุนการสอนและการเรียนรู้ในสถาบันการศึกษา โดยมีเป้าหมายเพื่อช่วยให้คนที่สนใจการสอน และการเรียนรู้สามารถระบุลักษณะการเรียนรู้ของตนเองได้ โดยประกอบด้วยคำถามที่ออกแบบมาเพื่อวัดความถนัดในการเรียนรู้ และเมื่อเสร็จสิ้นการตอบคำถามจะได้รับผลลัพธ์ว่ามีลักษณะการเรียนรู้อยู่ในกลุ่มไหน โดยประกอบด้วยดัชนีความสัมพันธ์ทั้งหมด 4 หมวดหมู่ ได้แก่ การรับรู้ (Sensing) และคิด (Intuitive) การคิด (Visual) และคิด (Verbal) การเรียนรู้ (Sequential) และ (Global) และการตัดสินใจ (Active) และ (Reflective)

ดังนั้นการใช้ระบบสนับสนุนการสอนสามารถช่วยให้อาจารย์ และผู้สอนสามารถตรวจสอบและปรับปรุงการสอนได้ให้เหมาะสมกับลักษณะการเรียนรู้ของนักเรียนหรือนักศึกษาในชั้นเรียน โดยสามารถแนะนำกลุ่มที่เหมาะสมของนักเรียนแต่ละคนในชั้นเรียน โดยที่ไม่ต้องเสียเวลาในการสังเกตุลักษณะของนักเรียนในชั้นเรียน และค่าใช้จ่ายเพิ่มเติมในการทดสอบ และตรวจสอบลักษณะการเรียนรู้ของนักเรียน

2.7 การวิเคราะห์กลุ่มข้อมูลด้วยเทคนิค Clustering

Clustering เป็นเทคนิคในการทำ Data Science ที่เรียกว่า “Unsupervised Learning” ซึ่งเป็นการหาข้อมูลเชิงลึกจากข้อมูล โดยข้อมูลนั้นจะไม่มีการระบุประเภทข้อมูล หรือแบ่งชนิดตัว

แปรในข้อมูลนั้นหลักการในการทำ clustering ก็คือจะจัดข้อสังเกตในชุดข้อมูลที่มี “ลักษณะเหมือนกัน” เข้าด้วยกัน

Unsupervised learning เป็นอัลกอริทึมการเรียนรู้ของเครื่องประเภทหนึ่งที่โมเดลได้รับการฝึกอบรมในชุดข้อมูลโดยไม่มีเอาต์พุตหรือการกำกับดูแลที่มีป้ายกำกับ เป้าหมายของการเรียนรู้แบบไม่มีผู้สอนคือการค้นหารูปแบบ ความสัมพันธ์ หรือโครงสร้างภายในข้อมูล แทนที่จะใช้ข้อมูลที่มีป้ายกำกับเพื่อทำการคาดคะเน เทคนิคทั่วไปของการเรียนรู้แบบไม่มีผู้สอน ได้แก่ การจัดกลุ่ม การลดขนาด และการตรวจจับความผิดปกติ ในเทคนิคเหล่านี้ ตัวแบบจะพยายามเรียนรู้โครงสร้างพื้นฐานของข้อมูลโดยไม่มีข้อมูลใด ๆ มาก่อนเกี่ยวกับฉลากหรือหมวดหมู่ และขึ้นอยู่กับผู้วิจัยที่จะตีความผลลัพธ์ และระบุรูปแบบที่มีความหมาย

Observations หรือ “Data Points” หมายถึงข้อมูลที่รวบรวมผ่านการสังเกต บันทึก และวัดลักษณะต่างๆ ของระบบ กระบวนการ หรือปรากฏการณ์ ข้อมูลนี้อาจเป็นข้อมูลเชิงคุณภาพหรือเชิงปริมาณ และสามารถบันทึกด้วยวิธีการต่างๆ เช่น การสังเกตโดยตรง การสำรวจ หรือการทดลอง การสังเกตมีความสำคัญในการวิจัยและการวิเคราะห์ข้อมูลเนื่องจากเป็นพื้นฐานสำหรับการทดสอบสมมติฐานและทำการสรุปเกี่ยวกับความสัมพันธ์และรูปแบบพื้นฐานในข้อมูลที่มีลักษณะเหมือนกัน คือ Data Points ที่อยู่ใกล้กัน ซึ่งวิธีในการหาเราต้องหา “ระยะทาง” หรือ Distances ระหว่าง Data Points โดย Data Points ที่มีระยะใกล้กันก็คือ Data Points ที่มีลักษณะเหมือนกัน การจัดกลุ่ม (หมายถึง คน สัตว์ สิ่งของ หรือ องค์กร ฯลฯ) หรือเป็นการจัดตัวแปร ออกเป็นกลุ่มย่อย ๆ ตั้งแต่ 2 กลุ่มขึ้นไป กลุ่มที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่เหมือนกัน หรือคล้ายกัน ส่วนกลุ่มที่อยู่ต่างกลุ่มกันจะมีลักษณะที่แตกต่างกัน

- ตัวแปรที่อยู่ในกลุ่มเดียวกัน จะมีความสัมพันธ์กันมากกว่าตัวแปรที่อยู่ต่างกลุ่มกัน
- ตัวแปรที่อยู่ต่างกลุ่มกัน จะมีความสัมพันธ์กันน้อยหรือไม่มีความสัมพันธ์กันเลย

2.8 การวัดระยะทาง (Distance Measure)

Distance Measure เป็นการวัดระยะห่าง (Distance) ระหว่างสองจุดในพื้นที่หรือข้อมูล โดยมักใช้ในการวิเคราะห์ข้อมูล และการจัดกลุ่ม (Clustering) เพื่อหาความคล้ายคลึง (Similarity) หรือความแตกต่าง (Dissimilarity) ของข้อมูล โดยมี Distance Measure หลายแบบ เช่น Euclidean Distance, Manhattan Distance ฯลฯ แต่แบบที่จะเลือกใช้ขึ้นอยู่กับลักษณะของข้อมูลและวัตถุประสงค์ของการวิเคราะห์ โดย Euclidean Distance เป็นการวัดระยะห่างโดยใช้ระยะสอง

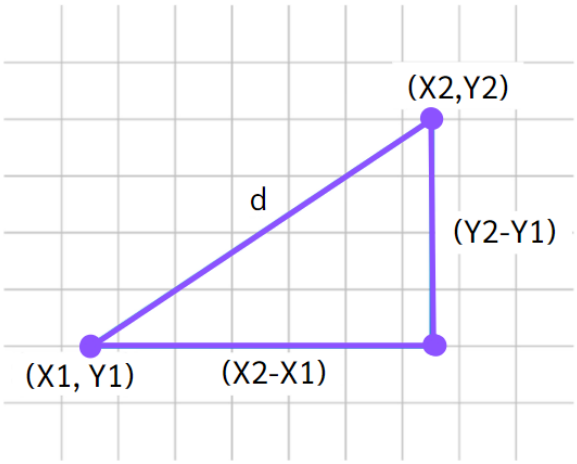
จุดที่ใช้ค่า X และ Y ของแต่ละจุด ในขณะที่ Manhattan Distance จะเป็นการวัดระยะห่างโดยใช้ระยะทางแนวตั้ง และแนวนอนของแต่ละจุด

2.8.1 การวัดระยะทางแบบ Euclidean Distance

Euclidean Distance เป็นวิธีการคำนวณระยะห่างระหว่างจุดสองจุดในระบบพิกัดสองมิติหรือสูงกว่านั้น โดยที่ระยะห่าง Euclidean Distance จะถูกวัดจากความยาวของเส้นตรงที่เชื่อมต่อระหว่างจุดสองจุดนั้นสำหรับสูตรการคำนวณ Euclidean Distance สำหรับจุดที่มีพิกัด (x1, y1) และ (x2, y2) ในระบบพิกัดสองมิติ มีดังนี้:

Euclidean Distance = $\sqrt{(X2 - X1)^2 + (Y2 - Y1)^2}$

เมื่อ x1 และ y1 คือตำแหน่งของจุดแรกในระบบพิกัดแนวตั้ง และแนวนอน
x2 และ y2 คือตำแหน่งของจุดที่สองในระบบพิกัดแนวตั้ง และแนวนอน



ภาพที่ 4 การวัดระยะทางแบบ Euclidean Distance

สำหรับวิธีการคำนวณระยะทางแบบ Euclidean Distance นั้นมีจะใช้ได้ดีเมื่อวัดระยะห่างระหว่างจุดสองจุดในพื้นที่สองมิติ โดยเฉพาะอย่างยิ่งเมื่อระยะห่างต้องการมีการนำไปใช้ในการทำนายหรือการจัดกลุ่มข้อมูล กรณีที่ข้อมูลมิติสูงกว่าสอง Euclidean Distance อาจไม่เหมาะสมเนื่องจากความซับซ้อนของการคำนวณ และจำนวนความสัมพันธ์ระหว่างตัวแปรที่สูงขึ้นอาจทำให้ผลการวัด

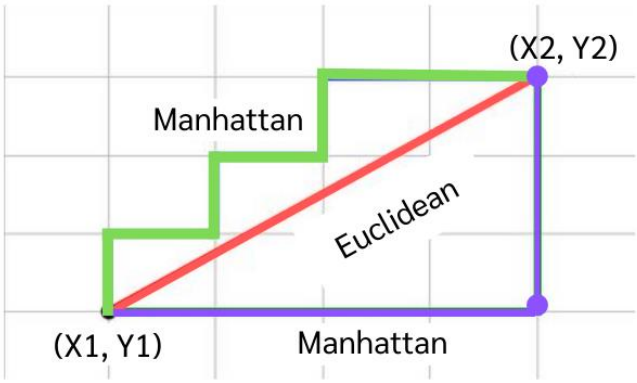
ระยะห่างไม่แม่นยำ กรณีนี้ควรใช้ Distance Measure อื่น ๆ เช่น Manhattan Distance ซึ่งมีความเหมาะสมกว่า Euclidean Distance ในการวัดระยะห่างของข้อมูล

2.8.2 การวัดระยะทางแบบ Manhattan Distance

Manhattan Distance เป็นการวัดระยะห่างระหว่างสองจุดในพื้นที่หรือข้อมูล โดยวิธีการคำนวณคือการหาผลรวมของความต่างของระยะแนวตั้ง และแนวนอนของสองจุด โดยทั่วไปจะใช้ในการคำนวณระยะทางในข้อมูลที่เป็นตารางหรือเมตริกซ์ที่มีตัวแปรมากกว่าหนึ่งตัว เช่น การจัดกลุ่มของคลัสเตอร์หรือคำนวณระยะห่างของรูปภาพ ซึ่ง Manhattan Distance จะมีความแม่นยำเมื่อข้อมูลเป็นลักษณะข้อมูลที่อยู่บนเวกเตอร์ที่มีสัมบูรณ์ และแยกชัดเจน โดยสามารถเขียนสูตรคำนวณได้ดังนี้

$$\text{Manhattan Distance} = |x1 - x2| + |y1 - y2|$$

เมื่อ x1 และ y1 คือตำแหน่งของจุดแรกในระบบพิกัดแนวตั้ง และแนวนอน
x2 และ y2 คือตำแหน่งของจุดที่สองในระบบพิกัดแนวตั้ง และแนวนอน



ภาพที่ 5 การวัดระยะทางแบบ Manhattan Distance

2.9 หลักเกณฑ์ในการรวมกลุ่ม (Linkage Methods)

Linkage Methods เป็นวิธีหนึ่งในการจัดกลุ่มข้อมูล (Clustering) ซึ่งใช้เพื่อคำนวณค่าระยะห่างระหว่างคลัสเตอร์ โดยจะนำค่าระยะห่างที่คำนวณได้มาเป็นเกณฑ์ในการเลือกคู่คลัสเตอร์ที่จะรวมกันต่อไป มีหลายวิธีในการคำนวณระยะห่างระหว่างคลัสเตอร์ใน Linkage Methods เช่น

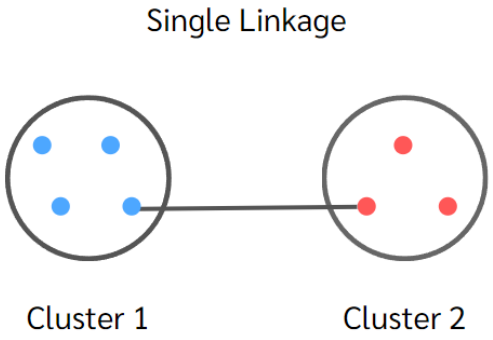
Single Linkage, Complete Linkage, Average Linkage, Ward's Method เป็นต้น โดยแต่ละวิธีจะมีวิธีการคำนวณค่าระยะห่างระหว่างคลัสเตอร์ต่างกันออกไป การเลือกวิธี Linkage ที่เหมาะสมจะช่วยให้การจัดกลุ่มข้อมูลของเรามีความแม่นยำ และมีความหมายต่อการวิเคราะห์ที่ต้องการใช้ข้อมูลในธุรกิจ วิทยาศาสตร์ หรืออื่นๆ อย่างมีประสิทธิภาพมากยิ่งขึ้น

2.9.1 หลักเกณฑ์ในการรวมกลุ่มแบบ Single Linkage Method

Single Linkage เป็นหนึ่งใน Linkage Methods ใน Hierarchical Clustering ซึ่งเป็นกระบวนการแบ่งกลุ่มข้อมูลโดยการใช้ความคล้ายคลึงของข้อมูล วิธีการ Single Linkage จะใช้ค่าระยะทางที่น้อยที่สุดระหว่างสมาชิกในกลุ่มที่แตกต่างกันเป็นเกณฑ์ในการจับคู่สมาชิก สามารถนำไปใช้ในงานวิจัยด้านการจัดกลุ่มลูกค้า หรือการจัดกลุ่มพันธุ์พืช และอื่นๆ อีกหลายงานด้วยกัน โดย Single Linkage Method เป็นวิธีที่รวมกลุ่ม 2 กลุ่มเข้าด้วยกันโดยพิจารณาจากระยะห่างที่สั้นที่สุด

$d_{mj} = \min (d_{kj}, d_{lj})$

ข้อกำหนด	**คำอธิบาย**
d_{mj}	ระยะห่างระหว่างกลุ่ม m และ j
m	คลัสเตอร์ที่ผสานซึ่งประกอบด้วยคลัสเตอร์ k และ l โดยที่ m = (k,i)
d_{kj}	ระยะห่างระหว่างกลุ่ม k และ j
d_{lj}	ระยะห่างระหว่างกลุ่ม l และ j



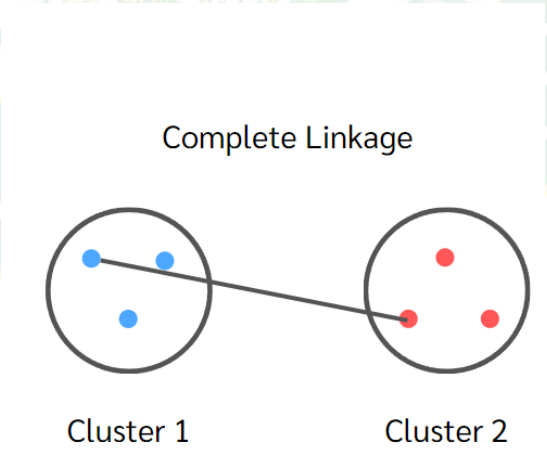
ภาพที่ 6 การรวมกลุ่มแบบ Single Linkage Method

2.9.2 หลักเกณฑ์ในการรวมกลุ่มแบบ Complete Linkage Method

Complete Linkage เป็นวิธีการทำ Hierarchical Clustering โดยใช้ค่าความคล้ายกันสูงสุด (Furthest Neighbor) ระหว่างคลัสเตอร์ที่จะรวมกัน โดยใช้ระยะทางที่ไกลที่สุดระหว่างสมาชิกสองกลุ่มเป็นค่าความคล้ายกันของคลัสเตอร์ ในขั้นตอนแรกทุกคลัสเตอร์จะถูกพิจารณาว่าเป็นกลุ่มตัวเอง จากนั้น จะทำการหาค่าความคล้ายกันของทุกคู่ของคลัสเตอร์แล้วรวมคลัสเตอร์ที่มีความคล้ายกันสูงที่สุดเข้าด้วยกันเป็นคลัสเตอร์ใหม่ และทำซ้ำจนกว่าจะได้จำนวนคลัสเตอร์ตามที่กำหนดหรือเป็นไปจนกว่าจะไม่มีการเชื่อมต่อได้อีกต่อไป โดย Complete Linkage Method วิธีนี้จะรวมกลุ่ม 2 กลุ่มเข้าด้วยกันโดยพิจารณาจากระยะห่างที่ยาวที่สุด

$d_{mj} = \max (d_{kj}, d_{lj})$

ตัวแปร	**คำอธิบาย**
d_{mj}	ระยะห่างระหว่างกลุ่ม m และ j
m	คลัสเตอร์ที่ผสานซึ่งประกอบด้วยคลัสเตอร์ k และ l โดยที่ m = (k,i)
d_{kj}	ระยะห่างระหว่างกลุ่ม k และ j
d_{lj}	ระยะห่างระหว่างกลุ่ม l และ j



ภาพที่ 7 การรวมกลุ่มแบบ Complete Linkage Method

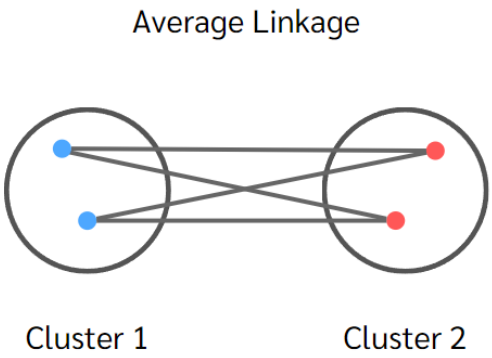
2.9.3 หลักเกณฑ์ในการรวมกลุ่มแบบ Average Linkage Method

Average Linkage เป็นวิธีหนึ่งใน Linkage Methods ที่ใช้ในการหาค่าระยะห่างระหว่างคลัสเตอร์สำหรับ Hierarchical Clustering โดยคำนวณค่าเฉลี่ยของค่าระยะห่างระหว่าง

ข้อมูลจากคลัสเตอร์ทั้งหมด วิธีนี้จะสร้างคลัสเตอร์ที่มีการกระจายของข้อมูลที่น้อยกว่าวิธี Single Linkage แต่อาจจะทำให้เกิดความคลาดเคลื่อนในการแบ่งกลุ่มได้ง่ายขึ้น ทำให้ต้องทำการวิเคราะห์ผลลัพธ์เพื่อเลือกวิธีที่เหมาะสมกับข้อมูลที่ใช้ในการวิเคราะห์ และจำเป็นต่อการแก้ไขปัญหาในการวิเคราะห์ข้อมูลในภายหลัง การคำนวณค่าระยะห่างจะใช้หลักการเดียวกับใน Single Linkage ดังนั้นการเลือกวิธี Linkage Method นั้นจะขึ้นอยู่กับการออกแบบ และวัตถุประสงค์ของการวิเคราะห์แต่ละครั้งด้วย โดย Average linkage method เป็นวิธีที่หาระยะห่างระหว่างสองกลุ่มโดยค่าเฉลี่ยของระยะทาง (average distance) ระหว่างการสังเกตุกลุ่มที่หนึ่งกับการสังเกตุกลุ่มอื่น

$$d_{mj} = \frac{N_k d_{kj} + N_l d_{lj}}{N_m}$$

ตัวแปร	**คำอธิบาย**
d_{mj}	ระยะห่างระหว่างกลุ่ม m และ j
m	คลัสเตอร์ที่ผสานซึ่งประกอบด้วยคลัสเตอร์ k และ l โดยที่ m = (k,i)
d_{kj}	ระยะห่างระหว่างกลุ่ม k และ j
d_{lj}	ระยะห่างระหว่างกลุ่ม l และ j
N_k	จำนวนการสังเกตในคลัสเตอร์ k
N_l	จำนวนการสังเกตในกลุ่ม l
N_m	จำนวนการสังเกตในคลัสเตอร์ m



ภาพที่ 8 การรวมกลุ่มแบบ Average Linkage Method

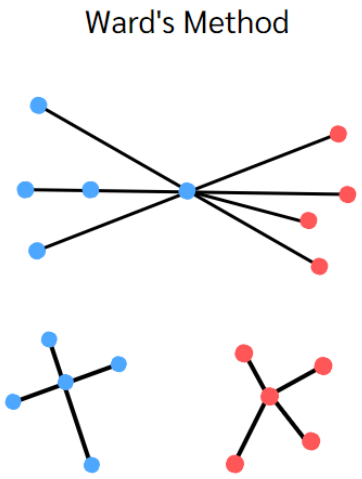
2.9.4 หลักเกณฑ์ในการรวมกลุ่มแบบ Ward’s Linkage Method

Ward linkage Method เป็นหนึ่งในวิธีการจัดกลุ่ม (clustering) ในการวิเคราะห์ข้อมูล ซึ่งใช้วิธีการคำนวณความคล้ายคลึงระหว่างกลุ่มโดยใช้วิธีการวัดระยะห่างระหว่างจุด (distance measure) เพื่อทำให้สามารถจัดกลุ่มได้ตามความคล้ายคลึงระหว่างข้อมูลที่กำหนดไว้ โดยจะใช้วิธีการคำนวณความคล้ายคลึงระหว่างกลุ่มโดยวิธีการหาค่าเฉลี่ยของความแตกต่างของจุดภายในกลุ่ม (within-cluster variance) จากนั้นนำค่าเฉลี่ยที่ได้มาวัดระยะห่างระหว่างกลุ่ม โดยจะเลือกกลุ่มที่มีค่า within-cluster variance น้อยที่สุดเป็นกลุ่มที่ใช้ในการรวมกลุ่มต่อไป จนกระทั่งได้กลุ่มเดียวเท่านั้น

Ward linkage Method เป็นวิธีการจัดกลุ่มที่มีความแม่นยำสูง และเหมาะสมสำหรับข้อมูลที่มีการกระจายตัวที่ไม่สม่ำเสมอ และมีขนาดของกลุ่มที่ไม่เท่ากัน

$$d_{mj} = \frac{(N_j + N_k)d_{kj} + (N_j + N_l)d_{lj} - N_j d_{kl}}{N_j + N_m}$$

ตัวแปร	**คำอธิบาย**
d_{mj}	ระยะห่างระหว่างกลุ่ม m และ j
m	คลัสเตอร์ที่ผสานซึ่งประกอบด้วยคลัสเตอร์ k และ l โดยที่ m = (k,i)
d_{kj}	ระยะห่างระหว่างกลุ่ม k และ j
d_{lj}	ระยะห่างระหว่างกลุ่ม l และ j
d_{kl}	ระยะห่างระหว่างกลุ่ม k และ l
N_j	จำนวนการสังเกตในคลัสเตอร์ j
N_k	จำนวนการสังเกตในคลัสเตอร์ k
N_l	จำนวนการสังเกตในคลัสเตอร์ l
N_m	จำนวนการสังเกตในคลัสเตอร์ m



ภาพที่ 9 การรวมกลุ่มแบบ Ward Linkage Method

2.9.5 หลักเกณฑ์ในการรวมกลุ่มแบบ Median Linkage Method

Median Linkage Method เป็นวิธีที่รวมกลุ่ม 2 กลุ่มเข้าด้วยกัน โดยให้แต่ละกลุ่มสำคัญเท่ากัน (ให้น้ำหนักเท่ากัน) Median Linkage Method จะใช้ค่ามัธยฐานเป็นค่ากลางของจุดศูนย์กลาง (Centroid) ถ้าระยะห่าง ระหว่างค่ามัธยฐานของ Clustering จะใช้ค่ามัธยฐานเป็นค่ากลางของจุดศูนย์กลาง (Centroid) ถ้าระยะห่าง ระหว่างค่ามัธยฐานของกลุ่ม คู่ใดต่ำจะรวมกลุ่มคู่นั้นเข้าด้วยกัน

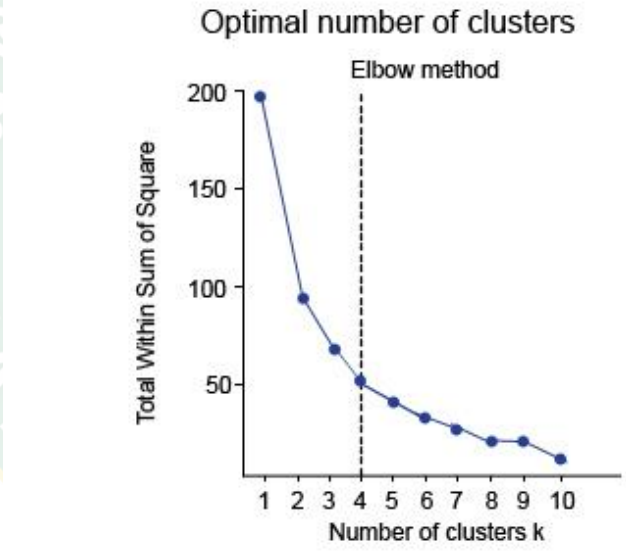
$$d_{mj} = \frac{d_{kj} + d_{lj}}{2} - \frac{d_{kl}}{4}$$

ตัวแปร	**คำอธิบาย**
d_{mj}	ระยะห่างระหว่างกลุ่ม m และ j
m	คลัสเตอร์ที่ผสานซึ่งประกอบด้วยคลัสเตอร์ k และ l โดยที่ m = (k,i)
d_{kj}	ระยะห่างระหว่างกลุ่ม k และ j
d_{lj}	ระยะห่างระหว่างกลุ่ม l และ j
d_{kl}	ระยะห่างระหว่างกลุ่ม k และ l

2.10 การกำหนดคลัสเตอร์ที่เหมาะสม (Determining Optimal Cluster)

การกำหนดจำนวนคลัสเตอร์ที่เหมาะสมที่สุดสำหรับชุดข้อมูลเป็นขั้นตอนสำคัญในกระบวนการจัดกลุ่มข้อมูล จำนวนคลัสเตอร์ที่เหมาะสมที่สุดคือกลุ่มที่แสดงถึงโครงสร้างและความสัมพันธ์ในข้อมูลได้ดีที่สุด มีหลายวิธีในการกำหนดจำนวนคลัสเตอร์ที่เหมาะสม ได้แก่

1) Elbow method เป็นวิธีการประเมินจำนวนคลัสเตอร์ในการทำ clustering ด้วยใช้ Within-Cluster-Sum-of-Squares (WCSS) เป็นตัววัดประสิทธิภาพ โดยหาค่า WCSS ของคลัสเตอร์ของ k จำนวน และพล็อตกราฟ WCSS ของ k เป็นแกน y และ k เป็นแกน x แล้วเราจะพบว่ามีจุดที่ WCSS เริ่มต้นเพิ่มขึ้น และจะมีจุดที่ WCSS กระตุก (elbow) และหลังจากนั้น WCSS เพิ่มขึ้นอย่างเร็วขึ้นไม่มาก จำนวน k ที่แสดง elbow นั้นเป็นจำนวนคลัสเตอร์ที่เหมาะสม ดังตัวอย่างภาพที่ 10



ภาพที่ 10 การหาจำนวนกลุ่มที่เหมาะสมด้วย Elbow Method

2) Silhouette: วิธีนี้เกี่ยวข้องกับการคำนวณความคล้ายคลึงกันระหว่างจุดข้อมูลแต่ละจุดและคลัสเตอร์ของตัวเอง ตลอดจนความคล้ายคลึงกันระหว่างจุดข้อมูลแต่ละจุด และคลัสเตอร์ที่ใกล้เคียงที่สุด จำนวนคลัสเตอร์ที่เหมาะสมจะกำหนดโดยการเลือกจำนวนคลัสเตอร์ที่มีคะแนนเงาเฉลี่ยสูงสุด

3) Gap Statistic: วิธีนี้เกี่ยวข้องกับการเปรียบเทียบผลรวมกำลังสองภายในคลัสเตอร์ที่สังเกตได้กับผลรวมกำลังสองภายในคลัสเตอร์ที่คาดไว้ภายใต้การแจกแจงอ้างอิงค่าว่าง และเลือกจำนวนคลัสเตอร์ที่มีช่องว่างมากที่สุด

สำหรับวิธีการเฉพาะสำหรับการกำหนดจำนวนคลัสเตอร์ที่เหมาะสมที่สุดอาจแตกต่างกันไปขึ้นอยู่กับลักษณะของข้อมูล และเป้าหมายของการวิเคราะห์ สิ่งสำคัญคือต้องพิจารณาการแลกเปลี่ยนระหว่างการใช้เกินพอดี และน้อยเกินไป ตลอดจนความสามารถในการตีความของผลลัพธ์

2.11 Sum of Squared Errors (SSE)

ผลรวมของข้อผิดพลาดกำลังสอง (SSE) คือการวัดความแตกต่างระหว่างค่าที่สังเกตได้ และค่าที่คาดการณ์ในชุดข้อมูล เป็นเมตริกการประเมินที่ใช้กันทั่วไปในอัลกอริทึมการจัดกลุ่มข้อมูล และการเรียนรู้ของเครื่อง โดยเฉพาะอย่างยิ่งในอัลกอริทึมการจัดกลุ่ม เช่น K-Means

สำหรับการจัดกลุ่ม K-Means SSE ถูกกำหนดเป็นผลรวมของระยะทางกำลังสองระหว่างจุดข้อมูลแต่ละจุดกับเซนทรอยด์ที่ใกล้ที่สุด โดยที่เซนทรอยด์เป็นจุดศูนย์กลางของคลัสเตอร์ เป้าหมายคือการลด Sum of Squared Errors (SSE) เพื่อให้มีโครงสร้างคลัสเตอร์ที่แน่นขึ้น และการแสดงความสัมพันธ์ภายในข้อมูลได้แม่นยำยิ่งขึ้น โดยสามารถใช้ Sum of Squared Errors (SSE) เพื่อกำหนดจำนวนคลัสเตอร์ที่เหมาะสมที่สุดได้โดยการพล็อตค่า SSE เทียบกับจำนวนคลัสเตอร์ และเลือกจำนวนคลัสเตอร์ที่การเปลี่ยนแปลงใน Sum of Squared Errors (SSE) เริ่มลดระดับลง (หรือที่เรียกว่า "จุดหักศอก") ค่า Sum of Squared Errors (SSE) ที่ต่ำกว่าบ่งชี้ว่าโมเดลเข้ากับข้อมูลได้ดีกว่า เนื่องจาก Sum of Squared Errors (SSE) วัดค่าความเบี่ยงเบนของจุดข้อมูลจากศูนย์กลางคลัสเตอร์

สำหรับ Sum of Squared Errors (SSE) เป็นเพียงหนึ่งในหลาย ๆ เมตริกที่สามารถใช้เพื่อประเมินประสิทธิภาพของอัลกอริทึมการจัดกลุ่ม และควรใช้ร่วมกับเมตริกอื่นๆ เช่น คะแนน Silhouette หรือดัชนี Calinski-Harabasz เพื่อรับการประเมินที่ครอบคลุมยิ่งขึ้นของผลลัพธ์

2.12 อัตราการเปลี่ยนแปลงของ Sum of Squared Errors (SSE) หรือ Relative Change of SSE (%Change)

อัตราการเปลี่ยนแปลงของ Sum of Squared Errors (SSE) คือความแตกต่างสัมพัทธ์ในค่า SSE ระหว่างสองขั้นตอนของอัลกอริทึมหรือโมเดล ซึ่งแสดงเป็นเปอร์เซ็นต์ของค่า Sum of Squared Errors (SSE) เริ่มต้น ในบริบทของการจัดกลุ่มข้อมูล และอัลกอริทึม เปอร์เซ็นต์การเปลี่ยนแปลงใน

SSE จะถูกใช้เพื่อประเมินการปรับปรุงในโครงสร้างคลัสเตอร์เมื่อจำนวนคลัสเตอร์เพิ่มขึ้น เพื่อดูอัตราการเปลี่ยนแปลงของ SSE ตั้งต้น และ SSE ลำดับถัดไปในอัลกอริทึมหรือโมเดล

สำหรับอัตราการเปลี่ยนแปลงของ SSE จะใช้เพื่อประเมินว่าการปรับปรุงโครงสร้างของคลัสเตอร์เมื่อเพิ่มจำนวนคลัสเตอร์จะทำให้ค่า SSE ลดลงเท่าไหร่ ซึ่งเป็นข้อมูลที่ใช้ในการหาจำนวนคลัสเตอร์ที่เหมาะสมสำหรับชุดข้อมูลที่กำลังจะจัดกลุ่ม การเปรียบเทียบอัตราการเปลี่ยนแปลงของ SSE จะช่วยให้เราวัดประสิทธิภาพของอัลกอริทึมหรือโมเดลได้ดียิ่งขึ้น และตัดสินใจเลือกใช้อัลกอริทึมหรือโมเดลที่ดีที่สุดสำหรับงานการจัดกลุ่มข้อมูลนั้น ๆ

2.13 รูปแบบการเรียนรู้ Index of Learning Styles (ILS)

Felder-Silverman Learning Style Model หรือที่รู้จักในชื่อ Index of Learning Styles (ILS) เป็นโมเดลที่จัดประเภทนักเรียนออกเป็นสี่รูปแบบการเรียนรู้ที่แตกต่างกัน: ประสาทสัมผัส การหยั่งรู้ ภาพ และคำพูด แบบจำลองนี้มีพื้นฐานมาจากแนวคิดที่ว่าบุคคลมีความชอบที่แตกต่างกันในการรับ และประมวลผลข้อมูล ซึ่งสามารถใช้เพื่อช่วยให้ผู้สอนเข้าใจรูปแบบการเรียนรู้ที่นักเรียนชื่นชอบได้ดีขึ้น และปรับวิธีการสอนให้เหมาะกับจุดแข็ง และความชอบของแต่ละคนได้ดียิ่งขึ้นตามลักษะรูปแบบการเรียนรู้ ดังตารางที่ 1 และรูปภาพที่ 11

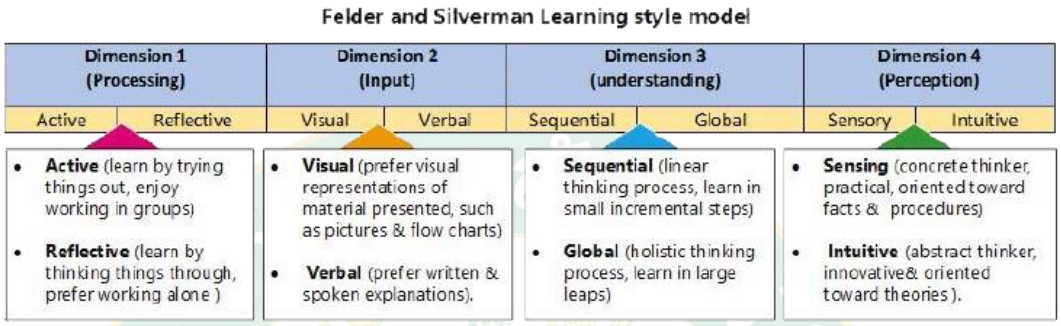
ประเภทของรูปแบบทั้งสี่มิติ (Dimension) ประกอบไปด้วย

1. Sensing - Intuitive เป็นตัวกำหนดรูปแบบการรับรู้หรือรับข้อมูล
2. Visual - Verbal เป็นตัวกำหนดของรูปแบบการนำเสนอข้อมูล
3. Active - Reflective เป็นตัวกำหนดรูปแบบการประมวลผลข้อมูล
4. Sequential - Global เป็นตัวกำหนดรูปแบบการจัดระเบียบ และการทำความเข้าใจข้อมูล

ตารางที่ 1 ตารางอธิบายลักษณะของรูปแบบการเรียนรู้แบบ Felder Silverman

ประเภทของผู้เรียน	คำอธิบาย
Sensing	ชอบความคิดที่เป็นรูปธรรม ใช้ได้จริง เกี่ยวข้องกับข้อเท็จจริง และขั้นตอน
Intuitive	ชอบการคิดเชิงมโนทัศน์ สร้างสรรค์ เกี่ยวข้องกับทฤษฎี และความหมาย
Visual	ชอบการแสดงภาพ รูปภาพ ไดอะแกรม และผังงาน
Verbal	ชอบการอธิบายที่เป็นลายลักษณ์อักษร และการพูด

Active	ชอบทดลองทำสิ่งต่าง ๆ ทำงานร่วมกับผู้อื่นเป็นกลุ่ม
Reflective	ชอบคิดอะไรไปเรื่อยเปื่อย ทำงานคนเดียว หรือกับคู่หูที่คุ้นเคย
Sequential	ชอบการคิดเชิงเส้น มีระเบียบ เรียนรู้ทีละน้อย
Global	ชอบการคิดแบบองค์รวม นักคิดเชิงระบบ เรียนรู้แบบก้าวกระโดด



ภาพที่ 11 ลักษณะรูปแบบการเรียนรู้ Felder-Silverman Learning Preferences

2.14 วรรณกรรมที่เกี่ยวข้องกับงานวิจัย

A Novel Algorithm for Course Learning Object Recommendation Based on Student Learning Styles (Nafea และคณะ, 2019) ได้ศึกษาอัลกอริทึมสำหรับการแนะนำหลักสูตรการเรียนรู้ตามรูปแบบการเรียนรู้ของนักเรียน เพื่อแก้ปัญหาในการค้นหาแหล่งเรียนรู้ที่เหมาะสมกับระบบอีเลิร์นนิง เพื่อให้ตรงกับรูปแบบการเรียนรู้ ระบบแนะนำข้อเสนอส่วนบุคคล และถ่ายทอดวัตถุประสงค์การเรียนรู้ที่เหมาะสมกับความชอบของนักเรียน และพิจารณาทดลองเพื่อตรวจสอบตัวชี้วัดความคล้ายคลึงกันที่ดีที่สุด ที่จะใช้ในระบบผู้แนะนำสำหรับการเรียนรู้ วิธีการจะขึ้นอยู่กับรูปแบบการเรียนรู้สไตล์ Felder และ Silverman ซึ่งใช้เพื่อเป็นตัวแทนรูปแบบการเรียนรู้ของนักเรียน และวัตถุประสงค์การเรียนรู้ โดยใช้ Skeleton Recommender Algorithm ออกแบบเมทริกที่คล้ายกัน คือ Sim1 และ Sim2 จากนั้นทดลองหาเมทริกที่มีความเหมือนจริงที่สุดสำหรับ Sim1 และ Sim2 จากกลุ่มความเหมือนกันทั่วไปที่ใช้ในระบบแนะนำ จากนั้นใช้ K-mean เพื่อสร้าง K กลุ่มของการเรียนรู้ตามวัตถุประสงค์การเรียนรู้ของแต่ละโปรไฟล์ ที่คล้ายคลึงเมททริกSim1 โดยใช้เพื่อค้นหากลุ่มการเรียนรู้ที่ใกล้ที่สุดกับรูปแบบการเรียนรู้ของนักเรียนจากนั้น Sim2จะถูกใช้เพื่อกำหนดวัตถุประสงค์ของการเรียนรู้อันดับต้นๆในกลุ่มนั้นซึ่งมีลักษณะคล้ายกับรูปแบบการเรียนรู้ของนักเรียนที่กระตือรือร้นที่สุด และคะแนนการทำนายของของวัตถุประสงค์การเรียนรู้สำหรับนักเรียน

จะถูกคำนวณโดยใช้ Sim2 พบว่าอัลกอริทึมทำงานได้ดีที่สุดสำหรับ Sim1 คือ cosine และSim2 คือ Pearson correlation coefficient

Detection of learning styles in the focus group (Švarcová และJelínková, 2016) ได้ศึกษาการตรวจจับรูปแบบการเรียนรู้ของนักศึกษามหาวิทยาลัย เพื่อค้นหาว่ารูปแบบการเรียนรู้แบบใดที่นักศึกษามหาวิทยาลัยชื่นชอบ โดยเฉพาะในยุคปัจจุบันที่เต็มไปด้วยเทคโนโลยีที่ทันสมัย งานวิจัยใช้การประเมินเชิงปริมาณของผลการวิจัยจากแบบสอบถามมาตรฐาน ILS (ดัชนีรูปแบบการเรียนรู้) ของ Felder และ Silverman รูปแบบการเรียนรู้นั้นฝังแน่นอยู่ในสภาพทางชีวภาพของแต่ละบุคคล ดังนั้นจึงไม่ง่ายที่จะเปลี่ยนแปลง ซึ่งเหมาะสมที่จะปรับกระบวนการสอนให้เข้ากับรูปแบบการเรียนรู้ของแต่ละบุคคลให้ตรงตามรูปแบบการเรียนรู้โดยเฉพาะแต่ต้องทราบถึงการรู้รูปแบบการเรียนรู้ของนักศึกษาแต่ละคนก่อน กระบวนการสอนจะต้องปรับให้เข้ากับความรู้เหล่านี้เพื่อให้ผลการเรียนรู้มีประสิทธิภาพมากที่สุด

E-Learning personalization based on hybrid recommendation strategy and learning style identification (Klašnja-Milićević และคณะ, 2011) ได้ศึกษาการออกแบบประสบการณ์การศึกษาในระบบอีเลิร์นนิงให้เหมาะสมกับความต้องการและเป้าหมายของผู้เรียน ศึกษาเทคนิคต่างๆ เพื่อช่วยอาจารย์ปรับปรุงระบบอีเลิร์นนิง โดยโมดูลข้อเสนอแนะของระบบการสอน Protus ระบบรู้จักรูปแบบการเรียนรู้ที่แตกต่าง และนิสัยของผู้เรียนบันทึกในเชิร์ฟเวอร์ และประมวลผลกลุ่มตามรูปแบบการเรียนรู้ที่แตกต่างกันกำหนดจากเนื้อหาการเรียนรู้ตามทฤษฎี ILS จากนั้นใช้ AppioriAll เพื่อค้นหาลำดับของรูปแบบในการวิเคราะห์นิสัย และความสนใจของผู้เรียน ผลการทดลองแสดงให้เห็นว่าการรวมกันของการทดสอบรูปแบบการเรียนรู้ของผู้เรียน และการขุดลำดับที่บ่อยในเว็บด้วยอัลกอริทึม AprioriAll มีความเหมาะสมในการใช้รูปแบบการแนะนำ เพื่อแนะนำกิจกรรมการเรียนรู้ออนไลน์กับผู้เรียนตามรูปแบบการเรียนรู้ และความชอบ

การทำเหมืองข้อมูล: แนวคิดและเทคนิค (Han และคณะ, 2011) การอธิบายหลักการและเทคนิคของการทำเหมืองข้อมูล โดยเริ่มต้นจากการอธิบายหลักการพื้นฐาน และหลักการต่าง ๆ ของการทำเหมืองข้อมูล เช่น การแยกแยะประเภทของข้อมูล การเตรียมข้อมูล การสกัดคุณลักษณะ การลดมิติ การกรอง และการจัดกลุ่มข้อมูล โดยเน้นเรื่องของการจัดกลุ่มข้อมูล (clustering) และการทำนาย (prediction) โดยรวมถึงการใช้เครื่องมือทางคณิตศาสตร์ และสถิติ เพื่อช่วยในการทำเหมืองข้อมูล ซึ่งมีการอธิบายขั้นตอน และเทคนิคของการใช้งานของอัลกอริทึมต่าง ๆ ที่สามารถใช้ในการทำ

เหมืองข้อมูลรวมถึงการทำ clustering ด้วย k-means และ hierarchical clustering และการทำ predictive modeling ด้วย regression และ decision tree

การประยุกต์ใช้เทคโนโลยีกลุ่มเพื่อตรวจสอบรูปแบบการเรียนการสอนของนักศึกษาวิศวกรรมศาสตร์ (Abdelhadi, 2017) ได้ศึกษาการจำแนกนักเรียนวิศวกรรมในห้องเรียน โดยการแบ่งกลุ่มตามความต้องการรูปแบบการเรียนรู้ของ Felder and Silverman ร่วมกับแนวคิดเทคโนโลยีกลุ่ม (Group Technology) มาใช้กับวิธีการสอนที่ถูกต้องเพื่อใช้เกิดประโยชน์สูงสุดจากเป้าหมายการสอน โดยเทคโนโลยีกลุ่ม เป็นวิธีการที่ทำให้ทราบถึงความคล้ายคลึงกันของคุณลักษณะของวัตถุที่รวมกันเป็นกลุ่ม พบว่าการใช้เทคโนโลยีกลุ่มในช่วงต้นภาคการศึกษาจะช่วยให้ผู้สอนใช้รูปแบบการสอนที่เหมาะสมกับนักเรียนเพื่อบรรลุการสอนที่ต้องการ นอกจากนี้ผู้สอนยังสามารถสังเกตพฤติกรรมของนักเรียนเป็นกลุ่มเพื่อเตรียมเนื้อหาตามหลักสูตร และการมอบหมายงานให้กับนักเรียนสามารถทำเพื่อให้เหมาะสมกับรูปแบบการเรียนรู้ของนักเรียนในกลุ่ม และสามารถเปลี่ยนแปลงได้เสมอตามกลุ่มที่เกิด

การตรวจสอบความเหมาะสมของการจัดกลุมข้อมูลอนุกรมเวลา (สูงเนิน และคณะ, 2018) การตรวจสอบความเหมาะสมของการจัดกลุ่มข้อมูลอนุกรมเวลาในชุดข้อมูล โดยใช้วิธีการตัดสินใจแบบเชิงกฎกำกับ (decision tree) และวิธีการกระจายตัวแบบกฎกำกับ (rule-based clustering) โดยการใช้เครื่องมือตรวจสอบความเหมาะสมในการจัดกลุ่ม โดยการพิจารณาค่าผลรวมความผิดพลาด (Sum of Squared Error: SSE) มาใช้กับกรณีการจัดกลุ่มแบบ Unsupervised Learning อีกทั้งยังใช้กราฟแสดงความสัมพันธ์ระหว่างค่า k และค่า SSE เพื่อหาจุดเปลี่ยนความชันหรือ %Change ในการหาจำนวนกลุ่มที่เหมาะสม

Clustering analysis of learning style on anggana high school student (Lailiyah และคณะ, 2019) ได้ศึกษาเรื่องการไร้ความสามารถของนักเรียนในการรับความรู้ที่ถ่ายทอดโดยผู้สอนนั้นไม่ได้เกิดจากความไม่เข้าใจ และผู้สอนไม่สามารถสอนได้เช่นกัน แต่เป็นเพราะความไม่ตรงกันของทักษะการเรียนรู้ระหว่างนักเรียน และอาจารย์ เพื่อให้นักเรียนรู้สึกไม่สบายใจกับผู้สอนท่านใดท่านหนึ่ง ดังนั้นจึงจำเป็นที่จะต้องทำการวิจัยนี้เพื่อวิเคราะห์กลุ่มของรูปแบบการเรียนรู้ของนักเรียนด้วยวิธี K-Means และFuzzy C-Means ซึ่งมีวัตถุประสงค์เพื่อให้ทราบประสิทธิผลของการเรียนรู้แบบกลุ่มต่อการพัฒนาการดูดซับ และการพัฒนาผลสัมฤทธิ์ทางการเรียนของนักเรียน วิธีการที่ใช้ในการจัดรูปแบบการเรียนรู้ด้วยกระบวนการเหมืองข้อมูลในกระบวนการทำงาน และพัฒนาความรู้ พบว่าการจำแนกประเภทของรูปแบบการเรียนรู้ของนักเรียนโดยใช้ K-Means และ FCM สามารถแบ่ง

ออกเป็น 4 กลุ่ม นักเรียนหลายคนชอบที่จะเรียนรู้ด้วยการฟังซึ่งช่วยในการมองเห็น มากกว่าการเรียนรู้เพียงแค่อ่านการฝึกฝนด้วยตนเอง ข้อสรุปนี้มาจากการรวมกลุ่มของนักเรียนที่ชื่นชอบการเรียนรู้ด้วยเสียง และสื่อ การวิจัยนี้สามารถช่วยอาจารย์ผู้สอนค้นหาวิธีการสอนที่ถูกต้องให้กับนักเรียนในชั้นเรียน

Learning and teaching styles in engineering education (Felder และ Silverman, 1988) ได้ศึกษาเกี่ยวกับรูปแบบการเรียนรู้ของนักศึกษาในหลากหลายรูปแบบ ซึ่งจำนวนนักเรียนที่ได้รับการเรียนรู้ในชั้นเรียนนั้นถูกควบคุมโดยผู้สอน และการเตรียมตัวล่วงหน้า แต่ยังอยู่กับรูปแบบการเรียนรู้ของนักเรียนหรือรูปแบบการสอนของผู้สอน ความไม่ตรงกันของรูปแบบการเรียนรู้ทั่วไปของนักศึกษาวิศวกรรม และรูปแบบการสอนแบบดั้งเดิมของอาจารย์วิศวกรรม พบว่ารูปแบบการเรียนรู้ของนักศึกษาวิศวกรรมส่วนใหญ่และรูปแบบการสอนของอาจารย์วิศวกรรมส่วนใหญ่ไม่เข้ากันในหลายมิติ นักศึกษาวิศวกรรมส่วนใหญ่เป็นนักเรียนที่มองเห็นได้สัมผัสอุปนัย และกระตือรือร้น และนักเรียนที่มีความคิดสร้างสรรค์ ความไม่ตรงกันเหล่านี้นำไปสู่ประสิทธิภาพการทำงานของนักเรียนที่ไม่ดี และการสูญเสียต่อสังคมของวิศวกรที่ยอดเยี่ยม แม้ว่ารูปแบบที่หลากหลายที่นักเรียนเรียนรู้จะมีมากมาย แต่การรวมกันของจำนวนค่อนข้างน้อย เทคนิคในวิทยาการของผู้สอนควรเพียงพอที่จะตอบสนองความต้องการของนักเรียนส่วนใหญ่หรือทั้งหมดในชั้นเรียนใด ๆ เทคนิค และคำแนะนำที่ให้ไว้ในหน้านี้ควรตอบสนองวัตถุประสงค์นี้ อย่างไรก็ตามความคิดนี้ไม่ได้ใช้เทคนิคทั้งหมดในทุกชั้นเรียน แต่เลือกที่จะเป็นไปได้ และลองใช้ ด้วยวิธีนี้รูปแบบการสอนที่มีประสิทธิภาพสำหรับนักเรียน และความสะดวกสบายสำหรับอาจารย์จะมีการพัฒนาอย่างเป็นธรรมชาติ และไม่เจ็บปวดโดยมีผลกระทบอย่างมากต่อคุณภาพการเรียนรู้ที่เกิดขึ้นในภายหลัง

Learning styles and strategies (Felder และ Soloman, 2000) บทความนี้เน้นการเรียนรู้และกลยุทธ์ในการสอนที่สอดคล้องกับความชอบ และรูปแบบการเรียนรู้ของนักเรียน โดยใช้รูปแบบการเรียนรู้ของ Felder-Silverman ที่แบ่งนักเรียนออกเป็นสี่มิติ: sensory/intuitive, visual/verbal, active/reflective, และ sequential/global โดยกลยุทธ์การเรียนรู้ที่แตกต่างกัน เช่น การเรียนรู้แบบกระตือรือร้น การเรียนรู้แบบร่วมมือ และการเรียนรู้โดยใช้ปัญหาเป็นฐาน จะมีประสิทธิผลสำหรับรูปแบบการเรียนรู้ที่แตกต่างกัน เพื่อปรับปรุงผลการเรียน อาจารย์ควรตระหนักถึงความหลากหลายของรูปแบบการเรียนรู้ และใช้วิธีการสอน และกลยุทธ์ที่หลากหลายซึ่งสามารถรองรับรูปแบบการเรียนรู้ที่แตกต่างกันได้ การประเมินด้วยดัชนีรูปแบบการเรียนรู้ (ILS) เป็นวิธีที่สามารถใช้ระบุความชอบในการเรียนรู้ของนักเรียน และปรับการสอนให้เหมาะสม โดยรวมแล้วการ

ทำความเข้าใจ และการจัดรูปแบบการเรียนรู้ และกลยุทธ์ของนักเรียนเป็นสิ่งสำคัญที่ช่วยปรับปรุงผลการเรียนของนักเรียน และสร้างส่วนร่วมในการเรียนรู้

Applications, reliability and validity of the index of learning styles (Felder และ Spurlin, 2005) บทความนี้มุ่งเสนอข้อมูลเกี่ยวกับการใช้ดัชนีรูปแบบการเรียนรู้ (Index of Learning Styles: ILS) เครื่องมือวัดแบบทางจิตวิทยาที่ช่วยให้ผู้ใช้สามารถระบุรูปแบบการเรียนรู้ของตนเองได้ โดยมุ่งเน้นที่การใช้ ILS ในการวิจัย และการใช้งานจริง โดยมีเป้าหมายเพื่ออธิบายความสอดคล้องของ ILS กับการเรียนรู้ และการพิจารณาความน่าเชื่อถือ และความพึงพอใจในการใช้ ILS ในการวิจัย การเรียนรู้ของนักเรียนมักขึ้นอยู่กับความสอดคล้องระหว่างรูปแบบการเรียนรู้ของนักเรียนกับสิ่งที่ถูกสอน และกล่าวถึงแต่ละมิติของ ILS ที่ประกอบด้วย 1) sensory/intuitive, 2) visual/verbal, 3) active/reflective, และ 4) sequential/global ซึ่งการวัด ILS จะช่วยให้ผู้ใช้สามารถระบุประเภทนักเรียนแต่ละประเภทได้อย่างชัดเจน และช่วยให้อาจารย์ปรับการสอนให้เหมาะสมกับรูปแบบการเรียนรู้ของนักเรียนได้

A psychometric study of the index of learning styles© (Litzinger และคณะ, 2007) นำเสนอการศึกษาคุณภาพของดัชนีรูปแบบการเรียนรู้ (Index of Learning Styles: ILS) โดยใช้ข้อมูลจากนักเรียนในระดับประถมศึกษา มัธยมศึกษาตอนต้น และมัธยมศึกษาตอนปลายของสหรัฐอเมริกา ผลการวิจัยพบว่า ILS มีความน่าเชื่อถือ และความถูกต้องทางประชากรวิทยา และมีความเหมาะสมในการใช้วัดรูปแบบการเรียนรู้ของนักเรียน แต่อย่างไรก็ตามผลการวิจัยยังไม่สามารถสนับสนุนข้อสมมติฐานที่ว่า ILS สามารถใช้ในการวัดรูปแบบการเรียนรู้ของนักเรียนทุกประเภทได้ โดยเฉพาะอย่างยิ่งสำหรับนักเรียนที่มีพฤติกรรมการเรียนรู้ที่แตกต่างกันอย่างมาก การศึกษาเหล่านี้เป็นประโยชน์สำหรับผู้ใช้ ILS ในการวัดรูปแบบการเรียนรู้ของนักเรียน โดยการวิจัยนี้จะช่วยให้ผู้ใช้สามารถปรับปรุงการใช้งาน ILS ให้เหมาะสม และมีประสิทธิภาพต่อการวัดและปรับปรุงการสอนให้เหมาะสมกับความต้องการของนักเรียนได้

Data classification: algorithms and applications (Aggarwal, 2014) อธิบาย และวิเคราะห์ของอัลกอริทึมต่างๆ ที่ใช้ในการจัดกลุ่มข้อมูล (data classification) และการประยุกต์ใช้งานในสายงานต่างๆ โดยเฉพาะอย่างยิ่งในการวิเคราะห์ข้อมูล และสถิติ เช่น การวิเคราะห์ และจัดกลุ่มข้อมูลโดยใช้การเรียนรู้แบบจำลอง (model-based learning) โดยเน้นที่การใช้แนวคิดของการแยกกลุ่มข้อมูลด้วยความสามารถของโมเดลการเรียนรู้เชิงกฎ (rule-based learning) และการใช้เทคนิคของการจัดกลุ่มข้อมูลแบบเชิงพหุนาม (clustering) และการแบ่งกลุ่มข้อมูลแบบต้นไม้

(decision trees) เพื่อจัดกลุ่มข้อมูลให้อยู่ในกลุ่มเดียวกันตามลักษณะการแบ่งกลุ่มที่ต้องการ และการใช้แนวคิดของการจัดกลุ่มข้อมูลด้วยเทคนิคการจัดแยกระดับ (hierarchical clustering) และการจัดกลุ่มข้อมูลแบบความสัมพันธ์ (association rule mining) ที่เน้นการจัดกลุ่มข้อมูลตามลักษณะการมีความสัมพันธ์กันระหว่างข้อมูลต่าง ๆ เป็นต้น

Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system (Cha และคณะ, 2006) การพัฒนาระบบสอนอัจฉริยะ (Intelligent Tutoring System) ที่สามารถปรับแต่งส่วนติดต่อผู้ใช้ (User Interface) โดยอัตโนมัติตามรูปแบบการเรียนรู้ของผู้ใช้งาน โดยใช้วิธีการวินิจฉัยสไตล์การเรียนรู้ (Learning Style Diagnosis) จากพฤติกรรมการใช้งานของผู้ใช้งาน โดยผู้วิจัยได้พัฒนาระบบสอนอัจฉริยะโดยใช้วิธีการติดตาม และบันทึกการใช้งาน (User Interaction Tracking) ของผู้ใช้งาน โดยการติดตามจะใช้การวัดเวลาที่ผู้ใช้ใช้กับแต่ละส่วนของส่วนติดต่อผู้ใช้ (User Interface) และการวัดปริมาณการใช้งานของแต่ละส่วน โดยผลการติดตามจะถูกนำมาใช้ในการวินิจฉัยสไตล์การเรียนรู้ของผู้ใช้ โดยใช้เครื่องมือสอบถามเพิ่มเติมเกี่ยวกับลักษณะการเรียนรู้ของผู้ใช้

Data clustering: 50 years beyond K-means (Jain, 2010) การจัดกลุ่มข้อมูลที่ทำให้มีความแม่นยำ และเป็นไปตามหลักการ รวมถึงการอธิบายเทคนิคใหม่ ๆ ที่อาศัยการประมวลผลแบบพร้อมกัน และการจัดกลุ่มข้อมูลแบบซ้อนทับ งานวิจัยจึงสรุปได้ว่าการจัดกลุ่มข้อมูลเป็นกระบวนการที่สำคัญ และมีความสำคัญในการวิเคราะห์ข้อมูล โดยการจัดกลุ่มข้อมูลนั้นมีอยู่หลายวิธีการที่แตกต่างกันในเชิงวิธีการใช้งาน และความแม่นยำ แต่ว่างานวิจัยดังกล่าวได้ชี้ให้เห็นถึงความสำคัญของการตรวจสอบ และประเมินวิธีการจัดกลุ่มข้อมูลเพื่อที่จะเลือกวิธีการที่เหมาะสมในการใช้งาน นอกจากนี้ยังเสนอเทคนิคต่างๆ ที่สามารถนำมาใช้ในการจัดกลุ่มข้อมูลได้ เช่น การปรับแต่งฟังก์ชันค่าความเหมือนของข้อมูล (similarity function), การใช้คลัสเตอร์ประสาทเทียม (artificial neural network) และการเลือกเหตุการณ์ที่เหมาะสม

A comparative efficiency of clustering using dynamic feature selection optimization of subspace clustering algorithms (Pimpaporn และMeesad, 2014) งานวิจัยนี้มีเป้าหมายเพื่อเปรียบเทียบประสิทธิภาพของการแบ่งกลุ่มข้อมูลโดยใช้การเลือกลักษณะ (feature selection) แบบไดนามิก ของอัลกอริทึมแบบการแบ่งส่วนของข้อมูล (subspace clustering algorithms) แต่ละอัลกอริทึม เพื่อหาวิธีที่มีประสิทธิภาพมากที่สุดในการแบ่งกลุ่มข้อมูล โดยมีเป้าหมายเพื่อลดความซับซ้อนของการแบ่งกลุ่มข้อมูล และเพิ่มประสิทธิภาพของการแบ่งกลุ่มข้อมูล

ผู้วิจัยใช้การประเมินผลการแบ่งกลุ่มโดยใช้ค่า Silhouette index และ DB index โดยเปรียบเทียบค่าผลลัพธ์กับอัลกอริทึมแบบธรรมดา และอัลกอริทึมแบบปรับแต่ง (dynamic feature selection optimization) ซึ่งเป็นแนวทางที่มีการเลือกลักษณะแบบไดนามิก และผลการวิจัยพบว่าการใช้การแบ่งส่วนของข้อมูลแบบปรับแต่งสามารถลดความซับซ้อนของการแบ่งกลุ่มข้อมูล และเพิ่มประสิทธิภาพในการแบ่งกลุ่มข้อมูลได้ในระดับสูงสุดเมื่อเทียบกับการใช้อัลกอริทึมแบบธรรมดา อีกทั้งยังแสดงให้เห็นถึงความสามารถในการลดขนาดของชุดข้อมูลโดยการเลือกลักษณะแบบไดนามิก ทำให้การแบ่งกลุ่มข้อมูลเป็นไปได้อย่างมีประสิทธิภาพ และเหมาะสมกับการประมวลผล

The clustering validity with silhouette and sum of squared errors (Thinsungnoena และคณะ, 2015) การเปรียบเทียบความสอดคล้องของการจัดกลุ่มข้อมูลโดยใช้วิธี silhouette และ sum of squared errors (SSE) ซึ่งพบว่าการใช้วิธี silhouette ให้ผลลัพธ์ที่ดีกว่า SSE ในการประเมินคุณภาพของการจัดกลุ่ม โดยมีค่า silhouette สูงขึ้นตามจำนวนกลุ่มที่เพิ่มขึ้น ในขณะที่ SSE มีค่าต่ำลงตามจำนวนกลุ่มที่เพิ่มขึ้น แต่การใช้ SSE ยังเป็นวิธีที่นิยมใช้ในการจัดกลุ่มข้อมูลในทางปฏิบัติ โดยผู้วิจัยจึงเสนอว่าควรใช้ทั้งสองวิธีร่วมกันเพื่อประเมินความสอดคล้องของการจัดกลุ่มข้อมูลอย่างเต็มประสิทธิภาพ

บทที่ 3
วิธีการดำเนินการวิจัย

การวิจัยนี้ผู้วิจัยมุ่งเน้นการจัดกลุ่มนักเรียนภายในชั้นเรียนอย่างมีประสิทธิภาพ และสามารถช่วยอาจารย์ผู้สอนในการประกอบการตัดสินใจสำหรับการเรียนการสอนภายในชั้นเรียน โดยมีขั้นตอนกระบวนการ วิธีวัดประสิทธิภาพของการแบ่งกลุ่มนักเรียนภายในชั้นเรียน รวมถึงระบบสนับสนุนในการช่วยตัดสินใจของอาจารย์ผู้สอน โดยมีขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM (Cross-Industry Standard Process for Data Mining) กับการแบ่งกลุ่มนักเรียนตามรูปแบบการเรียนรู้มาประยุกต์ใช้กับงานวิจัย ซึ่งประกอบด้วยกระบวนการทั้งหมด 6 ขั้นตอน กระบวนการแต่ละขั้นตอนนั้นจะต่อเนื่องกันโดยขั้นตอนในลำดับถัดไปจะต้องรอให้กระบวนการก่อนหน้าเสร็จสิ้นก่อน และเมื่อได้ผลลัพธ์จากกระบวนการสร้างโมเดลจำลองข้อมูล (Modeling) อาจจะมีกระบวนการย้อนกลับของการปรับปรุงขั้นตอน และกระบวนการเพื่อให้ได้ผลลัพธ์ที่ดีมีประสิทธิภาพ ของการวิเคราะห์ข้อมูลในงานวิจัย

3.1 การทำความเข้าใจกับปัญหา (Business Understanding)

ขั้นตอนแรกในกระบวนการ CRISP-DM ซึ่งเน้นไปที่การทำความเข้าใจกับปัญหา และแปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลของดเหมืองข้อมูล การวางแผนในการดำเนินการ ซึ่งในงานวิจัยนี้มีเป้าหมายในการแก้ปัญหาการแบ่งกลุ่มในชั้นเรียน โดยใช้กระบวนการเหมืองข้อมูลกับชุดข้อมูลรูปแบบการเรียนรู้ เพื่อช่วยแนะนำการแบ่งกลุ่มในชั้นเรียนที่เหมาะสม เข้าใจรูปแบบการเรียนรู้ของนักเรียน ลดการวางแผนการสอนให้กับอาจารย์ผู้สอน และการแนะนำให้นักเรียนได้ทราบถึงรูปแบบการเรียนรู้ของตนเอง โดยบทความงานวิจัยนี้ได้เสนอเทคนิคการแบ่งกลุ่มของนักเรียนในชั้นเรียน การหาวิธีการการแบ่งกลุ่มหลากหลายแนวทาง และมีประสิทธิภาพโดยอาศัยหลากหลายเทคนิคในกระบวนการการทำเหมืองข้อมูล รวมไปถึงการประเมินรูปแบบการเรียนรู้ของนักเรียน และนักศึกษา เพื่อหาจำนวนกลุ่มที่แนะนำของชั้นเรียน โดยมีการวัดประสิทธิภาพของการแบ่งกลุ่มโดยกระบวนการหาอัตราการเปลี่ยนแปลงของค่า Sum Square of Error (SSE) เพื่อให้ได้กลุ่มที่เหมาะสมของชุดข้อมูล และการแนะนำให้ผู้สอนได้ทราบถึงหมายเลขสมาชิกของนักเรียนในแต่ละกลุ่ม ในกรณีที่ทรัพยากรในการสอนมีจำกัด หรือผู้สอนต้องการที่จะแบ่งจำนวนกลุ่มแบบเจาะจง โดยกระบวนการประเมินรูปแบบการเรียนรู้ของนักเรียน กระบวนการหาวิธีการทางเหมืองข้อมูลที่เหมาะสม และกระบวนการประเมินประสิทธิภาพของการแบ่งกลุ่ม โดยนำข้อมูลที่ได้จากกระบวนการนั้นนำมาประยุกต์ใช้ให้เหมาะสมกับความต้องการของอาจารย์ผู้สอน

3.2 การเตรียมข้อมูล (Data Understanding)

ขั้นตอนที่สองในกระบวนการ CRISP-DM ซึ่งเน้นไปที่การเก็บรวบรวมข้อมูล และการตรวจสอบความถูกต้องของข้อมูลที่รวบรวม เพื่อประกอบการพิจารณาขอบเขตในการใช้ข้อมูล โดยเลือกชุดข้อมูลของนักศึกษามหาวิทยาลัย จัดเก็บข้อมูลชั้นเรียนนักศึกษาระดับปริญญาตรี ชั้นปีที่ 1 - 4 สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์มหาวิทยาลัยแม่โจ้ อำเภอสันทราย จังหวัดเชียงใหม่ ปีการศึกษา 2563 จำนวน 162 คน และสำหรับการทดลองชุดข้อมูลนักเรียน จัดเก็บข้อมูลชั้นเรียนระดับมัธยมศึกษาปีที่ 3 โรงเรียนชลประทานผาแตก อำเภอดอยสะเก็ด จังหวัดเชียงใหม่ ปีการศึกษา 2564 จำนวน 35 คน โดยแบ่งข้อมูลออกเป็น 2 ส่วนคือ

1. ข้อมูลพื้นฐานของนักเรียน นักศึกษา ได้แก่ คำนำหน้าชื่อ ชื่อ นามสกุล เพศ ระดับชั้นการศึกษา ระดับชั้นเรียน ห้องเรียน โดยมีโครงสร้างการจัดเก็บข้อมูลพื้นฐานของนักเรียน นักศึกษา ดังตารางที่ 2 และตัวอย่างข้อมูลดิบ (Raw Data) ข้อมูลพื้นฐานของนักเรียน และนักศึกษา ดังตารางที่ 3

2. ข้อมูลการประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) จากแบบประเมินเกี่ยวกับรูปแบบการเรียนรู้ทั้ง 8 ด้านตามรูปแบบการเรียนรู้ Index Learning Style (ILS) ของ (Felder และSilverman, 1988) และมีโครงสร้างการจัดเก็บข้อมูลรูปแบบการเรียนรู้ ดังตารางที่ 3 โดยเริ่มจากการเก็บรวบรวมข้อมูลหลังจากการทำแบบประเมินรูปแบบการเรียนรู้ซึ่งออกแบบตามรูปแบบการเรียน Index Learning Style (ILS) ของ (Felder และSilverman, 1988) ซึ่งมีตัวอย่างข้อมูลดิบ (Raw Data) รูปแบบการเรียนรู้ Index Learning Style (ILS) ดังตารางที่ 4 และตัวอย่างรูปแบบการเรียนรู้ ดังตารางที่ 5

โดยมีการอ้างอิงมาตรฐานในการกำหนดรหัสเช่น รหัสเพศ รหัสระดับการศึกษา จากประกาศของกระทรวงศึกษาธิการ เรื่องมาตรฐานการเก็บข้อมูลทางการศึกษา ซึ่งมีตัวอย่างของการจัดเก็บข้อมูลรหัสระดับการศึกษา (Education Level) ตามรหัสข้อมูลมาตรฐานกลาง (Code List) ดังตารางที่ 5 และตัวอย่างของการจัดเก็บข้อมูลรหัสชั้นปี (Grade Level) ตามรหัสข้อมูลมาตรฐานกลาง (Code List) ถัดมางานวิจัยนี้ยังเก็บข้อมูล ถัดมาจะเป็นกระบวนการตรวจสอบข้อมูลที่ได้รวบรวมเพื่อตรวจสอบความถูกต้องของข้อมูล และพิจารณาถึงขอบเขตของข้อมูลที่จะต้องใช้ในการวิเคราะห์

ตารางที่ 2 โครงสร้างการจัดเก็บข้อมูลพื้นฐานของนักเรียน นักศึกษา

ชื่อหัวข้อ	**คำอธิบาย**
Prefix	คำนำหน้าชื่อ

FirstName	ชื่อ
LastName	นามสกุล
Gender	รหัสเพศ (ชาย = 1, หญิง = 2)
EducationLevel	รหัสระดับชั้นการศึกษา (มัธยมศึกษาตอนต้น = 12, มัธยมศึกษาตอนปลาย = 13, ปริญญาตรี = 16)
GradeLevel	ระดับชั้นเรียน
ClassRoom	ห้องเรียน

ตารางที่ 3 ตัวอย่างข้อมูลดิบ (Raw Data) ข้อมูลพื้นฐานของนักเรียน และนักศึกษา

Prefix	FirstName	LastName	Gender	EducationLevel	GradeLevel	ClassRoom
เด็กหญิง	มาลี	รักเรียน	2	12	มัธยมศึกษาปีที่3	1
เด็กหญิง	เขมิกา	พุทธสิงห์	2	12	มัธยมศึกษาปีที่3	1
เด็กชาย	เสมา	ธรรมชาติ	1	12	มัธยมศึกษาปีที่3	1
เด็กชาย	สะอาด	มารยาทดี	1	12	มัธยมศึกษาปีที่3	1

ตารางที่ 4 โครงสร้างการจัดเก็บข้อมูลรูปแบบการเรียนรู้ Index Learning Style (ILS)

ชื่อหัวข้อ	**คำอธิบาย**
StudentID	รหัสนักเรียน
ActivePoint	คะแนนรูปแบบการเรียนรู้ด้าน Active
ReflectivePoint	คะแนนรูปแบบการเรียนรู้ด้าน Reflective
SensingPoint	คะแนนรูปแบบการเรียนรู้ด้าน Sensing
IntuitivePoint	คะแนนรูปแบบการเรียนรู้ด้าน Intuitive
VisualPoint	คะแนนรูปแบบการเรียนรู้ด้าน Visual
VerbalPoint	คะแนนรูปแบบการเรียนรู้ด้าน Verbal

SequentialPoint	คะแนนรูปแบบการเรียนรู้ด้าน Sequential
GlobalPoint	คะแนนรูปแบบการเรียนรู้ด้าน Global

ตารางที่ 5 ตัวอย่างข้อมูลดิบ (Raw Data) รูปแบบการเรียนรู้ Index Learning Style (ILS)

StudentID	Active Point	Reflective Point	Sensing Point	Intuitive Point	Visual Point	Verbal Point	Sequential Point	Global Point
1	0	1	3	0	0	1	0	1
2	3	0	3	0	5	0	3	0
3	0	3	1	0	0	5	0	5
4	0	1	0	1	5	0	5	0

ตารางที่ 6 ตัวอย่างข้อมูลรหัสระดับการศึกษา (Education Level) ตามรหัสข้อมูลมาตรฐานกลาง (Code List)

ที่	รหัส	ระดับการศึกษา (CLS_EDUCATION_LEVEL)	หมายเหตุ
1	00	เตรียมอนุบาล	
2	10	ก่อนประถมศึกษา	
3	11	ประถมศึกษา	
4	12	มัธยมศึกษาตอนต้น	
5	13	มัธยมศึกษาตอนปลาย	
6	14	ประกาศนียบัตรวิชาชีพเทคนิค (ปวท.)	
7	15	อนุปริญญา	
8	16	ปริญญาตรี	

ตารางที่ 7 ตัวอย่างข้อมูลรหัสชั้นปี (Grade Level) ตามรหัสข้อมูลมาตรฐานกลาง (Code List)

ที่	รหัส	ระดับชั้น (GRADE_LEVEL)	หมายเหตุ
1	211	ประถมศึกษาปีที่ 1/เกรด 1	
2	212	ประถมศึกษาปีที่ 2/เกรด 2	

3	213	ประถมศึกษาปีที่ 3/เกรด 3	
4	214	ประถมศึกษาปีที่ 4/เกรด 4	
5	215	ประถมศึกษาปีที่ 5/เกรด 5	
6	216	ประถมศึกษาปีที่ 6/เกรด 6	
7	311	มัธยมศึกษาปีที่ 1 /เกรด 7/ นาฏศิลป์ชั้นที่ 1	
8	312	มัธยมศึกษาปีที่ 2 /เกรด 8/ นาฏศิลป์ชั้นที่ 2	
9	313	มัธยมศึกษาปีที่ 3 /เกรด 9/ นาฏศิลป์ชั้นที่ 3	
10	411	มัธยมศึกษาปีที่ 4/เกรด10	
11	412	มัธยมศึกษาปีที่ 5/เกรด11/เตรียมทหารชั้นปีที่ 1	
12	413	มัธยมศึกษาปีที่ 6/เกรด12/เตรียมทหารชั้นปีที่ 2/กศน.มัธยมศึกษา	
13	601	ปริญญาตรีชั้นปีที่ 1 (สามัญ)	
14	602	ปริญญาตรีชั้นปีที่ 2 (สามัญ)	
15	603	ปริญญาตรีชั้นปีที่ 3 (สามัญ)	
16	604	ปริญญาตรีชั้นปีที่ 4 (สามัญ)	
17	605	ปริญญาตรีชั้นปีที่ 5 (สามัญ)	
18	606	ปริญญาตรีชั้นปีที่ 6 (สามัญ)	
19	607	ปริญญาตรีชั้นปีที่ 7 (สามัญ)	
20	608	ปริญญาตรีชั้นปีที่ 8 (สามัญ)	

3.3 การแก้ไข และดัดแปลงข้อมูล (Data Preparation)

กระบวนการเตรียมข้อมูลก่อนนำไปวิเคราะห์ โดยปกติแล้วข้อมูลที่มาจากแหล่งต่าง ๆ มักจะมีความไม่สมบูรณ์ของข้อมูล และมีการจัดเก็บขาดหรือเกินขอบเขตที่จะนำมาใช้วิเคราะห์จริง การเตรียมข้อมูลจึงเป็นการเตเตรียมพร้อมสำหรับการวิเคราะห์ และตรวจสอบความถูกต้องของข้อมูล เพื่อให้พร้อมกับขั้นตอนของการวิเคราะห์ และนำไปใช้ในการสร้างโมเดลประกอบการตัดสินใจอย่างมีประสิทธิภาพ

เนื่องจากข้อมูลหลังจากการตอบแบบประเมินนั้นมีข้อมูลพื้นฐาน ซึ่งเป็นส่วนบุคคลของนักเรียน นักศึกษา และข้อมูลรูปแบบการเรียนรู้ Index Learning Style (ILS) อยู่ภายในแถว และคอลัมน์เดียวกันจึงเกิดการคัดเลือกข้อมูลเฉพาะส่วน เพื่อให้สามารถนำไปวิเคราะห์โดยใช้เทคนิคเหมืองข้อมูล โดยแยกตารางสำหรับการวิเคราะห์ข้อมูลออกจากตารางหลัก ซึ่งภายในตารางนั้น

ประกอบด้วย StudentID เพื่อระบุตัวตนของนักเรียน และคะแนนรูปแบบกาเรียนรู้ Index Learning Style (ILS) เท่านั้นในการนำข้อมูลไปวิเคราะห์ในกระบวนการต่อไป ดังแสดงในภาพที่ 12

ภาพที่ 12 การแปลงข้อมูลหลังการประเมิน ให้อยู่ในรูปแบบสำหรับการวิเคราะห์

3.4 กระบวนการสร้าง และพัฒนาโมเดล (Modeling)

กระบวนการสร้างโมเดลเป็นการวิเคราะห์ข้อมูล และสร้างสมมติฐานว่ามีความสัมพันธ์กันระหว่างตัวแปรในข้อมูลนั้น ๆ โดยการเลือกวิธีการที่เหมาะสมกับข้อมูล โดยการนำข้อมูลมาสร้างโมเดลเพื่อใช้ในการทำนายผลหรือการจัดกลุ่มข้อมูล มีการแบ่งกลุ่มตามรูปแบบการเรียนรู้อยู่ด้วยกัน โดยงานวิจัยนี้เลือกใช้การแบ่งกลุ่มแบบUnsupervised Learning (การเรียนรู้แบบไม่มีผู้สอน) ซึ่งในการแบ่งกลุ่มแบบ Unsupervised Learning จะไม่มีการใช้ข้อมูลที่มีคำตอบในการสร้างโมเดล โดยจะมีขั้นตอนหลักๆ ดังนี้

1. การเตรียมข้อมูล: การเตรียมข้อมูลโดยการแยกข้อมูลออกเป็น Training Data เพื่อใช้ในการสร้างโมเดล ขั้นตอนการเตรียมแบบประเมินเพื่อแบบประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) โดย Felder-Silverman Learning จากแบบประเมินต้นฉบับ แปลไทย โดยมีผู้ร่วมตอบแบบประเมินเพื่อเก็บข้อมูล ดังนี้

- ชุดข้อมูล นักศึกษาชั้นปีที่ 1 จำนวน 60 คน
- ชุดข้อมูล นักศึกษาชั้นปีที่ 2 จำนวน 24 คน
- ชุดข้อมูล นักศึกษาชั้นปีที่ 3 จำนวน 48 คน
- ชุดข้อมูล นักศึกษาชั้นปีที่ 4 จำนวน 29 คน
- ชุดข้อมูล นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 35 คน

โดยแบบประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) โดย Richard M. Felder และ Barbara A. Soloman แบบประเมินป้อนคำตอบแบบ 2 ตัวเลือก โดยเลือกเพียงคำตอบเดียว สำหรับ "a" และ "b" ที่ตรงกับผู้ประเมิน ให้เลือกอันที่ใช้บ่อยกว่าแต่ละคำถาม ดังรูปที่ 13 ตัวอย่างคำถามแบบประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) สำหรับแบบประเมินของงานวิจัยได้มีกระบวนการแปลภาษาจากภาษาอังกฤษ เป็นภาษาไทยเพื่อให้นักเรียน นักศึกษาสามารถตอบแบบประเมินได้อย่างถูกต้อง

INDEX OF LEARNING STYLES*
Richard M. Felder
Barbara A. Soloman

DIRECTIONS

Enter your answers to every question on the ILS scoring sheet. Please choose only one answer for each question. If both "a" and "b" seem to apply to you, choose the one that applies more frequently.

1. I understand something better after I
   a) try it out.
   b) think it through.
2. I would rather be considered
   a) realistic.
   b) innovative.
3. When I think about what I did yesterday, I am most likely to get
   a) a picture.
   b) words.
4. I tend to
   a) understand details of a subject but may be fuzzy about its overall structure.
   b) understand the overall structure but may be fuzzy about details.

ภาพที่ 13 ตัวอย่างคำถามแบบประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS)

สำหรับการคิดคำนวณคะแนนของแบบประเมิน ตามรูปแบบการเรียนรู้แต่ละด้านในช่องกรอกคะแนนประเมิน ดังภาพที่ 14

1. การให้คะแนนแบบทดสอบ โดยการใส่ “1” ในช่องว่างในตารางตามรูปภาพที่ 1 (เช่น หากตอบ “a” ในคำถามที่ 3 ให้ใส่ “1” ในคอมลัมน์ “a” ในช่องคำถาม 3)
2. รวมคอลัมน์ และเขียนผลรวมในช่องว่างที่ระบุ

3. สำหรับแต่ละสเกลทั้งสี่ ให้ลบผลรวมที่น้อยกว่าออกจากสเกลที่ใหญ่กว่า เขียนความแตกต่าง (1 ถึง 11) และตัวอักษร (a หรือ b) ด้วยผลรวมที่มากกว่า

ตัวอย่างเช่น ถ้าภายใต้ "ACT/REF" มีคำตอบ 4 "a" และ 7 "b" จะเขียน "3b" บนบรรทัดล่างสุดภายใต้หัวข้อนั้น (3 = 7- 4 และผลรวม "b" คือค่าที่มากกว่าของทั้งสอง)

ACT/REF			SEN/INT			VIS/VRB			SEQ/GLO		
Q	a	b	Q	a	b	Q	a	b	Q	a	b
1	__	__	2	__	__	3	__	__	4	__	__
5	__	__	6	__	__	7	__	__	8	__	__
9	__	__	10	__	__	11	__	__	12	__	__
13	__	__	14	__	__	15	__	__	16	__	__
17	__	__	18	__	__	19	__	__	20	__	__
21	__	__	22	__	__	23	__	__	24	__	__
25	__	__	26	__	__	27	__	__	28	__	__
29	__	__	30	__	__	31	__	__	32	__	__
33	__	__	34	__	__	35	__	__	36	__	__
37	__	__	38	__	__	39	__	__	40	__	__
41	__	__	42	__	__	43	__	__	44	__	__
Total (sum X's in each column)											
ACT/REF			**SEN/INT**			**VIS/VRB**			**SEQ/GLO**		
	a	b		a	b		a	b		a	b
	__	__		__	__		__	__		__	__
(Larger - Smaller) + Letter of Larger (see below*)											
	____		____			____			____		

ภาพที่ 14 ตารางแบบประเมินรูปแบบการเรียนรู้ ILS

คำอธิบายของคะแนนจากการประเมินรูปแบบการเรียนรู้เบื้องต้น มีดังนี้

- หากคะแนนของคุณในระดับ 1 - 3 แสดงว่าคุณมีความชอบเล็กน้อยสำหรับมิติใดมิติหนึ่งหรืออีกมิติหนึ่ง แต่โดยพื้นฐานแล้วคุณมีความสมดุลที่ดี (ตัวอย่างเช่น 3a ในหมวด ACT/REF บ่งชี้ถึงความต้องการที่ไม่รุนแรงสำหรับการเรียนรู้แบบกระตือรือร้น (Active learning)
- หากคะแนนของคุณในมาตราส่วนคือ 5 - 7 แสดงว่าคุณมีความชอบในระดับปานกลางสำหรับหนึ่งมิติของมาตราส่วน และจะเรียนรู้ได้ง่ายขึ้นในสภาพแวดล้อมการสอนที่สนับสนุนมิตินั้น

- หากคะแนนของคุณในระดับ 9-11 แสดงว่าคุณมีความชอบด้านใดด้านหนึ่งของมาตราส่วนมาก คุณอาจมีปัญหาในการเรียนรู้อย่างแท้จริงในสภาพแวดล้อมที่ไม่สนับสนุนความชอบนั้น

ซึ่งรูปแบบการรายงานผลลัพธ์หลังจากทำแบบประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) จะเป็น ดังภาพที่ 15 และคำอธิบายถึงคะแนนความชอบในแต่ละด้านการเรียนรู้ คำแนะนำ และกระบวนการที่เหมาะสมกับการเรียนรู้ในด้านละด้านอ้างอิงจาก (Felder และ Soloman, 2000) คำอธิบายรู้แบบการเรียนรู้ Index Learning Style (ILS)

INDEX OF LEARNING STYLES -- REPORT OF RESULTS

ACT												**REF**
	11a	9a	7a	5a	3a	1a	1b	3b	5b	7b	9b	11b
SEN												**INT**
	11a	9a	7a	5a	3a	1a	1b	3b	5b	7b	9b	11b
VIS												**VRB**
	11a	9a	7a	5a	3a	1a	1b	3b	5b	7b	9b	11b
SEQ												**GLO**
	11a	9a	7a	5a	3a	1a	1b	3b	5b	7b	9b	11b

ภาพที่ 15 ตารางสรุปผลลัพธ์ รูปแบบการเรียนรู้ Index Learning Style

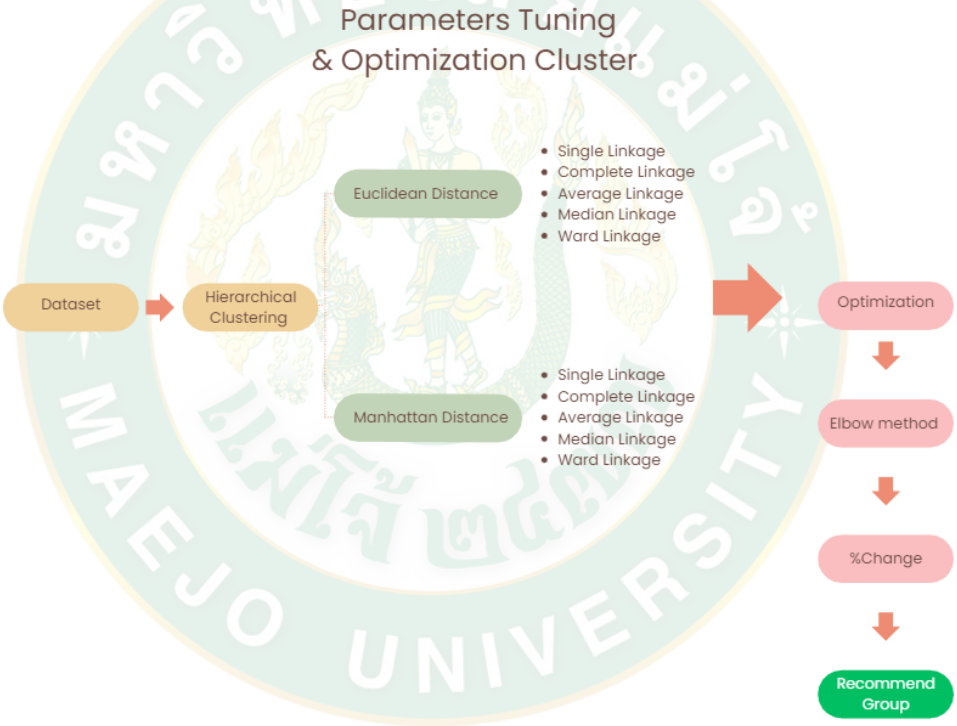
3.6 แนวคิดวิธีการทดสอบโมเดล

แนวคิดการสอบทดสอบโมเดลการแบ่งกลุ่มด้วย Parameters Tuning เป็นกระบวนการปรับแต่งพารามิเตอร์ของโมเดลเพื่อให้มีประสิทธิภาพในการทำงานมากขึ้น ในการแบ่งกลุ่มตามรูปแบบการเรียนรู้ การที่จะปรับแต่งพารามิเตอร์เหล่านี้จะช่วยปรับปรุงโมเดลให้สามารถทำงานได้อย่างมีประสิทธิภาพมากขึ้นเพื่อให้ได้ผลการแบ่งกลุ่มที่ดีกว่าด้วยความแม่นยำสูง โดยในการแบ่งกลุ่มตามรูปแบบการเรียนรู้ Hierarchical Clustering จะมีพารามิเตอร์สำคัญ 2 ตัว ได้แก่ distance และ linkage

1. Distance Measure: เป็นวิธีการคำนวณระยะห่างระหว่างจุดของข้อมูล ที่ใช้ในการคำนวณหาความคล้ายคลึงระหว่างกลุ่มข้อมูล โดยมีวิธีการคำนวณระยะห่างในการแบ่งกลุ่ม Hierarchical ได้แก่ Euclidean Distance, Manhattan Distance

2. Linkage Method: เป็นวิธีการคำนวณความคล้ายคลึงระหว่างกลุ่มข้อมูลโดยใช้ค่าที่เป็นไปได้ของระยะห่างระหว่างข้อมูลจาก distance โดย Linkage สามารถแบ่งได้เป็น 5 ประเภท ได้แก่ Single Linkage, Complete Linkage, Average Linkage, Median Linkage และ Ward Linkage

ซึ่งในกรณีที่ต้องการปรับแต่งพารามิเตอร์สำหรับ Hierarchical Clustering จะทำได้โดยการทดสอบรูปแบบการเรียนรู้แต่ละแบบด้วยการใช้ค่า Distance และ Linkage ที่แตกต่างกัน และเลือกค่าที่ให้ผลการแบ่งกลุ่มที่ดีที่สุด (Optimization) โดยใช้ Elbow method และ %Change ดังแนวคิดการวัดประสิทธิภาพข้อมูล ดังภาพที่ 16



ภาพที่ 16 กระบวนการปรับแต่งพารามิเตอร์ของโมเดลการแบ่งกลุ่ม

3.7 แนวคิดวิธีการวัดประสิทธิภาพข้อมูล

จากการแบ่งกลุ่มของชุดข้อมูลโดยใช้พารามิเตอร์ในรูปแบบที่แตกต่างกันกับชุดข้อมูล นำมาซึ่งผลลัพธ์จากการแบ่งกลุ่มของชุดนั้น ๆ ถัดมาจะเป็นกระบวนการวัดประสิทธิภาพของการแบ่งกลุ่มชุดข้อมูลที่เหมาะสม (Optimization) ซึ่งผู้วิจัยได้มีแนวคิดในการวัดประสิทธิภาพของการแบ่งกลุ่มชุดข้อมูล ดังนี้

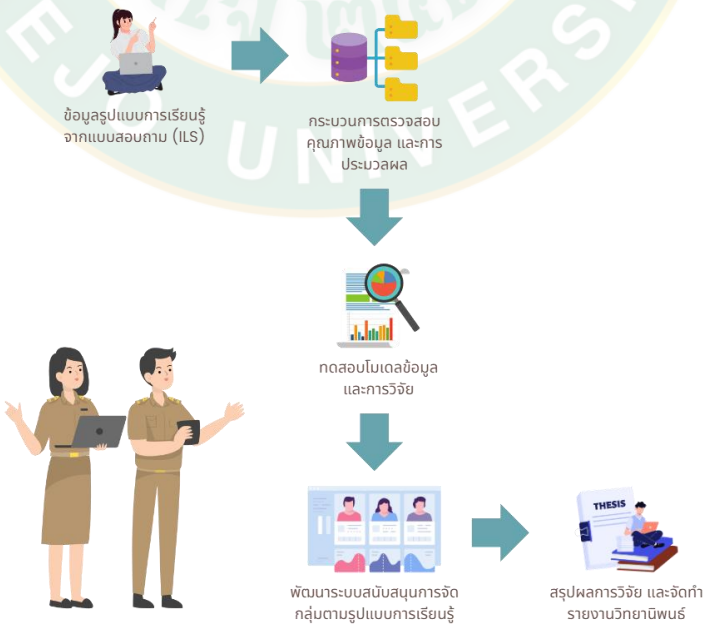
1. Elbow method เป็นวิธีการใช้กราฟเส้นโค้งที่บอกถึงความแตกต่างระหว่างจำนวนกลุ่มและความสอดคล้องของข้อมูลภายในกลุ่ม โดยจะวาดกราฟเส้นต่อเนื่องระหว่างค่าของ SSE กับจำนวนกลุ่ม และหาจุดที่กราฟเหลี่ยมคล้ายข้อศอก (elbow point) ซึ่งจะเป็นจำนวนกลุ่มที่เหมาะสมสำหรับแบ่งกลุ่มข้อมูล

2. ค่าเปลี่ยนแปลงของ SSE (%Change) จะใช้วัดประสิทธิภาพของการแบ่งกลุ่มที่ได้จาก Elbow method โดยการวัดการเปลี่ยนแปลงของค่าเฉลี่ยของคะแนนในแต่ละกลุ่ม หากการแบ่งกลุ่มมีประสิทธิภาพดี %Change จะต้องมีค่าเฉลี่ยของคะแนนในแต่ละกลุ่มที่แตกต่างกันมาก ๆ ซึ่งในงานวิจัยได้มีการกำหนดค่าของ %Change ที่ยอมรับได้ในการวัดความเหมาะสมของกลุ่ม

ดังนั้น การใช้ Elbow method ร่วมกับ %Change สามารถช่วยในการประเมินประสิทธิภาพของการแบ่งกลุ่มนักเรียนตามรูปแบบการเรียนรู้ได้ โดยการหาจำนวนกลุ่มที่เหมาะสมจากค่า Elbow และตรวจสอบค่า %Change ของคะแนนในแต่ละกลุ่มเพื่อยืนยันความสอดคล้องกับการแบ่งกลุ่มนั้น หากการแบ่งกลุ่มด้วยจำนวนกลุ่มนี้มีประสิทธิภาพสูง ก็สามารถนำไปใช้ในการสร้างแบบจำลองเรียนรู้แบบต่าง ๆ ที่เหมาะสมกับลักษณะของรูปแบบการเรียนรู้

3.8 สรุปผลการวิจัย และจัดทำรายงานวิทยานิพนธ์

ภาพรวมสรุปกระบวนการวิจัยเพื่อการตอบสนองวัตถุประสงค์ และสมมติฐานที่กำหนดไว้ในแผนงานวิจัย นำมาซึ่งผลการวิจัย และพัฒนาปรับปรุงองค์ความรู้ใหม่ ๆ ทางด้านการศึกษาผ่านการจัดทำรายงานวิทยานิพนธ์ ดังรูปภาพที่ 17



ภาพที่ 17 แผนผังสรุปกระบวนการวิจัย

บทที่ 4
ผลการดำเนินการวิจัย

งานวิจัยได้ดำเนินงานตามวัตถุประสงค์ ซึ่งเกี่ยวกับการวิเคราะห์ข้อมูลโดยการโมเดลวิธีการทางการทำเหมืองข้อมูล (Data Mining) ด้วยการแบ่งกลุ่มโดยใช้วิธี Hierarchical Clustering เพื่อทดลองหาพารามิเตอร์ที่เหมาะสมสำหรับการแบ่งกลุ่ม และหาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ในชุดข้อมูลสร้างระบบที่สนับสนุนการแบ่งกลุ่มในชั้นเรียนโดยรูปแบบการเรียนรู้ โดยผู้วิจัยได้ทำการเก็บข้อมูลจากนักเรียน นักศึกษา จากแบบทดสอบรูปแบบการเรียนรู้ Index Learning Style (ILS) เพื่อให้ทราบถึงพารามิเตอร์ และการกำหนดค่าที่เหมาะสมกับชุดข้อมูลด้านการศึกษา โดยผู้วิจัยได้ดำเนินการทดลองงานวิจัยตามกระบวนการ CRISP-DM ในกระบวนการดังต่อไปนี้

4.1 การวิเคราะห์ปัญหา และการเลือกปัญหาจากข้อมูลที่มี

งานวิจัยนี้กล่าวถึงการแบ่งกลุ่มในชั้นเรียน และรูปแบบการเรียนรู้ของนักเรียนนักศึกษา และการแก้ไขปัญหาในการแบ่งกลุ่มภายในชั้นเรียนให้สอดคล้องกับรูปแบบการเรียนรู้ ทักษะ จำนวนกลุ่มที่เหมาะสมภายในชั้นเรียนนั้น ๆ โดยมีแนวคิดในการที่อาจารย์ผู้สอนนั้นสามารถรับรู้ถึงคุณลักษณะของนักเรียนแต่ละคน การแบ่งจำนวนกลุ่มที่เหมาะสม รวมถึงลักษณะรูปแบบการเรียนรู้ของนักเรียนที่คล้ายคลึงกันภายในกลุ่ม นอกเหนือจากการที่โมเดลของงานวิจัยนี้จะแนะนำกลุ่มที่เหมาะสมกับชั้นเรียนให้กับอาจารย์ งานวิจัยได้คำนึงถึงสภาพแวดล้อมห้องเรียนจริงซึ่งอาจารย์ผู้สอนนั้นสามารถปรับพารามิเตอร์จำนวนกลุ่มให้เหมาะสมกับอุปกรณ์ในการเรียน การสอน หรือทรัพยากรที่มีแบบจำกัด การแนะนำถึงลักษณะของนักเรียน รวมถึงกลุ่มของนักเรียนเพื่อการวางแผน และประยุกต์การเรียนการสอนให้เหมาะสมกับผู้เรียนให้ดียิ่งขึ้น

4.2 การเตรียมข้อมูล

จากการวิเคราะห์ปัญหาในเรื่องของการแบ่งกลุ่มในชั้นเรียน และการกำหนดวิธีการแก้ไขปัญหา กระบวนการถัดมาคือการเตรียมข้อมูลหรือตัวแปรต่าง ๆ ซึ่งเกี่ยวข้องกับการนำมาวิเคราะห์ข้อมูล โดยข้อมูลหลักที่งานวิจัยนี้เลือกใช้คือข้อมูลรูปแบบการเรียนรู้ Index Learning Style (ILS) ซึ่งหลักจาจกนักเรียนประเมินรูปแบบการเรียนรู้เสร็จสิ้นจะได้ข้อมูลดิบ (Raw Data) ซึ่งจะต้องเข้าสู่กระบวนการทางเหมืองข้อมูลเพื่อเตรียมข้อมูลให้พร้อมสำหรับกระบวนวิเคราะห์ทางเหมืองข้อมูล ซึ่งตัวอย่างของข้อมูลหลังผ่านกระบวนการเป็น ดังภาพที่ 18

	C1	C2	C3	C4	C5	C6	C7	C8	C9
	StudentID	ACT	REF	SEN	INT	VIS	VRB	SEQ	GLO
1	1	0	1	3	0	0	1	0	1
2	2	3	0	3	0	5	0	3	0
3	3	0	3	1	0	0	5	0	5
4	4	0	1	0	1	5	0	5	0
5	5	0	1	0	3	3	0	0	3
6	6	5	0	9	0	5	0	0	1
7	7	1	0	3	0	0	5	0	5
8	8	0	3	3	0	9	0	1	0
9	9	0	1	0	7	3	0	0	9
10	10	0	1	0	1	0	1	0	1

ภาพที่ 18 ชุดข้อมูลของนักเรียนก่อนนำไปวิเคราะห์

4.3 การเลือกเทคนิคที่เหมาะสม

การแบ่งกลุ่มข้อมูลเป็นหนึ่งในเทคนิคการวิเคราะห์ข้อมูลที่สำคัญ เพื่อให้เกิดความเข้าใจ และการนำข้อมูลไปใช้อย่างมีประสิทธิภาพ งานวิจัยนี้ได้พิจารณาทดลองเทคนิคการแบ่งกลุ่มข้อมูลหลายพารามิเตอร์ได้แก่

1.พารามิเตอร์ Distance จำนวน 2 ตัวแปรได้แก่ Euclidean Distance และ Manhattan Distance

2.พารามิเตอร์ Linkage จำนวน 5 ตัวแปรได้แก่ Single Linkage, Complete Linkage, Average Linkage, Median Linkage และWard Linkage

3. พารามิเตอร์ด้านชุดข้อมูล จำนวน 5 ชุดข้อมูลได้แก่ ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1, ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2, ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3, ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 4 และชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3

จากการทดลองการแบ่งกลุ่มกับหลายพารามิเตอร์ และชุดข้อมูลในงานวิจัยนี้เพื่อเปรียบเทียบหาเทคนิคที่เหมาะสมกับการแบ่งกลุ่มในชั้นเรียน ซึ่งจะช่วยแก้ปัญหาลดระยะเวลาในการแบ่งกลุ่ม และมีวิธีการแบ่งกลุ่มอย่างมีประสิทธิภาพ

4.4 ผลการทดสอบความแม่นยำของตัวต้นแบบ (Evaluation)

การทำทดสอบความแม่นยำของตัวต้นแบบกับพารามิเตอร์ที่เหมาะสม จากการทดลองพบตัวอย่างของการแบ่งชุดข้อมูลทั้ง 5 ชุดข้อมูลกับพารามิเตอร์ที่ค่อนข้างเหมาะสมกับชุดข้อมูลรูปแบบการเรียนรู้ ได้แก่การใช้พารามิเตอร์ Euclidean Distance และ Ward Linkage โดยการจำแนกกลุ่มออกเป็น 7 กลุ่มหรือ K = 7 ดังตัวอย่างการแบ่งกลุ่มด้วยพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลดังต่อไปนี้

4.4.1 ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3

ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 35 คน โดยการแบ่งกลุ่มแบบ Hierarchical Clustering ด้วยการใช้พารามิเตอร์ Euclidean Distance, Ward Linkage โดยเลือก Cluster ที่ 7 เพราะมีความเหมาะสมในการแบ่งกลุ่มของชุดข้อมูลเพราะมี SSE ที่ค่อนข้างน้อย และอัตราการเปลี่ยนแปลงของ SSE หรือ %Change ที่เหมาะสม โดยมีกระบวนการรวมกลุ่มของการแบ่งกลุ่มแบบ Hierarchical Clustering ดังนี้

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	34	91.263	1.4142	10	32	10	2
2	33	87.644	2.0000	5	34	5	2
3	32	87.644	2.0000	19	33	19	2
4	31	87.644	2.0000	20	22	20	2
5	30	84.867	2.4495	21	25	21	2
6	29	84.232	2.5523	19	29	19	3
7	28	82.526	2.8284	27	35	27	2
8	27	82.526	2.8284	2	23	2	2
9	26	82.526	2.8284	12	17	12	2
10	25	82.526	2.8284	6	14	6	2
11	24	82.526	2.8284	4	11	4	2
12	23	81.651	2.9701	1	10	1	3
13	22	81.006	3.0745	18	20	18	3
14	21	79.776	3.2736	26	27	26	3
15	20	79.760	3.2761	16	19	16	4
16	19	76.884	3.7417	3	7	3	2
17	18	70.580	4.7620	15	18	15	4

18	17	70.241	4.8169	6	30	6	3
19	16	69.173	4.9898	8	12	8	3
20	15	65.700	5.5520	5	16	5	6
21	14	63.976	5.8310	13	24	13	2
22	13	63.352	5.9320	5	31	5	7
23	12	62.693	6.0387	2	4	2	4
24	11	57.251	6.9195	13	21	13	4
25	10	54.150	7.4215	2	8	2	7
26	9	43.217	9.1912	1	15	1	7
27	8	42.236	9.3499	2	26	2	10
28	7	35.789	10.3935	1	28	1	8
29	6	35.407	10.4552	6	13	6	7
30	5	22.506	12.5435	3	9	3	3
31	4	7.880	14.9109	1	5	1	15
32	3	-8.586	17.5761	1	3	1	18
33	2	-94.925	31.5514	1	2	1	28
34	1	-136.955	38.3545	1	6	1	35

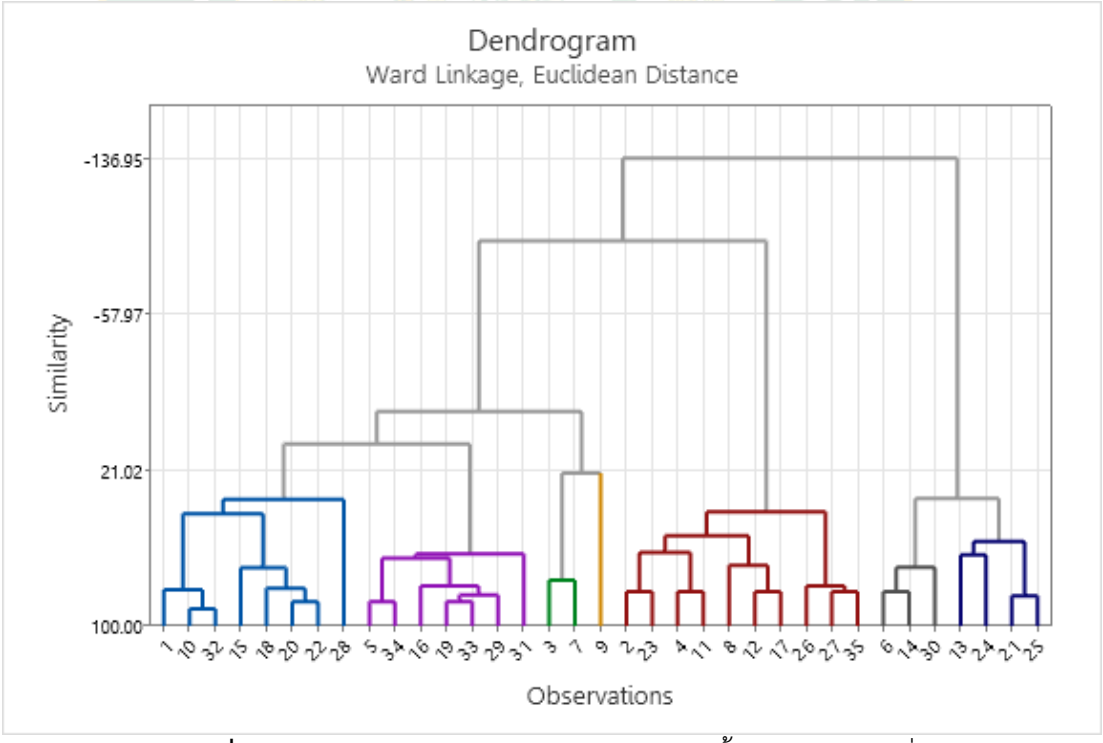
ถัดมาจะเป็นตารางผลรวม หลังจากกระบวนการแบ่งกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ซึ่งจะพิจารณาในส่วนของ Within cluster sum of squares หรือ SSE เพื่อพิจารณาประสิทธิภาพในการแบ่งกลุ่มข้อมูลที่เหมาะสม

Final Partition

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	8	94.500	3.21053	6.25999
Cluster2	10	106.000	3.19569	4.12311
Cluster3	2	7.000	1.87083	1.87083

Cluster4	7	41.143	2.30462	3.68117
Cluster5	3	16.000	2.11765	2.82843
Cluster6	1	0.000	0.00000	0.00000
Cluster7	4	41.000	3.11386	3.96863

แผนภาพ Dendrogram แสดงถึงผลลัพธ์ของการแบ่งกลุ่มซึ่งสามารถบอกถึงความสัมพันธ์แบบลำดับขั้นหลังจากกระบวนการแบ่งกลุ่มแบบ Hirarchical Clustering กับพารามิเตอร์ Ward Linkage และ Euclideance Distance โดยแสดงให้เห็นถึงหมายเลขของสมาชิกในแต่ละกลุ่มข้อมูล (Cluster) โดยจะแทนสีที่แตกต่างกันในแต่ละกลุ่มซึ่งคุณลักษณะของสมาชิกในแต่ละกลุ่ม (Cluster) นั้นมีความคล้ายคลึงกัน และเมื่อเทียบคุณลักษณะของสมาชิกระหว่างกลุ่มนั้นมีความแตกต่างกัน ซึ่งแผนภาพ Dendrogram จะช่วยให้มองเห็นภาพของการแบ่งกลุ่มแบบลำดับขั้นได้ดียิ่งขึ้นในมิติของการแบ่งกลุ่มในจำนวนต่าง ๆ ดังภาพที่ 19



ภาพที่ 19 Dendrogram การจัดกลุ่มของนักเรียนชั้นมัธยมศึกษาปีที่ 3

สำหรับผลลัพธ์ของการแบ่งกลุ่มชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยจะมี StudentID เป็นหมายเลขที่ใช้แทนรายชื่อของเด็กนักเรียน และค่าคะแนนประเมินรูปแบบการเรียนรู้ Index

Learning Style (ILS) ทั้ง 8 ด้าน สุดท้ายที่คอลัม Member จะบ่งบอกถึงหมายเลขสมาชิกกลุ่ม (Cluster) ของนักเรียน ในการแบ่งกลุ่มของชุดข้อมูลนั้น ๆ ดังภาพที่ 20

	StudentID	ACT	REF	SEN	INT	VIS	VRB	SEQ	GLO		Member
1	1	0	1	3	0	0	1	0	1		1
2	2	3	0	3	0	5	0	3	0		2
3	3	0	3	1	0	0	5	0	5		3
4	4	0	1	0	1	5	0	5	0		2
5	5	0	1	0	3	3	0	0	3		4
6	6	5	0	9	0	5	0	0	1		5
7	7	1	0	3	0	0	5	0	5		3
8	8	0	3	3	0	9	0	1	0		2
9	9	0	1	0	7	3	0	0	9		6
10	10	0	1	0	1	0	1	0	1		1

ภาพที่ 20 ผลสมาชิกกลุ่มชุดข้อมูลชั้นมัธยมศึกษาปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม

4.4.2 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1

ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 จำนวน 60 คน โดยการแบ่งกลุ่มแบบ Hierarchical Clustering ด้วยการใช้พารามิเตอร์ Euclidean Distance, Ward Linkage โดยเลือก Cluster ที่ 7 เพราะมีความเหมาะสมในการแบ่งกลุ่มของชุดข้อมูลเพราะมี SSE ที่ค่อนข้างน้อย และอัตราการเปลี่ยนแปลงของ SSE หรือ %Change ที่เหมาะสม โดยมีกระบวนการรวมกลุ่มของการแบ่งกลุ่มแบบ Hierarchical Clustering ดังนี้

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	59	100.000	0.000	29	48	29	2
2	58	95.997	1.000	2	53	2	2
3	57	95.997	1.000	16	52	16	2
4	56	95.997	1.000	28	47	28	2
5	55	95.997	1.000	8	44	8	2
6	54	95.997	1.000	27	40	27	2
7	53	95.997	1.000	23	34	23	2

8	52	95.997	1.000	5	30	5	2
9	51	95.997	1.000	7	11	7	2
10	50	94.339	1.414	43	58	43	2
11	49	94.339	1.414	14	55	14	2
12	48	94.339	1.414	18	46	18	2
13	47	94.339	1.414	3	38	3	2
14	46	94.339	1.414	17	19	17	2
15	45	93.066	1.732	12	15	12	2
16	44	92.938	1.764	27	59	27	3
17	43	92.938	1.764	28	33	28	3
18	42	92.938	1.764	5	26	5	3
19	41	92.466	1.882	27	57	27	4
20	40	91.593	2.100	23	56	23	3
21	39	91.576	2.104	25	43	25	3
22	38	91.297	2.174	3	10	3	3
23	37	90.175	2.454	2	36	2	3
24	36	89.325	2.667	4	29	4	3
25	35	88.830	2.790	8	41	8	3
26	34	88.677	2.828	21	51	21	2
27	33	88.677	2.828	20	50	20	2
28	32	86.820	3.292	14	45	14	3
29	31	85.779	3.552	16	49	16	3
30	30	85.021	3.742	9	39	9	2
31	29	84.276	3.928	4	25	4	6
32	28	82.617	4.342	2	7	2	5
33	27	82.097	4.472	31	37	31	2
34	26	81.429	4.639	18	21	18	4
35	25	79.984	5.000	1	42	1	2
36	24	79.369	5.154	6	14	6	4
37	23	78.936	5.262	4	17	4	8

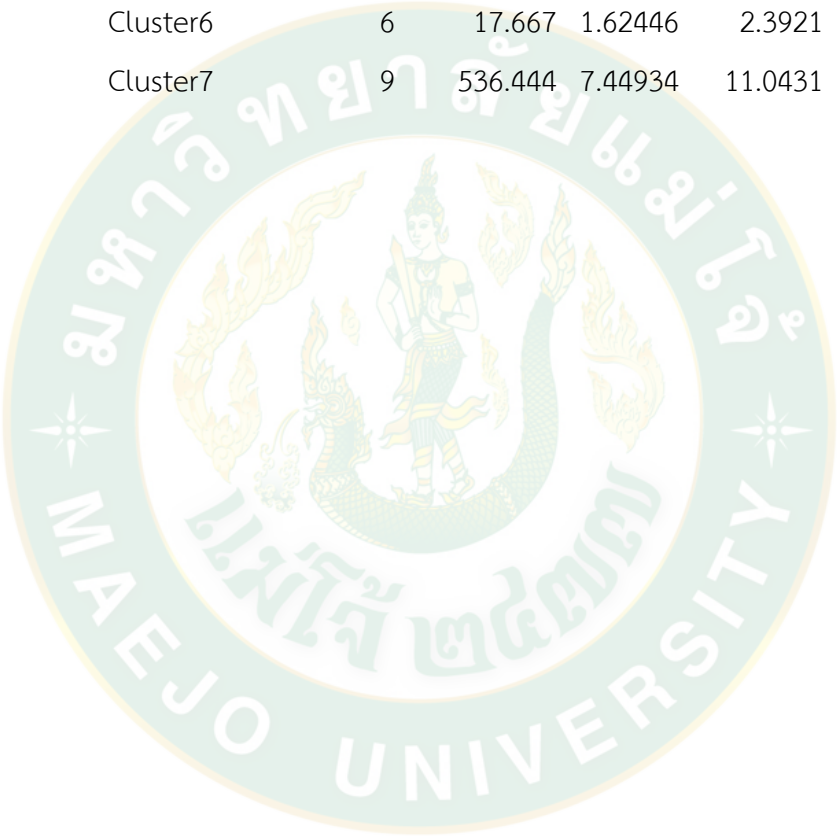
38	22	78.074	5.477	35	54	35	2
39	21	78.074	5.477	24	32	24	2
40	20	77.384	5.650	16	23	16	6
41	19	76.705	5.819	8	28	8	6
42	18	74.353	6.407	3	13	3	4
43	17	72.408	6.892	12	18	12	6
44	16	70.502	7.368	4	27	4	12
45	15	63.340	9.158	1	22	1	3
46	14	48.367	12.898	9	35	9	4
47	13	45.531	13.606	1	31	1	5
48	12	35.105	16.211	20	60	20	3
49	11	27.832	18.028	9	24	9	6
50	10	3.321	24.150	9	20	9	9
51	9	-0.359	25.070	2	5	2	8
52	8	-30.096	32.498	1	12	1	11
53	7	-51.262	37.785	3	16	3	10
54	6	-65.573	41.360	2	6	2	12
55	5	-114.709	53.634	4	8	4	18
56	4	-126.807	56.656	4	9	4	27
57	3	-192.115	72.970	1	2	1	23
58	2	-242.157	85.471	3	4	3	37
59	1	-464.227	140.944	1	3	1	60

ตารางผลรวม หลังจากกระบวนการแบ่งกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 ซึ่งจะพิจารณาในส่วนของ Within cluster sum of squares หรือ SSE เพื่อพิจารณาประสิทธิภาพในการแบ่งกลุ่มข้อมูลที่เหมาะสม

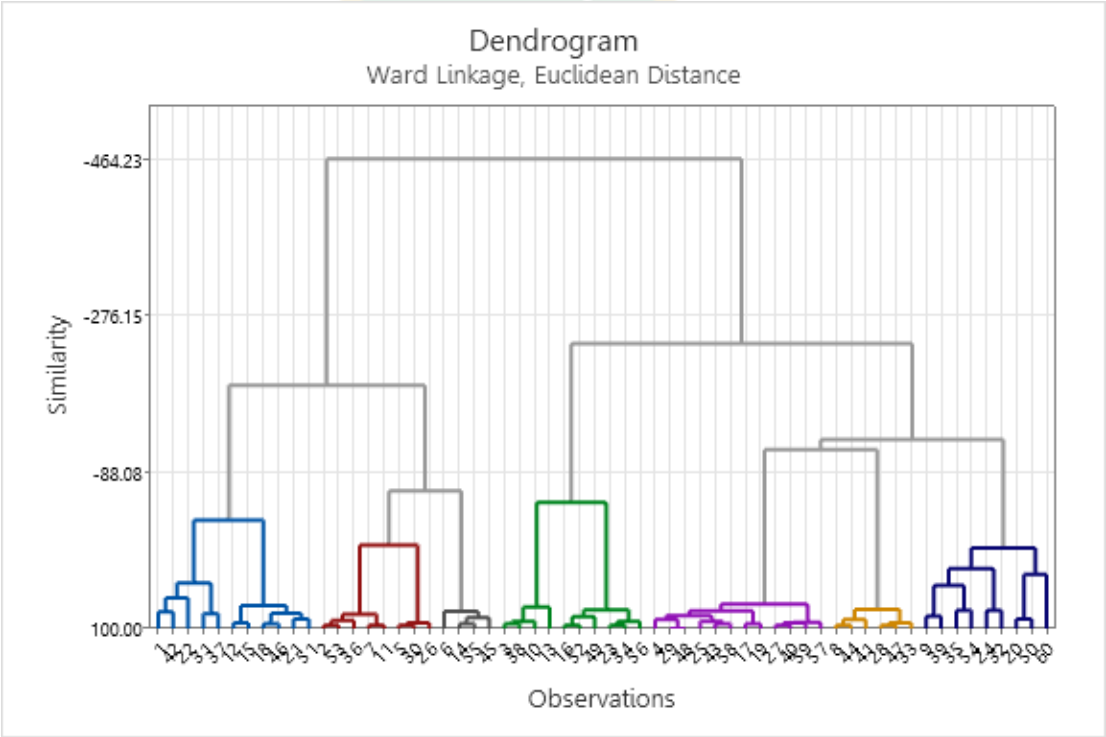
Final Partition

	Number of observations	Within cluster sum of squares	Average distance	Maximum distance

		sum of squares	from centroid	from centroid
Cluster1	11	410.909	5.70219	10.5791
Cluster2	8	128.625	3.90699	5.0544
Cluster3	10	274.900	5.05878	7.7078
Cluster4	12	43.000	1.77940	2.7234
Cluster5	4	18.000	2.03657	3.0000
Cluster6	6	17.667	1.62446	2.3921
Cluster7	9	536.444	7.44934	11.0431



แผนภาพ Dendrogram แสดงถึงผลลัพธ์ของการแบ่งกลุ่มซึ่งสามารถบอกถึงความสัมพันธ์แบบลำดับขั้นหลังจากกระบวนการแบ่งกลุ่มแบบ Hirarchical Clustering กับพารามิเตอร์ Ward Linkage และ Euclideance Distance โดยแสดงให้เห็นถึงหมายเลขของสมาชิกในแต่ละกลุ่มข้อมูล (Cluster) โดยจะแทนสีที่แตกต่างกันในแต่ละกลุ่มซึ่งคุณลักษณะของสมาชิกในแต่ละกลุ่ม (Cluster) นั้นมีความคล้ายคลึงกัน และเมื่อเทียบคุณลักษณะของสมาชิกระหว่างกลุ่มนั้นมีความแตกต่างกัน ซึ่งแผนภาพ Dendrogram จะช่วยให้มองเห็นภาพของการแบ่งกลุ่มแบบลำดับขั้นได้ดียิ่งขึ้นในมิติของการแบ่งกลุ่มในจำนวนต่าง ๆ ดังภาพที่ 21



ภาพที่ 21 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 1

ผลลัพธ์ของการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 โดยจะมี StudentID เป็นหมายเลขที่ใช้แทนรายชื่อของเด็กนักเรียน และค่าคะแนนประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) ทั้ง 8 ด้าน สุดท้ายที่คอลัม Member จะบ่งบอกถึงหมายเลขสมาชิกกลุ่ม (Cluster) ของนักเรียน ในการแบ่งกลุ่มของชุดข้อมูลนั้น ๆ ดังภาพที่ 22

	StudentID	ACT	REF	SEN	INT	VIS	VRB	SEQ	GLO		Member
1	1	5	0	5	0	0	5	0	5		1
2	2	0	7	5	0	0	8	8	0		2
3	3	0	6	7	0	7	0	0	7		3
4	4	9	0	9	0	9	0	9	0		4
5	5	0	7	7	0	0	7	0	5		2
6	6	0	6	0	7	0	5	6	0		5
7	7	0	8	8	0	0	7	8	0		2
8	8	8	0	9	0	10	0	0	7		6
9	9	5	0	6	0	5	0	6	0		7
10	10	0	8	7	0	8	0	0	6		3

ภาพที่ 22 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม

4.4.3 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2

ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2 จำนวน 24 คน โดยการแบ่งกลุ่มแบบ Hierarchical Clustering ด้วยการใช้พารามิเตอร์ Euclidean Distance, Ward Linkage โดยเลือก Cluster ที่ 7 เพราะมีความเหมาะสมในการแบ่งกลุ่มของชุดข้อมูลเพราะมี SSE ที่ค่อนข้างน้อย และอัตราการเปลี่ยนแปลงของ SSE หรือ %Change ที่เหมาะสม โดยมีกระบวนการรวมกลุ่มของการแบ่งกลุ่มแบบ Hierarchical Clustering ดังนี้

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	23	93.532	1.4142	4	8	4	2
2	22	92.078	1.7321	11	24	11	2
3	21	92.078	1.7321	18	23	18	2
4	20	92.078	1.7321	1	15	1	2
5	19	89.890	2.2103	3	18	3	3
6	18	88.796	2.4495	19	20	19	2
7	17	87.063	2.8284	7	12	7	2
8	16	86.454	2.9616	2	3	2	4
9	15	85.536	3.1623	13	17	13	2

10	14	84.156	3.4641	10	21	10	2
11	13	78.583	4.6824	5	10	5	3
12	12	78.474	4.7063	6	13	6	3
13	11	64.102	7.8485	11	16	11	3
14	10	59.426	8.8707	6	19	6	5
15	9	49.225	11.1010	11	22	11	4
16	8	49.044	11.1407	1	2	1	6
17	7	47.286	11.5251	7	9	7	3
18	6	38.540	13.4371	4	14	4	3
19	5	-3.384	22.6031	5	11	5	7
20	4	-3.914	22.7190	4	7	4	6
21	3	-51.986	33.2291	1	6	1	11
22	2	-83.101	40.0318	4	5	4	13
23	1	-169.312	58.8803	1	4	1	24

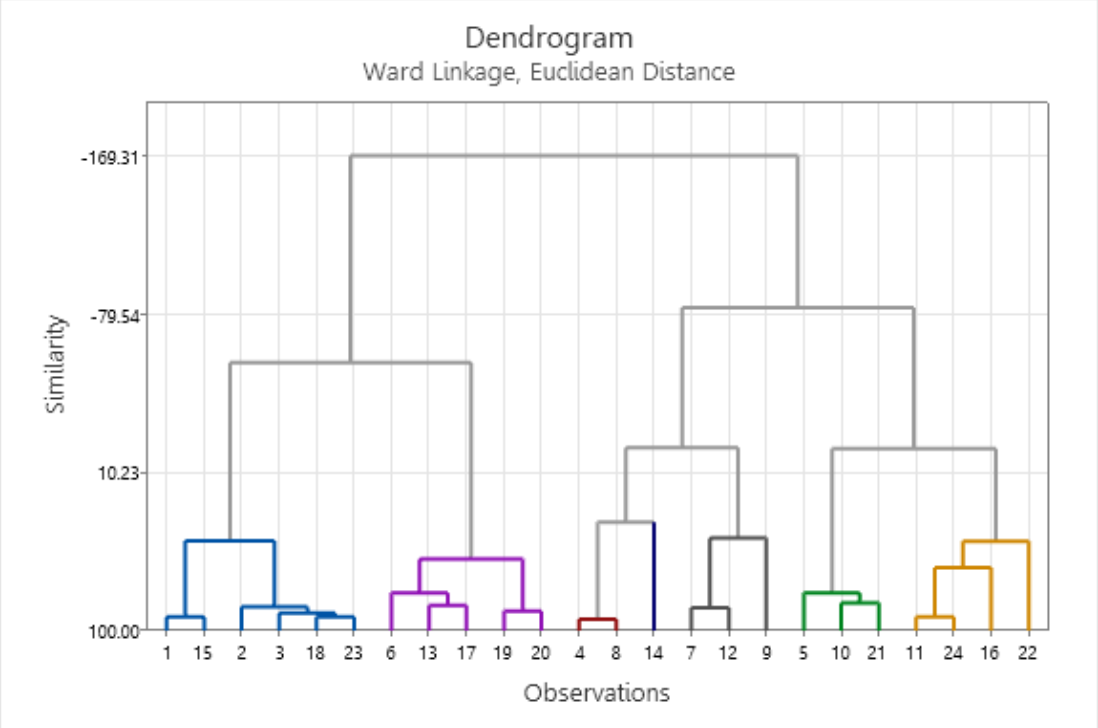
ตารางผลรวม หลังจากกระบวนการแบ่งกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2 ซึ่งจะพิจารณาในส่วนของ Within cluster sum of squares หรือ SSE เพื่อพิจารณาประสิทธิภาพในการแบ่งกลุ่มข้อมูลที่เหมาะสม

Final Partition

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	6	47.5000	2.65764	4.03113
Cluster2	2	1.0000	0.70711	0.70711
Cluster3	3	17.3333	2.32940	2.74874
Cluster4	5	52.4000	3.03043	4.63897
Cluster5	3	61.3333	4.27825	6.18241
Cluster6	4	82.5000	4.23943	6.39336

Cluster7	1	0.0000	0.00000	0.00000

แผนภาพ Dendrogram แสดงถึงผลลัพธ์ของการแบ่งกลุ่มซึ่งสามารถบอกถึงความสัมพันธ์แบบลำดับขั้นหลังจากกระบวนการแบ่งกลุ่มแบบ Hirarchical Clustering กับพารามิเตอร์ Ward Linkage และ Euclideance Distance โดยแสดงให้เห็นถึงหมายเลขของสมาชิกในแต่ละกลุ่มข้อมูล (Cluster) โดยจะแทนสีที่แตกต่างกันในแต่ละกลุ่มซึ่งคุณลักษณะของสมาชิกในแต่ละกลุ่ม (Cluster) นั้นมีความคล้ายคลึงกัน และเมื่อเทียบคุณลักษณะของสมาชิกระหว่างกลุ่มนั้นมีความแตกต่างกัน ซึ่งแผนภาพ Dendrogram จะช่วยให้มองเห็นภาพของการแบ่งกลุ่มแบบลำดับขั้นได้ดียิ่งขึ้นในมิติของการแบ่งกลุ่มในจำนวนต่าง ๆ ดังภาพที่ 23



ภาพที่ 23 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 2

ผลลัพธ์ของการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 โดยจะมี StudentID เป็นหมายเลขที่ใช้แทนรายชื่อของเด็กนักเรียน และค่าคะแนนประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) ทั้ง 8 ด้าน สุดท้ายที่คอลัม Member จะบ่งบอกถึงหมายเลขสมาชิกกลุ่ม (Cluster) ของนักเรียน ในการแบ่งกลุ่มของชุดข้อมูลนั้น ๆ ดังภาพที่ 24

	StudentID	ACT	REF	SEN	INT	VIS	VRB	SEQ	GLO		Member
1	1	8	0	8	0	10	0	10	0		1
2	2	7	0	8	0	6	0	8	0		1
3	3	6	0	6	0	6	0	6	0		1
4	4	8	0	6	0	0	9	7	0		2
5	5	4	0	0	5	5	0	0	5		3
6	6	0	7	8	0	10	0	10	0		4
7	7	8	0	0	8	0	6	7	0		5
8	8	7	0	6	0	0	8	7	0		2
9	9	6	0	0	7	0	8	0	4		5
10	10	7	0	0	4	6	0	0	6		3

ภาพที่ 24 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 2 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม

4.4.4 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3

ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3 จำนวน 48 คน โดยการแบ่งกลุ่มแบบ Hierarchical Clustering ด้วยการใช้พารามิเตอร์ Euclidean Distance, Ward Linkage โดยเลือก Cluster ที่ 7 เพราะมีความเหมาะสมในการแบ่งกลุ่มของชุดข้อมูลเพราะมี SSE ที่ค่อนข้างน้อย และอัตราการเปลี่ยนแปลงของ SSE หรือ %Change ที่เหมาะสม โดยมีกระบวนการรวมกลุ่มของการแบ่งกลุ่มแบบ Hierarchical Clustering ดังนี้

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	47	100.000	0.000	14	43	14	2
2	46	96.038	1.000	17	44	17	2
3	45	96.038	1.000	6	42	6	2
4	44	96.038	1.000	10	40	10	2
5	43	96.038	1.000	31	35	31	2
6	42	96.038	1.000	18	26	18	2
7	41	96.038	1.000	13	24	13	2
8	40	93.137	1.732	39	45	39	2
9	39	93.137	1.732	27	28	27	2

10	38	93.137	1.732	3	22	3	2
11	37	93.137	1.732	11	20	11	2
12	36	93.010	1.764	13	23	13	3
13	35	92.076	2.000	34	46	34	2
14	34	91.463	2.155	10	33	10	3
15	33	91.463	2.155	6	8	6	3
16	32	90.295	2.449	2	36	2	2
17	31	89.911	2.546	3	30	3	3
18	30	89.517	2.646	1	48	1	2
19	29	88.944	2.790	17	25	17	3
20	28	88.793	2.828	9	12	9	2
21	27	87.969	3.036	9	11	9	4
22	26	87.939	3.044	10	14	10	5
23	25	87.893	3.056	32	39	32	3
24	24	86.275	3.464	29	38	29	2
25	23	85.759	3.594	6	19	6	4
26	22	85.549	3.647	3	31	3	5
27	21	84.073	4.020	9	47	9	5
28	20	82.114	4.514	16	32	16	4
29	19	80.997	4.796	2	3	2	7
30	18	80.851	4.833	27	41	27	3
31	17	79.884	5.077	4	18	4	3
32	16	77.567	5.662	2	37	2	8
33	15	75.711	6.130	16	34	16	6
34	14	71.606	7.166	6	27	6	7
35	13	71.155	7.280	5	7	5	2
36	12	64.761	8.894	2	10	2	13
37	11	63.027	9.332	9	13	9	8
38	10	45.253	13.818	4	21	4	4
39	9	37.799	15.699	17	29	17	5

40	8	35.994	16.154	6	15	6	8
41	7	20.363	20.100	2	16	2	19
42	6	19.078	20.424	1	5	1	4
43	5	-46.521	36.980	1	4	1	8
44	4	-49.185	37.653	1	17	1	13
45	3	-151.732	63.534	1	6	1	21
46	2	-207.904	77.711	2	9	2	27
47	1	-332.047	109.044	1	2	1	48

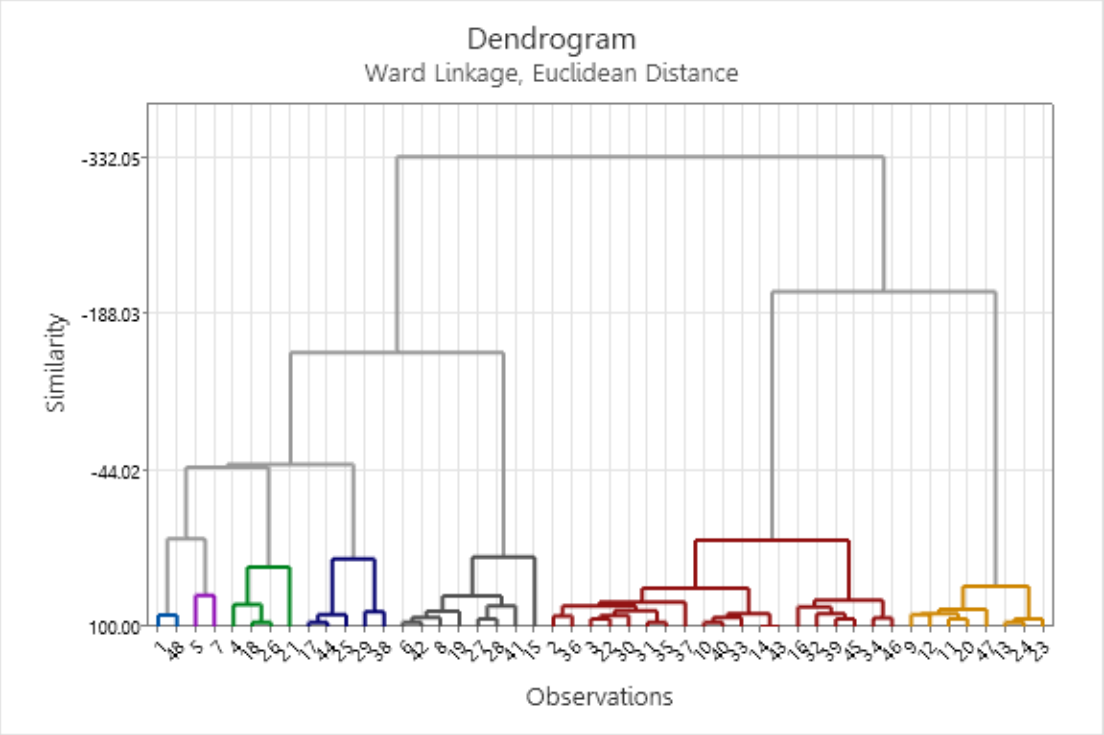
ตารางผลรวม หลังจากกระบวนการแบ่งกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3 ซึ่งจะพิจารณาในส่วนของ Within cluster sum of squares หรือ SSE เพื่อพิจารณาประสิทธิภาพในการแบ่งกลุ่มข้อมูลที่เหมาะสม

Final Partition

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	2	3.500	1.32288	1.32288
Cluster2	19	186.632	2.77356	5.66322
Cluster3	4	87.000	4.25194	7.53326
Cluster4	2	26.500	3.64005	3.64005
Cluster5	8	135.875	3.35056	9.15918
Cluster6	8	46.750	2.26776	3.79556
Cluster7	5	83.200	3.93656	5.34416

แผนภาพ Dendrogram แสดงถึงผลลัพธ์ของการแบ่งกลุ่มซึ่งสามารถบอกถึงความสัมพันธ์แบบลำดับขั้นหลังจากกระบวนการแบ่งกลุ่มแบบ Hirarchical Clustering กับพารามิเตอร์ Ward Linkage และ Euclideance Distance โดยแสดงให้เห็นถึงหมายเลขของสมาชิกในแต่ละกลุ่มข้อมูล (Cluster) โดยจะแทนสีที่แตกต่างกันในแต่ละกลุ่มซึ่งคุณลักษณะของสมาชิกในแต่ละกลุ่ม (Cluster)

นั้นมีความคล้ายคลึงกัน และเมื่อเทียบคุณลักษณะของสมาชิกระหว่างกลุ่มนั้นมีความแตกต่างกัน ซึ่งแผนภาพ Dendrogram จะช่วยให้มองเห็นภาพของการแบ่งกลุ่มแบบลำดับขั้นได้ดียิ่งขึ้นในมิติของการแบ่งกลุ่มในจำนวนต่าง ๆ ดังภาพที่ 25



ภาพที่ 25 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 1

ผลลัพธ์ของการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 โดยจะมี StudentID เป็นหมายเลขที่ใช้แทนรายชื่อของเด็กนักเรียน และค่าคะแนนประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) ทั้ง 8 ด้าน สุดท้ายที่คอลัม Member จะบ่งบอกถึงหมายเลขสมาชิกกลุ่ม (Cluster) ของนักเรียน ในการแบ่งกลุ่มของชุดข้อมูลนั้น ๆ ดังภาพที่ 26

	StudentID	ACT	REF	SEN	INT	VIS	VRB	SEQ	GLO		Member
1	1	9	0	0	9	0	9	0	8		1
2	2	5	0	7	0	9	0	4	0		2
3	3	7	0	6	0	7	0	6	0		2
4	4	6	0	7	0	10	0	0	5		3
5	5	6	0	6	0	0	6	0	5		4
6	6	6	0	0	6	8	0	7	0		5
7	7	5	0	5	0	0	5	5	0		4
8	8	7	0	0	7	7	0	8	0		5
9	9	0	5	8	0	8	0	8	0		6
10	10	7	0	7	0	8	0	7	0		2

ภาพที่ 26 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม

4.4.5 ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 4

ชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 4 จำนวน 29 คน โดยการแบ่งกลุ่มแบบ Hierarchical Clustering ด้วยการใช้พารามิเตอร์ Euclidean Distance, Ward Linkage โดยเลือก Cluster ที่ 7 เพราะมีความเหมาะสมในการแบ่งกลุ่มของชุดข้อมูลเพราะมี SSE ที่ค่อนข้างน้อย และอัตราการเปลี่ยนแปลงของ SSE หรือ %Change ที่เหมาะสม โดยมีกระบวนการรวมกลุ่มของการแบ่งกลุ่มแบบ Hierarchical Clustering ดังนี้

Amalgamation Steps

Step	Number of clusters	Similarity level	Distance level	Clusters joined		New cluster	Number of obs. in new cluster
1	28	93.497	1.4142	22	29	22	2
2	27	93.497	1.4142	4	19	4	2
3	26	93.497	1.4142	14	15	14	2
4	25	92.036	1.7321	13	16	13	2
5	24	91.702	1.8047	4	6	4	3
6	23	91.549	1.8380	11	14	11	3
7	22	90.804	2.0000	9	18	9	2
8	21	89.719	2.2361	5	7	5	2
9	20	89.272	2.3332	10	11	10	4

10	19	88.737	2.4495	8	12	8	2
11	18	87.662	2.6833	3	22	3	3
12	17	84.047	3.4696	5	23	5	3
13	16	83.868	3.5085	4	24	4	4
14	15	78.147	4.7527	5	25	5	4
15	14	77.981	4.7888	3	27	3	4
16	13	75.119	5.4113	8	21	8	3
17	12	74.399	5.5678	20	28	20	2
18	11	70.384	6.4412	3	9	3	6
19	10	56.867	9.3808	2	26	2	2
20	9	50.478	10.7703	1	17	1	2
21	8	40.413	12.9594	4	10	4	8
22	7	31.590	14.8783	2	13	2	4
23	6	24.172	16.4916	1	3	1	8
24	5	-30.822	28.4519	4	20	4	10
25	4	-71.130	37.2184	4	8	4	13
26	3	-80.092	39.1675	2	5	2	8
27	2	-134.395	50.9775	2	4	2	21
28	1	-290.274	84.8791	1	2	1	29

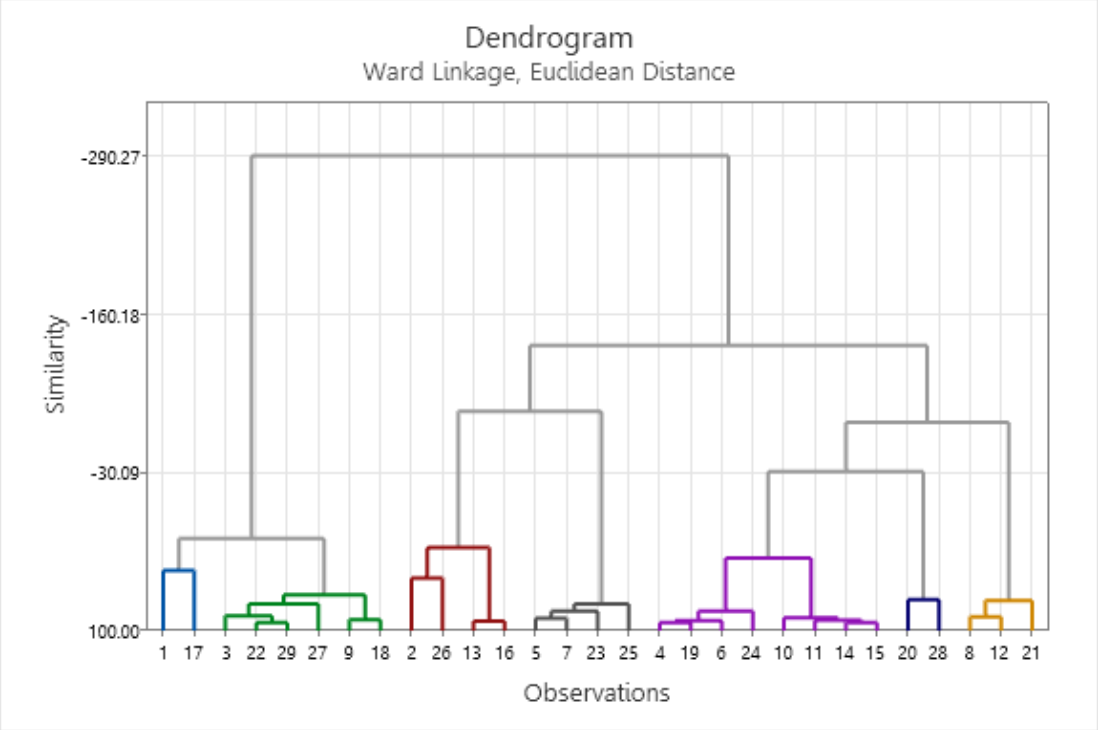
ตารางผลรวมหลังจากกระบวนการแบ่งกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 ซึ่งจะพิจารณาในส่วนของ Within cluster sum of squares หรือ SSE เพื่อพิจารณาประสิทธิภาพในการแบ่งกลุ่มข้อมูลที่เหมาะสม

Final Partition

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	2	58.000	5.38516	5.38516

Cluster2	4	127.750	5.55661	7.09313
Cluster3	6	33.833	2.17503	3.14024
Cluster4	8	56.250	2.50697	3.53111
Cluster5	4	19.250	2.14624	2.81736
Cluster6	3	16.667	2.28090	3.01846
Cluster7	2	15.500	2.78388	2.78388

แผนภาพ Dendrogram แสดงถึงผลลัพธ์ของการแบ่งกลุ่มซึ่งสามารถบอกถึงความสัมพันธ์แบบลำดับขั้นหลังจากกระบวนการแบ่งกลุ่มแบบ Hirarchical Clustering กับพารามิเตอร์ Ward Linkage และ Euclideance Distance โดยแสดงให้เห็นถึงหมายเลขของสมาชิกในแต่ละกลุ่มข้อมูล (Cluster) โดยจะแทนสีที่แตกต่างกันในแต่ละกลุ่มซึ่งคุณลักษณะของสมาชิกในแต่ละกลุ่ม (Cluster) นั้นมีความคล้ายคลึงกัน และเมื่อเทียบคุณลักษณะของสมาชิกระหว่างกลุ่มนั้นมีความแตกต่างกัน ซึ่งแผนภาพ Dendrogram จะช่วยให้มองเห็นภาพของการแบ่งกลุ่มแบบลำดับขั้นได้ดียิ่งขึ้นในมิติของการแบ่งกลุ่มในจำนวนต่าง ๆ ดังภาพที่ 27



ภาพที่ 27 Dendrogram การจัดกลุ่มของนักศึกษาปริญญาตรี ชั้นปีที่ 4

ผลลัพธ์ของการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 1 โดยจะมี StudentID เป็นหมายเลขที่ใช้แทนรายชื่อของเด็กนักเรียน และค่าคะแนนประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) ทั้ง 8 ด้าน สุดท้ายที่คอลัม Member จะบ่งบอกถึงหมายเลขสมาชิกกลุ่ม (Cluster) ของนักเรียน ในการแบ่งกลุ่มของชุดข้อมูลนั้น ๆ ดังภาพที่ 28

	StudentID	ACT	REF	SEN	INT	VIS	VRB	SEQ	GLO	Member
1	1	0	5	5	0	0	5	9	0	1
2	2	7	0	7	0	7	0	0	7	2
3	3	0	9	0	7	0	6	10	0	3
4	4	5	0	6	0	5	0	6	0	4
5	5	7	0	0	9	9	0	10	0	5
6	6	6	0	6	0	6	0	6	0	4
7	7	5	0	0	9	8	0	10	0	5
8	8	0	7	10	0	6	0	10	0	6
9	9	0	5	0	7	0	7	7	0	3
10	10	9	0	9	0	7	0	10	0	4

ภาพที่ 28 ผลสมาชิกกลุ่มชุดข้อมูลนักศึกษาปริญญาตรี ชั้นปีที่ 4 ของการแบ่งกลุ่มแบบโดยใช้พารามิเตอร์ Euclidean Distance, Ward Linkage จำนวน 7 กลุ่ม

4.5 การวัดค่าประสิทธิภาพในการจัดกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3

กระบวนการวัดค่าประสิทธิภาพในงานวิจัยนี้จะแบ่งการทดสอบเพื่อหาประสิทธิภาพแบ่งกลุ่มของข้อมูลออกเป็นกลุ่มโดยใช้ขั้นตอนการพิจารณาตารางแสดงการแบ่งกลุ่มโดยใช้ค่า Sum of Squared Errors (SSE) เพื่อเลือกพารามิเตอร์ Distance และ Linkage ที่เหมาะสมในการแบ่งกลุ่มของชุดข้อมูล โดยแบ่งกลุ่มตั้งแต่ 2 ถึง 10 กลุ่ม (K = 2-10) โดยแบ่งขั้นตอนดังนี้

4.5.1 ขั้นตอนที่ 1 การพิจารณาค่าเฉลี่ยของพารามิเตอร์ Linkage ที่เหมาะสม

การพิจารณาพารามิเตอร์ Linkage ที่เหมาะสมนั้นจะพิจารณาใช้ค่าเฉลี่ย Sum of Squared Errors หรือ SSE ของพารามิเตอร์ Linkage ที่มีค่าเฉลี่ยน้อยที่สุดเป็นอันดับแรก จากตารางการแบ่งกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดังตารางที่ 8 พบว่าพารามิเตอร์ Euclidean Distance และ Linkage Ward ให้ค่าเฉลี่ยของ Sum of Squared Errors หรือ SSE น้อยที่สุดที่ 397.43

4.5.2 ขั้นตอนที่ 2 การพิจารณาจำนวนคลัสเตอร์ของพารามิเตอร์ Linkage ที่เหมาะสม

เมื่อได้พารามิเตอร์ Linkage ที่เหมาะสมแล้ว ถัดมาจะพิจารณาจำนวนของกลุ่ม หรือคลัสเตอร์ (Cluster) ที่เหมาะสมโดยการพิจารณากำหนด %Change ที่ยอมรับได้โดยชุดข้อมูลของนักเรียนชั้นมัธยมศึกษาปีที่ 3 นั้นจะพิจารณาค่า %Change ที่ 15% และเลือกจำนวนคลัสเตอร์ที่เข้าเงื่อนไข %Change ที่พิจารณา พบว่าจำนวนของกลุ่มที่เหมาะสมของชุดข้อมูลอยู่ที่ 6 กลุ่ม (Cluster

6) ซึ่งมีค่าของ %Change ที่ 19.92% ซึ่งเป็นจุดที่เหมาะสมกับการแบ่งกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดังตารางที่ 8

ตารางที่ 8 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่ม ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Euclidean Distance และLinkage จำนวน 2 - 10 กลุ่ม

Number of Cluster	Sum of Squared Errors									
	Euclidean Distance									
	Average		Complete		Single		Ward		Median	
	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change
K=2	930.77		930.77		930.77		765.71		930.77	
K=3	865.09	7.59	665.27	39.91	865.09	7.59	574.83	33.21	865.09	7.59
K=4	755.90	14.44	486.30	36.80	755.90	14.44	478.65	20.09	748.24	15.62
K=5	494.96	52.72	426.46	14.03	693.67	8.97	415.36	15.24	724.64	3.26
K=6	351.14	40.96	351.14	21.45	490.70	41.36	346.36	19.92	481.33	50.55
K=7	310.43	13.12	301.57	16.44	448.85	9.33	305.64	13.32	456.13	5.52
K=8	289.43	7.26	275.06	9.64	425.88	5.39	260.86	17.17	434.13	5.07
K=9	251.76	14.96	249.06	10.44	418.88	1.67	228.19	14.32	336.48	29.02
K=10	234.76	7.24	225.36	10.52	384.42	8.97	201.31	13.35	329.48	2.12
Average	498		434		601		**397**		589	

	.25		.55		.57		**.43**		.59	

สำหรับการแบ่งกลุ่มข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้พารามิเตอร์ Manhattan Distance กับพารามิเตอร์ Linkage ที่หลากหลายนั้นพบว่าค่าเฉลี่ย Sum of Squared Errors หรือ SSE ของพารามิเตอร์ Linkage ที่น้อยที่สุด ดังตารางที่ 9 พบว่าพารามิเตอร์ Manhattan Distance และ Linkage Ward ให้ค่าเฉลี่ยของ Sum of Squared Errors หรือ SSE น้อยที่สุดอยู่ที่ 417.45

การพิจารณาจำนวนคลัสเตอร์ของพารามิเตอร์ Linkage ที่เหมาะสม เมื่อได้พารามิเตอร์ Linkage ที่เหมาะสมแล้ว ถัดมาจะพิจารณาจำนวนของกลุ่ม หรือคลัสเตอร์ (Cluster) ที่เหมาะสมโดยการพิจารณากำหนด %Change ที่ยอมรับได้โดยชุดข้อมูลของนักเรียนชั้นมัธยมศึกษาปีที่ 3 นั้นจะพิจารณาค่า %Change ที่ 15% และเลือกจำนวนคลัสเตอร์ที่เข้าเงื่อนไข %Change ที่พิจารณาพบว่าจำนวนของกลุ่มที่เหมาะสมของชุดข้อมูลอยู่ที่ 5 กลุ่ม (Cluster 5) ซึ่งมีค่าของ %Change ที่ 17.83% ซึ่งเป็นจุดที่เหมาะสมกับการแบ่งกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ดังตารางที่ 9 ซึ่งหากเปรียบเทียบระหว่างพารามิเตอร์ Euclidean Distance ดังตารางที่ 8 และ Manhattan Distance ดังตารางที่ 9

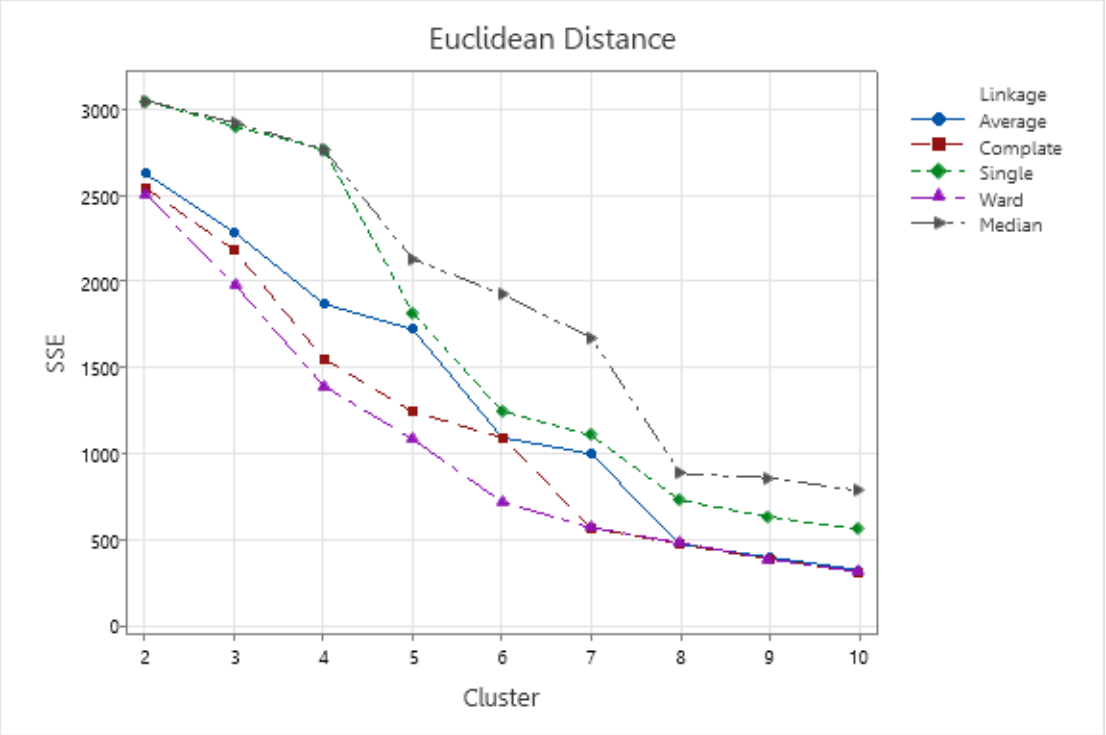
แต่เมื่อพิจารณาเปรียบเทียบระหว่างพารามิเตอร์ Euclidean Distance ดังตารางที่ 8 และพารามิเตอร์ Manhattan Distance ดังตารางที่ 9 พบว่าพารามิเตอร์ Euclidean Distance นั้นให้ค่าเฉลี่ยของ Sum of Squared Errors หรือ SSE ที่น้อยกว่าพารามิเตอร์ Manhattan Distance งานวิจัยนี้จึงเลือกวิธีการแบ่งกลุ่มข้อมูลโดยใช้ Euclidean Distance และ Ward Linkage ในการแบ่งกลุ่มของชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3

ตารางที่ 9 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่ม ชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Manhattan Distance และLinkage จำนวน 2-10 กลุ่ม

Number of Cluster	Sum of Squared Errors									
	Manhattan Distance									
	Average		Complete		Single		Ward		Median	
	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change
K=2	930.77		831.58		935.36		765.71		930.765	

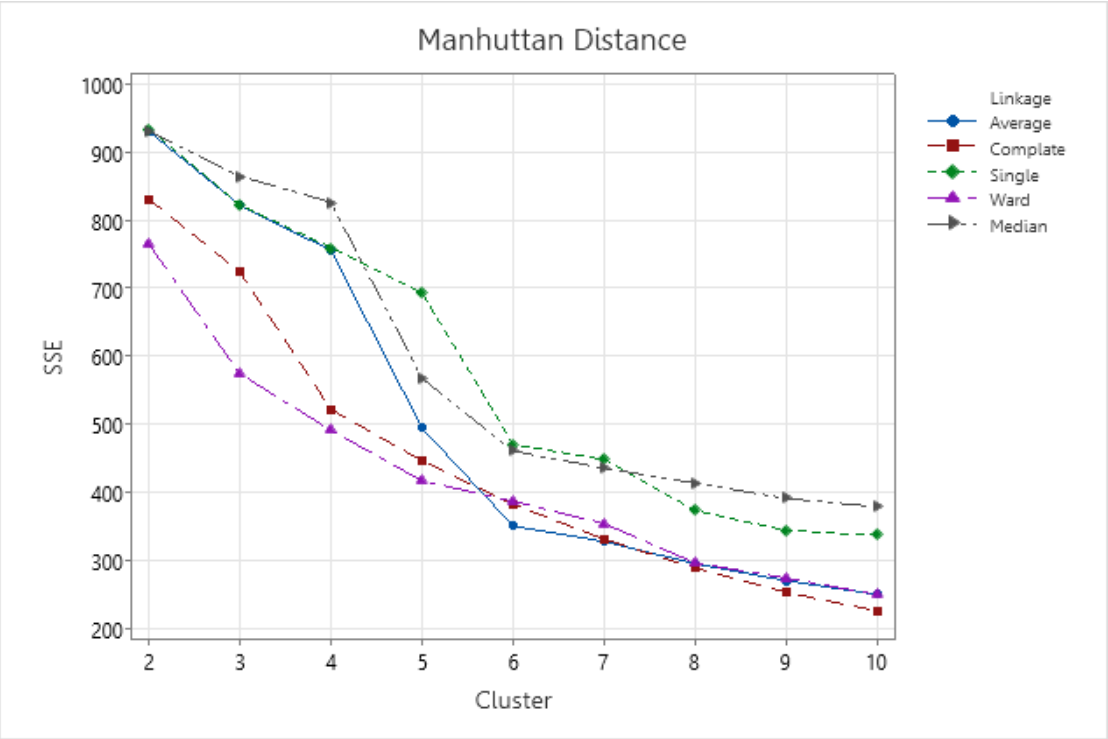
K=3	822.38	13.18	724.54	14.77	822.38	13.74	574.83	33.21	865.091	7.59
K=4	755.90	8.79	522.42	38.69	760.03	8.20	491.88	16.86	826.438	4.68
K=5	494.96	52.72	446.95	16.89	693.67	9.57	417.45	17.83	568.237	45.44
K=6	351.14	40.96	383.11	16.66	470.85	47.32	387.43	7.75	461.24	23.20
K=7	328.38	6.93	331.71	15.50	448.85	4.90	354.76	9.21	436.04	5.78
K=8	295.71	11.05	289.21	14.69	373.89	20.05	296.83	19.52	414.04	5.31
K=9	270.80	9.20	253.69	14.00	344.56	8.51	274.45	8.15	391.417	5.78
K=10	250.50	8.10	226.14	12.18	337.56	2.07	250.45	9.58	379.417	3.16
Average	500.06		445.49		576.35		423.75		585.85	

สำหรับการพิจารณากราฟของพารามิเตอร์ Euclidean Distance จะสร้างกราฟแสดงการแบ่งกลุ่มของชุดข้อมูลโดยใช้ SSE ในการประเมินประสิทธิภาพของการแบ่งกลุ่ม โดยใช้พารามิเตอร์ Euclidean Distance และ Linkage ต่าง ๆ ในการแบ่งกลุ่มระหว่าง 2 ถึง 10 กลุ่ม (K = 2-10) ซึ่งจะช่วยให้หาจุดที่เหมาะสมที่สุดในการแบ่งกลุ่มชุดข้อมูลด้วยวิธีการ Elbow และ %Change ดังแสดงในภาพที่ 29 และเมื่อพิจารณากราฟการแบ่งกลุ่มด้วย Euclidean Distance และ Linkage ต่าง ๆ พบว่าเส้นกราฟในส่วนของเส้น Ward Linkage มีอัตราการลดลงของ SSE ที่สูงมากเมื่อเปรียบเทียบกับกราฟอื่น ๆ และจำนวนคลัสเตอร์เพิ่มขึ้นโดยมีอัตราการลดลงที่ดีกว่าเส้น Linkage อื่น ๆ จนถึงจำนวนคลัสเตอร์ที่ 6 (Cluster 6) พบว่ากราฟมีความชันที่น้อยลงคล้ายจุดข้อศอก (Elbow) เมื่อพิจารณาร่วมกับตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่ม สำหรับชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในตารางที่ 8 นั้น เราพบว่าค่า SSE และ %Change สอดคล้องกันเมื่อพิจารณาร่วมกับกราฟ ซึ่งช่วยให้เราสามารถหาจุดที่เหมาะสมสำหรับการแบ่งกลุ่มข้อมูลชั้นมัธยมศึกษาปีที่ 3



ภาพที่ 29 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Euclidean Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change

สำหรับการพิจารณากราฟของพารามิเตอร์ Manhattan Distance และ Linkage ต่าง ๆ เพื่อแบ่งกลุ่มชุดข้อมูลด้วยวิธี Elbow และ %Change สำหรับการหาจุดที่เหมาะสมในการแบ่งกลุ่มระหว่าง 2 ถึง 10 กลุ่ม (K = 2-10) ดังภาพที่ 30 พบว่ากราฟการแบ่งกลุ่มโดยใช้ Manhattan Distance และ Ward Linkage เป็นการแบ่งกลุ่มที่เหมาะสมที่สุด เนื่องจากพารามิเตอร์ Ward Linkage มีอัตราการลดลงของ SSE ที่สูงสุดเมื่อเปรียบเทียบกับกราฟอื่น ๆ และกราฟมีความชันที่น้อยลงคล้ายจุดข้อศอกเมื่อพิจารณาร่วมกับตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มตารางที่ 9 ดังนั้น จำนวนคลัสเตอร์ที่เหมาะสมในการแบ่งกลุ่มชุดข้อมูลนักเรียนชั้นมัธยมศึกษาปีที่ 3 สำหรับพารามิเตอร์ Manhattan Distance คือ 5 กลุ่ม (Cluster 5)



ภาพที่ 30 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Manhattan Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change

4.6 การวัดค่าประสิทธิภาพในการจัดกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3

กระบวนการวัดค่าประสิทธิภาพในงานวิจัยนี้จะแบ่งการทดสอบเพื่อหาประสิทธิภาพแบ่งกลุ่มของข้อมูลโดยมีขั้นตอนพิจารณาตารางแสดงการแบ่งกลุ่มโดยใช้ค่า Sum of Squared Errors หรือ SSE ในการพิจารณา ซึ่งการแบ่งกลุ่มโดยใช้พารามิเตอร์ Euclidean Distance และLinkage ต่าง ๆ ในการแบ่งกลุ่มระหว่าง 2 ถึง 10 กลุ่ม หรือ (K = 2-10) เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มชุดข้อมูลตามลำดับดังนี้

4.6.1 ขั้นตอนที่ 1 การพิจารณาค่าเฉลี่ยของพารามิเตอร์ Linkage ที่เหมาะสม

การพิจารณาพารามิเตอร์ Linkage ที่เหมาะสมนั้นจะพิจารณาใช้ค่าเฉลี่ย Sum of Squared Errors หรือ SSE ของพารามิเตอร์ Linkage ที่มีค่าเฉลี่ยน้อยที่สุดเป็นอันดับแรก จากตารางการแบ่งกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ดังตารางที่ 10 พบว่าพารามิเตอร์ Euclidean Distance และ Linkage Ward ให้ค่าเฉลี่ยของ Sum of Squared Errors หรือ SSE น้อยที่สุดที่ 26.16

4.6.2 ขั้นตอนที่ 2 การพิจารณาจำนวนคลัสเตอร์ของพารามิเตอร์ Linkage ที่เหมาะสม

เมื่อได้พารามิเตอร์ Linkage ที่เหมาะสมแล้ว ถัดมาจะพิจารณาจำนวนของกลุ่ม หรือคลัสเตอร์ (Cluster) ที่เหมาะสมโดยการพิจารณากำหนด %Change ที่ยอมรับได้โดยชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 นั้นจะพิจารณาค่า %Change ที่ 25% และเลือกจำนวนคลัสเตอร์ที่เข้าเงื่อนไขที่พิจารณา พบว่าจำนวนของกลุ่มที่เหมาะสมของชุดข้อมูลอยู่ที่ 7 กลุ่ม (Cluster 7) ซึ่งมีค่าของ %Change ที่ 26.16% ซึ่งเป็นจุดที่เหมาะสมกับการแบ่งกลุ่มของชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ดังตารางที่ 10

ตารางที่ 10 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Euclidean Distance และLinkage จำนวน 2 - 10 กลุ่ม

Number of Cluster	Sum of Squared Errors									
	Euclidean Distance									
	Average		Complete		Single		Ward		Median	
	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change
K=2	2628.99	-	2548.23	-	3055.61	-	2508.85	-	3055.61	-
K=3	2286.97	14.955159 01	2181.25	16.824297 99	2901.9	5.2968744 62	1983.86	26.463056 87	2924.74	4.4745857 75
K=4	1871.89	22.174379 9	1552.801	40.471960 03	2771.18	4.7171241 13	1394.8	42.232578 15	2771.18	5.5413217 47
K=5	1722.89	8.6482596 1	1243.445	24.878945 19	1819.25	52.325408 82	1084.707	28.58772	2132.75	29.934591 49
K=6	1094.445	57.421341 41	1094.445	13.614206 29	1248.167	45.753733 27	718.457	50.977302 75	1926.84	10.686408 84

K=7	998.57	9.601229759	569.457	92.19098194	1106.293	12.82426988	569.457	26.16527675	1675.11	15.02766982
K=8	473.582	110.85472	473.582	20.24464612	728.493	51.86048459	482.761	17.95836863	885.629	89.14353527
K=9	397.915	19.01587022	386.8859	22.40869983	632.314	15.21063902	386.8859	24.78123395	859.129	3.084519321
K=10	324.715	22.54284526	311.2192	24.31299226	559.114	13.09214221	313.6859	23.33544479	785.929	9.313818424
Average	1311.107444		1151.257233		1646.924556		1049.273756		1890.768556	

สำหรับการแบ่งกลุ่มข้อมูลนักศึกษาระดับปริญญาตรีชั้นปีที่ 3 โดยใช้พารามิเตอร์ Manhattan Distance กับพารามิเตอร์ Linkage ที่หลากหลายนั้นพบว่าค่าเฉลี่ย Sum of Squared Errors หรือ SSE ของพารามิเตอร์ Linkage ที่น้อยที่สุด ดังตารางที่ 11 พบว่าพารามิเตอร์ Manhattan Distance และ Linkage Ward ให้ค่าเฉลี่ยของ Sum of Squared Errors หรือ SSE น้อยที่สุดอยู่ที่ 1065.167

การพิจารณาจำนวนคลัสเตอร์ของพารามิเตอร์ Linkage ที่เหมาะสม เมื่อได้พารามิเตอร์ Linkage ที่เหมาะสมแล้ว ถัดมาจะพิจารณาจำนวนของกลุ่ม หรือคลัสเตอร์ (Cluster) ที่เหมาะสมโดยการพิจารณากำหนด %Change ที่ยอมรับได้โดยชุดข้อมูลของนักศึกษาระดับปริญญาตรีชั้นปีที่ 3 นั้นจะพิจารณาค่า %Change ที่ 25% และเลือกจำนวนคลัสเตอร์ที่เข้าเงื่อนไข %Change ที่พิจารณาพบว่าจำนวนของกลุ่มที่เหมาะสมของชุดข้อมูลอยู่ที่ 7 กลุ่ม (Cluster 7) ซึ่งมีค่าของ %Change ที่ 25.33% ซึ่งเป็นจุดที่เหมาะสมกับการแบ่งกลุ่มของชุดข้อมูลนักศึกษาระดับปริญญาตรีชั้นปีที่ 3 ดังตารางที่ 11

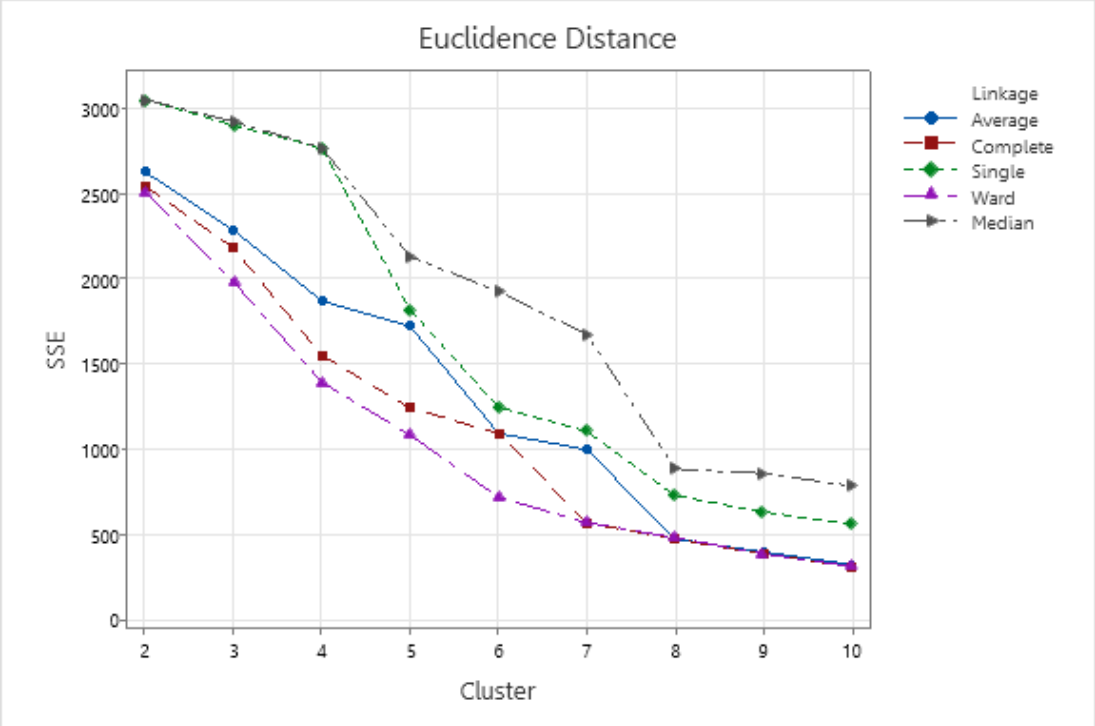
แต่เมื่อพิจารณาเปรียบเทียบระหว่างพารามิเตอร์ Euclidean Distance ดังตารางที่ 10 และพารามิเตอร์ Manhattan Distance ดังตารางที่ 11 พบว่าพารามิเตอร์ Euclidean Distance นั้นให้

ค่าเฉลี่ยของ Sum of Squared Errors หรือ SSE ที่น้อยกว่าพารามิเตอร์ Manhattan Distance งานวิจัยนี้จึงเลือกวิธีการแบ่งกลุ่มข้อมูลโดยใช้ Euclidean Distance และ Ward Linkage ในการแบ่งกลุ่มของชุดข้อมูลนักศึกษาระดับปริญญาตรีชั้นที่ 3

ตารางที่ 11 ตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ของการแบ่งกลุ่มแบบโดยใช้หลายพารามิเตอร์ Manhattan Distance และLinkage จำนวน 2 - 10 กลุ่ม

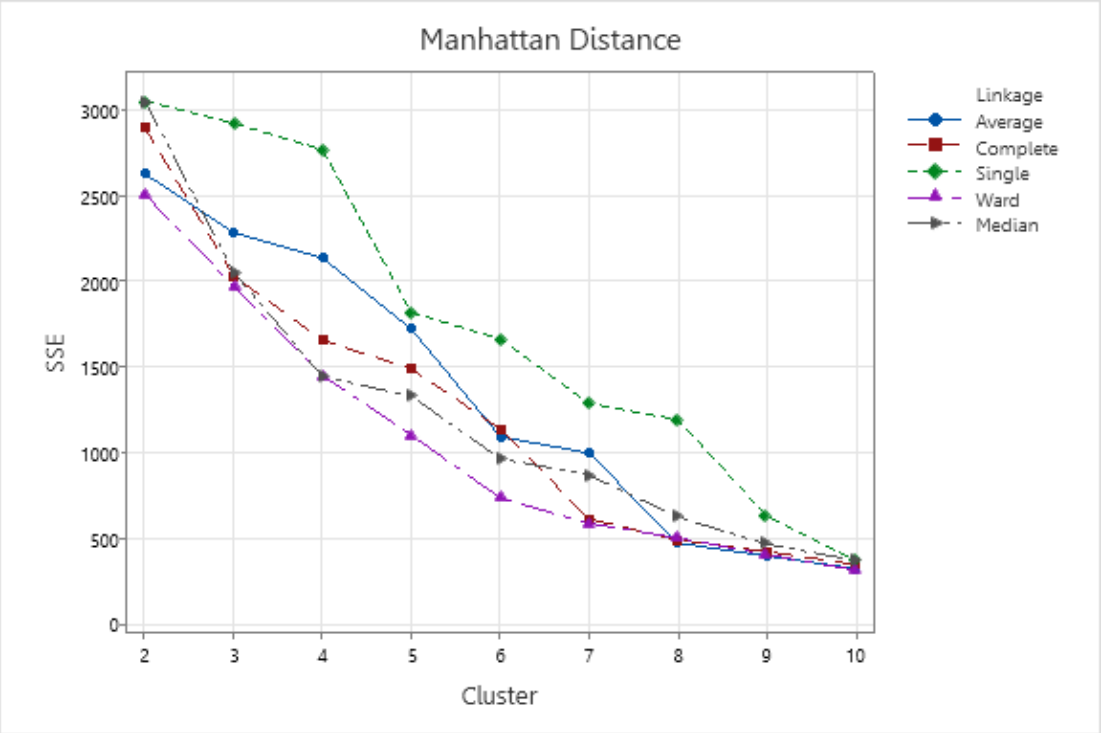
Number of Cluster	Sum of Squared Errors									
	Manhattan Distance									
	Average		Complete		Single		Ward		Median	
	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change	SSE	%Change
K=2	2628.99	-	2899.62	-	3055.61	-	2508.85	-	3055.61	-
K=3	2286.97	14.95516	2032.61	42.65501	2924.74	4.474586	1974.466	27.06474	2053.53	48.79792
K=4	2137.97	6.969228	1657.915	22.60037	2771.18	5.541322	1449.478	36.21911	1445.819	42.0323
K=5	1722.89	24.09208	1489.665	11.29449	1819.25	52.32541	1103.465	31.35695	1331.013	8.62546
K=6	1094.445	57.42134	1139.37	30.74462	1666.31	9.178364	737.215	49.68022	966.242	37.75152
K=7	998.57	9.60123	614.382	85.44977	1288.51	29.32069	588.215	25.33087	870.367	11.01547
K=8	473.582	110.8547	492.632	24.71419	1193.87	7.927161	502.742	17.00136	624.878	39.28591
K=9	397.915	19.01587	423.8333	16.23249	632.314	88.80967	407.1591	23.47557	467.283	33.72581
K=10	324.715	22.54285	348.1666	21.7329	371.415	70.24461	314.9091	29.29417	371.415	25.81156
Average	1340.672		1233.133		1747.022		1065.167		1242.906	

สำหรับการพิจารณากราฟของพารามิเตอร์ Euclidean Distance จะสร้างกราฟแสดงการแบ่งกลุ่มของชุดข้อมูลโดยใช้ SSE ในการประเมินประสิทธิภาพของการแบ่งกลุ่ม โดยใช้พารามิเตอร์ Euclidean Distance และ Linkage ต่าง ๆ ในการแบ่งกลุ่มระหว่าง 2 ถึง 10 กลุ่ม (K = 2-10) ซึ่งจะช่วยให้หาจุดที่เหมาะสมที่สุดในการแบ่งกลุ่มชุดข้อมูลด้วยวิธีการ Elbow และ %Change ดังแสดงในภาพที่ 31 และเมื่อพิจารณากราฟการแบ่งกลุ่มด้วย Euclidean Distance และ Linkage ต่าง ๆ พบว่าเส้นกราฟในส่วนของเส้น Ward Linkage มีอัตราการลดลงของ SSE ที่สูงมากเมื่อเปรียบเทียบกับกราฟอื่น ๆ และจำนวนคลัสเตอร์เพิ่มขึ้นโดยมีอัตราการลดลงที่ดีกว่าเส้น Linkage อื่น ๆ จนถึงจำนวนคลัสเตอร์ที่ 7 (Cluster 7) พบว่ากราฟมีความชันที่น้อยลงคล้ายจุดข้อศอก (Elbow) เมื่อพิจารณาร่วมกับตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่ม สำหรับชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 ในตารางที่ 10 นั้น เราพบว่าค่า SSE และ %Change สอดคล้องกันเมื่อพิจารณาร่วมกับกราฟ ซึ่งช่วยให้เราสามารถหาจุดที่เหมาะสมสำหรับการแบ่งกลุ่มข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3



ภาพที่ 31 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Euclidean Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change

สำหรับการพิจารณากราฟของพารามิเตอร์ Manhattan Distance และ Linkage ต่าง ๆ เพื่อแบ่งกลุ่มชุดข้อมูลด้วยวิธี Elbow และ %Change สำหรับการหาจุดที่เหมาะสมในการแบ่งกลุ่มระหว่าง 2 ถึง 10 กลุ่ม (K = 2-10) ดังภาพที่ 32 พบว่ากราฟการแบ่งกลุ่มโดยใช้ Manhattan Distance และ Ward Linkage เป็นการแบ่งกลุ่มที่เหมาะสมที่สุด เนื่องจากพารามิเตอร์ Ward Linkage มีอัตราการลดลงของ SSE ที่สูงสุดเมื่อเปรียบเทียบกับกราฟอื่น ๆ และกราฟมีความชันที่น้อยลงคล้ายจุดข้อศอกเมื่อพิจารณาร่วมกับตารางเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มตารางที่ 11 ดังนั้น จำนวนคลัสเตอร์ที่เหมาะสมในการแบ่งกลุ่มชุดข้อมูลนักศึกษาปริญญาตรีชั้นปีที่ 3 สำหรับพารามิเตอร์ Manhattan Distance คือ 7 กลุ่ม (Cluster 7)



ภาพที่ 32 กราฟแสดงการแบ่งกลุ่มโดยใช้ SSE ของการแบ่งกลุ่มโดยใช้พารามิเตอร์ Manhattan Distance และLinkage ต่าง ๆ เพื่อหาจุดที่เหมาะสมในการแบ่งกลุ่มด้วยวิธีการ Elbow และ %Change

4.7 การนำไปใช้งาน (Deployment)

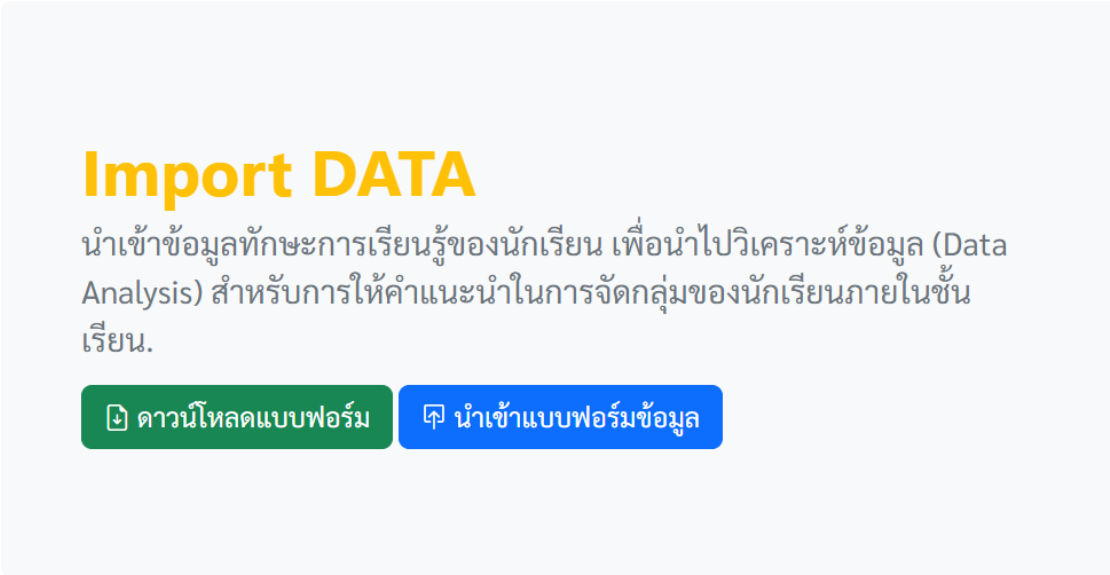
จากการกระบวนการสร้างโมเดลเพื่อวิเคราะห์ข้อมูลรูปแบบการเรียนรู้ของนักเรียนผ่านเทคนิคเหมืองข้อมูล โดยมีวัตถุประสงค์เพื่อให้คำแนะนำสำหรับอาจารย์ผู้สอนสำหรับการแบ่งกลุ่มใน

ชั้นเรียน และการรับรู้ถึงลักษณะการเรียนรู้ของเด็กนักเรียน นักศึกษา จึงนำมาสู่การออกแบบ และพัฒนาโปรแกรมในรูปแบบของเว็บไซต์เพื่อความสะดวกต่อการใช้งานระบบสนับสนุน และลดขั้นตอนกระบวนการในการทำงานของอาจารย์ผู้สอน ซึ่งมีกระบวนการใช้งาน ดังภาพที่ 33

ภาพที่ 33 ผังงาน FlowChart แสดงการทำงานของระบบสนับสนุนการสอน โดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้

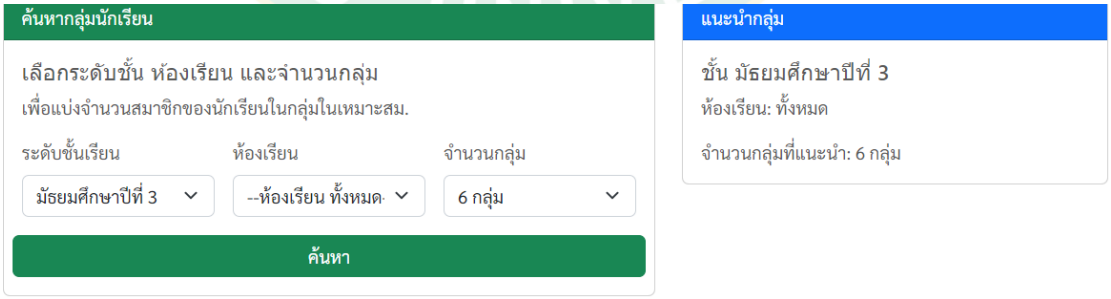
การออกแบบข้อมูลนำเข้าในส่วนการพัฒนาระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ ผู้วิจัยออกแบบระบบในรูปแบบเว็บไซต์ โดยภาษาที่ใช้พัฒนาคือภาษา PHP ในส่วนของฐานข้อมูลและการจัดเก็บข้อมูล และติดต่อกับฐานข้อมูลโดยใช้ My SQL เพื่อบันทึกข้อมูลลงในฐานข้อมูล

ซึ่งการนำเข้าข้อมูลรูปแบบการเรียนรู้ของนักเรียน จะใช้เป็นรูปแบบของการ import Excel ไฟล์ตามรูปแบบแบบฟอร์มที่ให้ดาวน์โหลด ดังภาพที่ 34



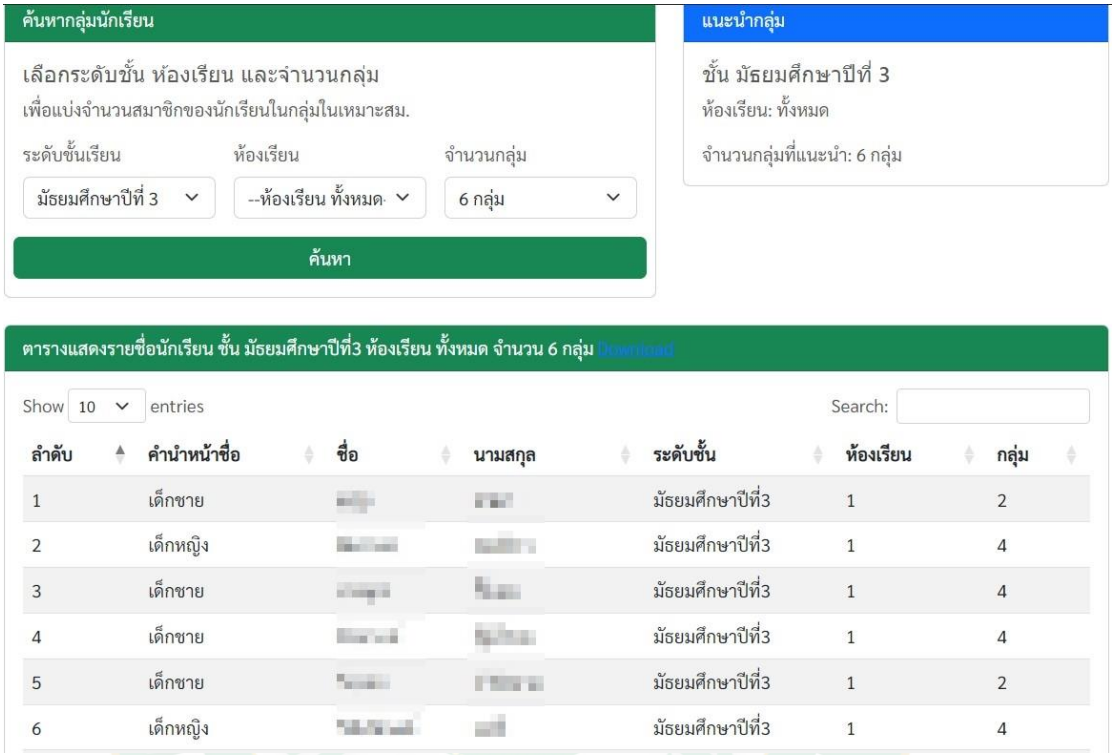
ภาพที่ 34 หน้าจอการนำเข้าข้อมูลทักษะการเรียนรู้ของนักเรียน และดาวน์โหลดแบบฟอร์มการนำเข้าข้อมูล

หน้าจอแสดงการแนะนำกลุ่มที่เหมาะสมกับชุดข้อมูลระดับชั้นเรียน ในกรณีที่มีการระบุชั้นเรียน ห้องเรียน จำนวนกลุ่ม และกดค้นหา ระบบจะแสดงหน้าต่างแนะนำกลุ่มที่เหมาะสมกับชั้นเรียนนั้น ๆ :ซึ่งนอกเหนือจากที่ระบบจะแนะนำกลุ่มให้กับผู้ใช้งานระบบแล้ว ผู้ใช้งานสามารถปรับแต่งจำนวนกลุ่มที่ต้องการ ห้องเรียนที่ต้องการตามความต้องการได้อีกด้วย ดังภาพที่ 35



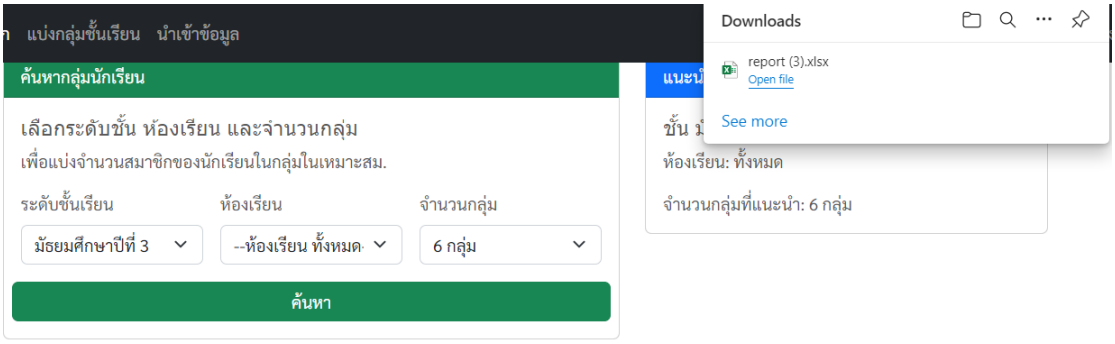
ภาพที่ 35 หน้าจอแสดงการแนะนำกลุ่มที่เหมาะสมกับชุดข้อมูลระดับชั้นเรียน

การค้นหารายชื่อนักเรียน ชั้นเรียน ห้องเรียน และจำนวนกลุ่มที่ต้องการพร้อมกับแสดงกลุ่มที่แนะนำสำหรับเด็กนักเรียนช่วยให้ผู้สอนสามารถรับรู้ถึงสมาชิกในแต่ละกลุ่มภายในชั้นเรียน เพื่อดำเนินการแบ่งกลุ่มนักเรียนตามความต้องการ ดังภาพที่ 36



ภาพที่ 36 หน้าจอแสดงผลการค้นหารายชื่อนักเรียน ชั้นเรียน ห้องเรียน และจำนวนกลุ่ม

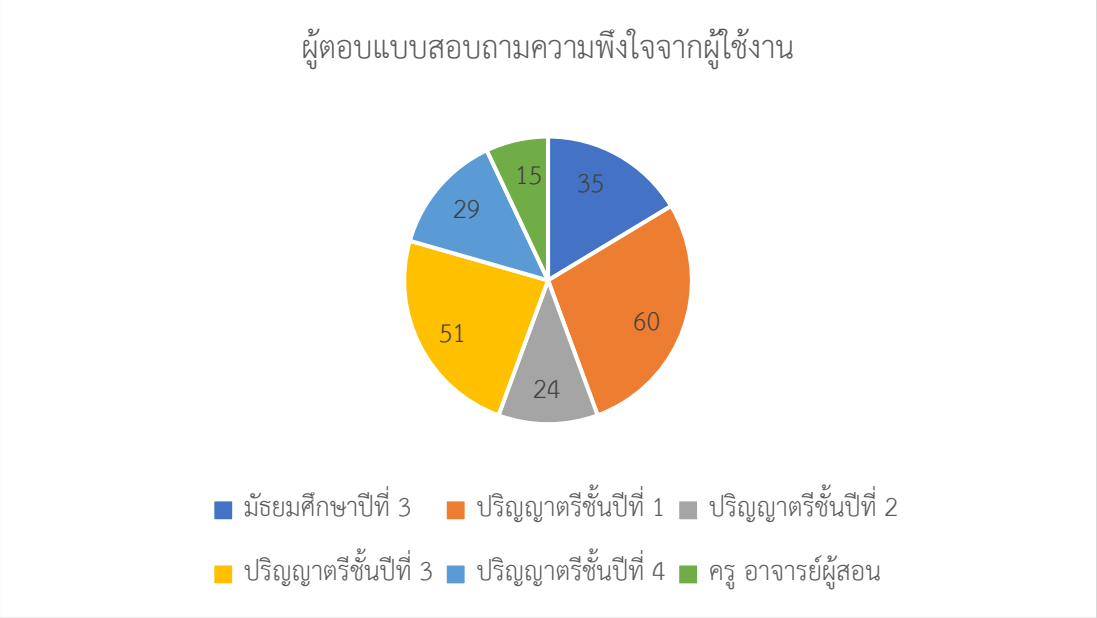
การค้นหานักเรียนโดยการระบุชั้นเรียน ห้องเรียน และจำนวนกลุ่มที่ต้องการแสดงรายชื่อ ระบบยังสามารถดาวน์โหลดไฟล์ในรูปแบบ Excel File รายชื่อเด็กนักเรียนเพื่อให้อาจารย์ผู้สอนสามารถนำไปใช้งานต่อได้ ดังภาพที่ 37



ภาพที่ 37 หน้าจอแสดงไฟล์ดาวน์โหลด Excel File รายชื่อนักเรียนจากการแบ่งกลุ่มชั้นเรียน

4.8 การประเมินความพึงพอใจจากผู้ใช้งาน

ตารางการประเมินความพึงพอใจต่อการใช้งานระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ มีจุดมุ่งหมายเพื่อตรวจสอบว่าการแบ่งกลุ่มนักเรียนในชั้นเรียนสามารถส่งผลต่อการเรียนรู้และผลการเรียนของนักเรียนได้หรือไม่ รวมถึงการสร้างความสัมพันธ์ระหว่างนักเรียน และอาจารย์ผู้สอน โดยมีผู้ตอบแบบสอบถามสถานะนักเรียน และนักศึกษาจำนวน 199 คน และอาจารย์ผู้สอนจำนวน 15 คนในสถานบันการศึกษา ดังภาพที่ 38



ภาพที่ 38 ผู้ตอบแบบสอบถามความพึงพอใจผู้ใช้งานระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้

แบบสอบถามความพึงพอใจผู้ใช้งานระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ ซึ่งมีคะแนนระดับความพึงพอใจดังนี้ มากที่สุด (5), มาก (4), ปานกลาง (3), น้อย (2), น้อยที่สุด (1) โดยแบ่งด้านการประเมินออกเป็น 3 ด้าน ได้แก่

1. ด้านความพึงพอใจต่อการแบ่งกลุ่ม พบว่ามีค่าเฉลี่ยคะแนนประเมินอยู่ที่ 4.16 คะแนน ซึ่งอยู่ในระดับพอใจมาก

2. ด้านการใช้งานระบบสนับสนุนการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ พบว่ามีค่าเฉลี่ยคะแนนประเมินอยู่ที่ 4.16 คะแนน ซึ่งอยู่ในระดับพอใจมาก

3. ด้านความพึงพอใจต่อระบบสนับสนุนการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ พบว่ามีค่าเฉลี่ยคะแนนประเมินอยู่ที่ 4.49 คะแนน ซึ่งอยู่ในระดับพอใจมาก

ซึ่งสรุปโดยภาพรวมของแบบประเมินความพึงพอใจต่อการใช้ระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้อยู่ในระดับ พอใจมาก ในทุกด้านการประเมิน โดยมีภาพรวมของการประเมินคะแนน ดังตารางที่ 12

ตารางที่ 12 ตารางสรุปคะแนนจากแบบสอบถามความพึงพอใจผู้ใช้งานระบบสนับสนุนการสอนโดยการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้

สถานะผู้ประเมิน	แบบประเมินด้านที่ 1	แบบประเมินด้านที่ 2	แบบประเมินด้านที่ 3
มัธยมศึกษาปีที่ 3	4.05	4.09	4.48
ปริญญาตรีชั้นปีที่ 1	4.13	4.04	4.44
ปริญญาตรีชั้นปีที่ 2	4.21	4.32	4.56
ปริญญาตรีชั้นปีที่ 3	4.17	4.16	4.42
ปริญญาตรีชั้นปีที่ 4	4.06	4.11	4.52
ครู และอาจารย์	4.32	4.21	4.5
ค่าเฉลี่ยคะแนน	**4.16**	**4.16**	**4.49**

บทที่ 5
สรุป และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ระบบสนับสนุนการสอนโดยจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ สามารถแนะนำการแบ่งกลุ่มของนักเรียนภายในชั้นเรียนที่เหมาะสมตามรูปแบบการเรียนรู้ของนักเรียน เพื่อเพิ่มประสิทธิผลของการศึกษาโดยปรับแต่งวิธีการสอนให้เหมาะสมกับความชอบในการเรียนรู้ของนักเรียนแต่ละคนด้วยการกำหนดรูปแบบการเรียนรู้ที่โดดเด่นของนักเรียนแต่ละคน ซึ่งผ่านการประเมินรูปแบบการเรียนรู้ Index Learning Style (ILS) เพื่อให้ผู้สอนสามารถปรับแต่งสื่อการสอน กลยุทธ์การสอน และวิธีการประเมินให้เหมาะสมกับความต้องการส่วนบุคคลของนักเรียนแต่ละคน

การวางแผนการสอนที่เหมาะสมกับนักเรียนแต่ละคนตามรูปการเรียนรู้นั้นจะเพิ่มประสิทธิภาพในการเรียนให้กับกับนักเรียนอย่างสูงสุด แต่จะเพิ่มภาระการวางแผนการสอนให้กับอาจารย์อย่างมากเพราะจะต้องวางแผนการสอนให้เท่ากับจำนวนนักเรียนแต่ละคนในชั้นเรียน ซึ่งการลดภาระของอาจารย์ในการวางแผนการสอน การทำความเข้าใจรูปแบบการเรียนรู้ของนักเรียนในชั้นเรียน การแบ่งกลุ่มให้เหมาะสมโดยที่ภายในกลุ่มมีความต้องการการรูปแบบเรียนรู้ที่คล้ายคลึงกันลดการวางแผนการสอน และภาระการสอนกับอาจารย์ผู้สอน

ท้ายที่สุดแล้วเป้าหมายคือการสร้างอัลกอริทึมในการสนับสนุนอาจารย์ผู้สอน สำหรับการแบ่งกลุ่มในเพิ่มการมีส่วนร่วมของนักเรียน ส่งเสริมทักษะทางสังคม และสนับสนุนรูปแบบการเรียนรู้ที่แตกต่างกัน การทำงานกลุ่มที่มีประสิทธิภาพยังต้องการความคาดหวังที่ชัดเจน และการสนับสนุนจากอาจารย์อย่างต่อเนื่อง เพิ่มประสิทธิภาพการเรียนให้กับนักเรียน และลดภาระของอาจารย์ในการวางแผนการสอน ความเข้าใจรูปแบบการเรียนรู้ของนักเรียนแต่ละคนในชั้นเรียน

เพื่อพัฒนาโมเดลต้นแบบในการแบ่งกลุ่มนักเรียนตามรูปแบบการเรียนรู้ โดยใช้การทดลองกับพารามิเตอร์ และชุดข้อมูลที่แตกต่างกัน เพื่อหาพารามิเตอร์ที่ให้ผลลัพธ์ของอัลกอริทึมที่เหมาะสมกับชุดข้อมูลทางการศึกษา การลดกระบวนการทำซ้ำของชุดข้อมูลกับพารามิเตอร์ในรูปแบบต่าง ๆ ช่วยลดระยะเวลาในการแบ่งกลุ่มชุดข้อมูล โดยพบว่าการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ของนักเรียนนั้น มีอัลกอริทึม และพารามิเตอร์ที่เหมาะสมคือ Euclidean Distance และ Ward Linkage ที่ให้ค่าเฉลี่ยของ SSE ที่น้อยกว่าพารามิเตอร์อื่น ๆ จากการทดลองกับชุดข้อมูลรูปแบบการเรียนรู้ของนักเรียนทั้ง 4 ชุดข้อมูล จำนวนกลุ่มที่เหมาะสมสำหรับการแนะนำในการแบ่งกลุ่มคือ 6-8 กลุ่ม ซึ่งให้

ค่าอัตราการเปลี่ยนแปลงของ SSE หรือ %Change ที่เข้าเงื่อนไข และเหมาะสม การประยุกต์ใช้อัลกอริทึม และพารามิเตอร์ที่แนะนำตามการทดลองของโมเดลต้นแบบนั้นจะช่วยในการหาพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลด้านการศึกษาได้อย่างมีประสิทธิภาพ ด้วยเหตุนี้ โมเดลต้นแบบนี้สามารถช่วยพัฒนากระบบสนับสนุนการจัดกลุ่มนักเรียนตามรูปแบบการเรียนรู้ได้อย่างเหมาะสม

5.2 ปัญหา และอุปสรรค

การปรับตัวกับลักษณะการเรียนรู้ของนักเรียน สำหรับการจัดกลุ่มตามรูปแบบการเรียนรู้จะต้องคำนึงถึงความแตกต่างของนักเรียนแต่ละคนในการเลือกกลุ่ม แต่การทำแบบนี้อาจเป็นอุปสรรคหรือทำให้ผู้สอนมีความยุ่งยากในการจัดกลุ่มให้เหมาะสมกับความต้องการของนักเรียน นอกจากนี้ ยังมีความเป็นไปได้ว่าผู้สอนอาจไม่สามารถปรับตัวเพื่อตอบสนองความต้องการของนักเรียนแต่ละคนได้อย่างเหมาะสม

การประสานงานกันระหว่างนักเรียนในกลุ่ม การจัดกลุ่มตามรูปแบบการเรียนรู้อาจทำให้เกิดความแตกต่างในระดับความรู้ และทักษะของนักเรียนในกลุ่มเดียวกัน ซึ่งอาจทำให้เกิดปัญหาในการสร้างบทเรียนที่เหมาะสม และการแบ่งปันความรู้ระหว่างนักเรียนในกลุ่ม นอกจากนี้ การสร้างความสัมพันธ์ที่ดีระหว่างนักเรียนในกลุ่มก็เป็นอีกอุปสรรคที่ต้องรับมือ โดยต้องมีการส่งเสริม และสนับสนุนให้นักเรียนเรียนรู้ร่วมกัน และแบ่งปันความรู้ในกลุ่มโดยเข้าใจถึงการเรียนรู้

การทดสอบระบบแนะนำการจัดกลุ่มตามรูปแบบการเรียนรู้เป็นระยะแรกของการทดลองกับพารามิเตอร์ที่หลากหลายที่ใช้ในการแบ่งกลุ่ม โดยมีปัญหาเกี่ยวกับการทดสอบกับชุดข้อมูลรูปแบบการเรียนรู้ที่น้อยเกินไป และระยะเวลาในการติดตามผลการนำการแบ่งกลุ่มไปใช้อย่างละเอียด อย่างไรก็ตาม การทดสอบด้วยวิธีการอื่น ๆ ยังคงให้ประสิทธิภาพที่ดี โดยสามารถแนะนำการแบ่งกลุ่มหรือจัดกลุ่มของข้อมูลได้อย่างเหมาะสมซึ่งอาจจะมีวิธีการวัดประสิทธิภาพที่เหมาะสมกับชุดข้อมูลในแต่ละรูปแบบ แต่ในกรณีที่ทดลองการแบ่งกลุ่มของชุดข้อมูลโดยใช้รูปแบบอื่น ๆ วิธีการทดสอบประสิทธิภาพอาจเปลี่ยนไปตามข้อกำหนด และยังมีปัญหาเรื่องระยะเวลาที่ใช้ในกระบวนการเหมืองข้อมูลกับการทดสอบระหว่างชุดข้อมูลกับพารามิเตอร์ต่าง ๆ เพื่อหาพารามิเตอร์ที่เหมาะสมกับการแบ่งกลุ่มของชุดข้อมูล ดังนั้นจึงต้องคำนึงถึงปัญหาดังกล่าวในการวิเคราะห์ผลการทดสอบ

5.3 ข้อเสนอแนะ

ระบบสนับสนุนผู้สอนโดยการจัดกลุ่มตามรูปแบบการเรียนรู้ เป็นแนวคิดจากความต้องการในพัฒนาด้านการศึกษาซึ่งเกี่ยวข้องกับการแนะนำจัดกลุ่ม หรือแบ่งกลุ่มภายในชั้นเรียนซึ่งเป็นงานวิจัยที่เกี่ยวข้องในด้านการศึกษา ซึ่งสามารถนำมาพัฒนา ต่อยอด ติดตาม ประเมินผล และขยายผลจากการใช้งานโมเดลของงานวิจัยนี้ กับงานวิจัยในรูปแบบอื่น ๆ ที่คล้ายคลึงกัน โดยการเรียนรู้ของนักเรียนแต่ละคนมีลักษณะ และสไตล์การเรียนรู้ที่แตกต่างกันไป ดังนั้นการจัดกลุ่มตามรูปแบบการเรียนรู้จะช่วยให้ผู้สอนสามารถปรับการสอนให้เหมาะสมกับความต้องการของนักเรียนได้มากขึ้น ในการจัดกลุ่มนักเรียน สามารถใช้เทคนิคต่างๆ เช่น การสำรวจพฤติกรรมการเรียนรู้ของนักเรียน การจัดกิจกรรมเพื่อพัฒนาทักษะ และความสามารถตามรูปแบบการเรียนรู้

การพัฒนาให้โมเดลการแบ่งกลุ่มให้รองรับกับชุดข้อมูลแบบประเมินอื่น นอกเหนือจากแบบประเมินรูปแบบการเรียนรู้ที่ได้เสนอไปของงานวิจัย และการทดสอบประสิทธิภาพของโมเดลที่เหมาะสมเป็นอีกหนึ่งปัจจัย กล่าวอีกนัยหนึ่งคือการเสนอการใช้งานโมเดลการแนะนำการแบ่งกลุ่มของงานวิจัยกับกระบวนการแบ่งกลุ่มอื่นทั้งในภาคส่วนของการศึกษา และภาคส่วนอื่น ๆ ที่นอกเหนือจากการศึกษา เพื่อที่จะสามารถจัดกลุ่มได้อย่างเหมาะสม

การสร้างระบบสนับสนุนผู้สอนโดยการจัดกลุ่มตามรูปแบบการเรียนรู้ สามารถทำได้โดยใช้เทคโนโลยีการเรียนรู้เชิงลึก (deep learning) เพื่อวิเคราะห์ข้อมูลพฤติกรรมการเรียนรู้ของนักเรียนและจัดกลุ่มตามลักษณะการเรียนรู้ของนักเรียน นอกจากนี้ ยังสามารถนำเอาเทคนิคการเรียนรู้แบบมีส่วนร่วม (collaborative learning) มาใช้ในการสร้างระบบสนับสนุนผู้สอน โดยให้นักเรียนเข้าร่วมกิจกรรมการเรียนรู้ร่วมกัน และแบ่งปันความรู้กัน เพื่อส่งเสริมการเรียนรู้

การทดสอบประสิทธิภาพของโมเดลสามารถทำได้โดยใช้วิธีการอื่น ๆ ที่ยังคงให้ประสิทธิภาพที่ดีเช่นเดียวกัน นอกจากนี้ยังสามารถแนะนำวิธีการแบ่งกลุ่มหรือจัดกลุ่มของข้อมูลได้อย่างเหมาะสม โดยใช้วิธีการวัดประสิทธิภาพที่เหมาะสมกับแต่ละรูปแบบของข้อมูล ในงานวิจัยนี้ได้มีการเสนอวิธีการวัดประสิทธิภาพการแบ่งกลุ่ม และการตรวจสอบประสิทธิภาพการแบ่งกลุ่มด้วย %Change อย่างไรก็ตาม ในกรณีที่ใช้รูปแบบการแบ่งกลุ่มของชุดข้อมูลอื่น ๆ วิธีการทดสอบประสิทธิภาพอาจจะต้องเปลี่ยนไปตามข้อกำหนดที่เหมาะสมกับแต่ละรูปแบบของข้อมูลนั้น ๆ

บรรณานุกรม

Abdelhadi, Abdelhakim. 2017. **Group technology application to investigate learning/teaching style of engineering students**. IEEE.

Aggarwal, Charu C. 2014. **Data classification: algorithms and applications**. CRC press.

Cha, Hyun Jin, Kim, Yong Se, Park, Seon Hee, Yoon, Tae Bok, Jung, Young Mo และLee, Jee-Hyong. 2006. **Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system**. Springer.

Felder, Richard M และSilverman, Linda K. 1988. Learning and teaching styles in engineering education. **Engineering education**,78(7), 674-681.

Felder, Richard M และSoloman, Barbara A. (2000). Learning styles and strategies.

Felder, Richard M และSpurlin, Joni. 2005. Applications, reliability and validity of the index of learning styles. **International journal of engineering education**,21(1), 103-112.

Han, Jiawei, Pei, Jian และKamber, Micheline. 2011. **Data mining: concepts and techniques**. Elsevier.

Jain, Anil K. 2010. Data clustering: 50 years beyond K-means. **Pattern recognition letters**,31(8), 651-666.

Klašnja-Milićević, Aleksandra, Vesin, Boban, Ivanović, Mirjana และBudimac, Zoran. 2011. E-Learning personalization based on hybrid recommendation strategy and learning style identification. **Computers & Education**,56(3), 885-899.

Lailiyah, Siti, Yulsilviana, Ekawati และAndrea, Reza. 2019. Clustering analysis of learning style on anggana high school student. **Telkomnika**,17(3).

Litzinger, Thomas A, Lee, Sang Ha, Wise, John C และFelder, Richard M. 2007. A psychometric study of the index of learning styles©. **Journal of engineering education**,96(4), 309-319.

Nafea, Shaimaa M, Siewe, François และHe, Ying. 2019. **A Novel Algorithm for Course Learning Object Recommendation Based on Student Learning Styles**. IEEE.

Pimpaporn, W และMeesad, P. 2014. A comparative efficiency of clustering using dynamic feature selection optimization of subspace clustering algorithms. **Information Technology Journal**,10(2), 43-51.

Švarcová, Eva และJelínková, Kristýna. 2016. Detection of learning styles in the focus group. **Procedia-Social and Behavioral Sciences**,217(177-182.

Thinsungnoena, Tippaya, Kaoungkub, Nuntawut, Durongdumronchaib, Pongsakorn, Kerdprasopb, Kittisak และKerdprasopb, Nittaya. 2015. The clustering validity with silhouette and sum of squared errors. **learning**,3(7.

สูงเนิน, ทิพ ยา ถิ น, สูงเนิน, มา โน ช ถิ น, ประสพ, กิตติ ศักดิ์ เกิด และประสพ, นิต ยา เกิด. 2018. การ ตรวจ สอบ ความ เหมาะสม ของ การ จัด กลุ่ม ข้อมูล อนุกรม เวลา. วารสาร วิทยาศาสตร์ และ เทคโนโลยี มหาวิทยาลัย อุบลราชธานี,20(2), 127-147.

ภาคพนวก ก

การนำเสนอผลงานวิชาการระดับนานาชาติ

Teaching support system by Clustering students according to learning styles

Narongsak Yotaman, Kitisak Osathanunkul, Paween Khoenkaw and Part Pramokchon*

Department of Digital Technology Innovation Science
Faculty of Science, Maejo University
Nongharn, Sansai, Chiang Mai, 50290, Thailand
narongsak_aoe@hotmail.com, kit_o@mju.ac.th, paween_k@gmaejo.mju.ac.th, part@gmaejo.mju.ac.th

***Abstract*— The student has different learning styles that indicate characteristics and skills of their learning. Assessment of student learning styles in the classroom will help the teacher know the cognitive skills of students. This research applies the hierarchical clustering technique for grouping students from similar learning styles in the classroom. In addition, the research investigates the parameters of hierarchical clustering that are appropriate for the learning style data and measures the efficiency of the proposed clustering method. Those parameters are the number of the cluster from 2 to 10 and two distance functions (Euclidean distance, Manhattan distance) and various linkage criteria (Average linkage, Complete linkage, Single linkage, Ward linkage, Median linkage). This research collects student learning styles into dataset using questionnaire for constructing student clusters. The experimental results show that grouping students into 7 clusters using the Euclidean distance function and the ward linkage criteria yields the highest efficiency in clustering (Error Sum of Squares, SSE = (569.457). Therefore, this method is suitable for the student learning style dataset collected in this research. The resulted clusters can help to identify the behaviors and learning skills of students that will encourage teachers to select appropriate methods and teaching tools.**

Keywords*— *Index learning styles, Student Grouping, Hierarchical Clustering

I. INTRODUCTION

In the classroom, student have different learning styles. Teachers can observe these characteristics and learning behaviors that there are similarities within the group and differences between groups. The learning style will be an indicator of the strengths or talents of that group of students in the future. Detection and analyzes groups of learning styles for students in the class to recommend the teacher know of the learning styles helping teachers to suitable plan to teach or enhance learning in the classroom effectively. Data clustering is a method of analyzing and partitioning data. The method brings data that have similar characteristics or similar to the same group (called clusters). The cluster determination is based on similarity approximation between the data by using calculating the distance measurement between vectors of data and use the criterion to combine between the two considered vectors.

This research presents an experiment to find a way to group dataset into appropriate and effective learning styles. The experiment consists of the process of measuring the similarity of data by calculating the distance of the data and finding the criteria for grouping various data clusters and evaluating the efficiency of the clustering methods. This process is to make more information within the cluster as consistent as possible and the information in each group to be most distinct. Moreover, the method can discover the number of groups suitable for the dataset. Therefore, the challenge for these grouping methods is to consider efficient partitioning techniques with the collected dataset.

The rest of paper is organized as follows. the next section will introduce about the literature review. Section 3 presents the progresses of proposed methods for clustering student learning styles, performance evaluating. Section 4 details on experimental results, and Section 5 gives the conclusions and future work.

II. RELATED WORK & LITERATURE REVIEWAS

Richard M. Felder and Barbara A. Soloman [1] propose the concept of the learning style index (called Index Learning Style, ILS) by observation the characteristics and methods that each student uses to learn, think, or solve problems. ILS defines the preference of leaning styles of student in 4 dimensions (Processing, Input, Understanding, Perception) which can be categorized into 8 styles (Active, Reflective, Visual, Verbal, Sequential, Global, Sensing, Intuitive). These 8 styles are used as features for collecting learning style dataset described in next section. Table 1 shows the detail of 4 preferences of leaning dimension and 8 related styles. The learning style of the students will indicate their strengths and possible trends of habits that may lead to obstacles in education.

Hierarchical clustering is the cluster analysis method that attempts to create the data group hierarchy. The strategy for hierarchical grouping is generally divided into two types: Agglomerative and Divisive [2].

The difference between the two methods is that the agglomerative method is a bottom-up approach: each observation (a data in dataset) starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. The divisive method is a top-down approach: all observation start in one cluster, and splits are performed recursively as one moves down the hierarchy. It is necessary to measure the differences between sets of observations. In most methods of hierarchical clustering, this is achieved by using appropriate indicators:Matrix of distance and Linkage criteria. [3]

Matrix of distance represents the distance between sets of observations. Choosing the right metric will affect the shape of the group, because some observations may be close to one another in the distance and farther in the other [4].

Linkage criteria define the distance between sets of observations (cluster to cluster) as a function of matrix comparison of distances. In general, the linkage criteria tend to apply to sets of observations A and B [5][6].

Abdelhakim Abdelhadi [7] proposed the concept of classifying students in the classroom according to the needs of the learning model by using the Learning style modeling, the

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Index Leaning Style (ILS) of the Felder-Silverman model. The model finds the learning style that students like, and then groups students into clusters according to the similarities of learning styles using clustering technology concepts. Based on clustering techniques, teachers can use appropriate teaching styles to teach students.

TABLE I. DESCRIBES THE DEFINITION PREFERENCE OF LEARNING STYLE BY THE FELDER-SILVERMAN MODEL.

Dimension 1 (Processing)	Dimension 2 (Input)	Dimension 3 (Understanding)	Dimension 4 (Perception)
Active (learn by trying things out, enjoy working in groups) **Reflective** (learn by thinking things through, prefer working alone)	**Visual** (prefer visual representations of material presented. Such as pictures & flow charts) **Verbal** (prefer written & spoken explanation)	**Sequential** (linear thinking process, learn in small in cemental steps) **Global** (holistic thinking process, learn in large leaps)	**Sensing** (concrete thinker, practical, oriented toward facts & procedures) **Intuitive** (abstract thinker, innovative & oriented toward theories)

III. PROPOSED METHOD

A. Data Selection

48 students of computer science program participated in the questionnaire to assess their score of preference of learning styles. There are four-dimensional learning styles, in which there are two opposite types with assessment points in each area 1-10 points. The reported score from the questionnaire indicates a preference of students for a specific category of leaning skill. If the score for the dimension is 1-3, students have both balanced skills in two types of that dimension and slight satisfaction for one type. If the score for the dimension is 4-6, students are moderately satisfied for one category of that dimension. Students may learn more easily in a failed environment. If the score for the dimension is 7-10, students are very satisfied with that category in the dimension. Students may have trouble learning in an environment that fails with some satisfaction.

The scores of preference of each student's learning style will be converted into an 8-element vector as in Table 2, and 48 sample students will be converted into a 48x8 matrix called the learning style dataset.

TABLE II. SAMPLE EVALUATION RESULTS FROM STUDENT LEARNING STYLES

No.	Processing		Perception		Input		Understanding	
	Active	*Reflective*	*Sensing*	*Intuitive*	*Visual*	*Verbal*	*Sequential*	*Global*
1	9	0	0	9	0	9	0	8

B. Data Clustering

First, the research process begins with grouping of student learning dataset using Agglomerative Hierarchical Clustering. Dividing the groups according to k values from 2 to 10 and adjusting the distance methods of data, such as Euclidean Distance and Manhattan Distance Several linkage criteria such as the single linkage, complete linkage, average linkage, Ward linkage, Median linkage [5][6] are involved in the experiment. Next, the proposed method will take the result clusters of the Hierarchical clustering that varies these essential parameters to find the efficiency of clustering. In order to find performance of each method of data clustering, the error sum of square (SSE) is applied to evaluate the grouping efficiency that is most suitable for the dataset

C. Clustering Validity Method

This section presents the methods of checking the efficiency of clustering by considering the error sum of different types of clustering methods. In general, Unsupervised Learning is necessary to use indicators to assess whether the group divided into the dataset is appropriate or not, using the sum of the errors as in Equation 1

$$\mathrm{SSE(c)}=\sum_{i=1}^{k}\sum_{x_j\in c_i}\left\|x_j-\mu_i\right\|^2 \quad (1)$$

When x_j = values of j
When $\mu_i = average$ of i

The research considers SSE values by graphing the relationship between SSE values and number of cluster *K* and finding the local change of slope of the SSE significantly, also known as the knee position. This position will be able to identify the number of groups suitable for the student dataset Tippaya, et.al. [8].

The proposed method selects the suitable *k* clusters that *k* clusters has low the SSE change rate calculated by Equation 2.

$$\%\mathrm{Change}=\frac{(SSE\ of\ K_{i-1}-SSE\ of\ K_i)\times 100}{SSE\ of\ K_i} \quad (2)$$

When $i \leq 2$
SSE of K_{i-1}=SSE value of K_{i-1}
SSE of K_{i-1}=SSE value of K_{i-1}

D. Selection an Appropriate Number of Cluster

The research varies the distance function and linkage criteria. This section proposed the algorithm for choosing the number of clusters that are appropriate for the dataset. The algorithm selects the best clustering result based on the Error Sum of Squares (SSE) and change rate of error sum of square (% Change) in each clustering.

- Step 1 From varying *k* value the number of clusters from 2 to *N*, consider the best linkage criteria which average of the Error Sum of Squares (SSE) is the smallest value.
- Step 2 Define the cutoff threshold of SSE % Change, *T*,
- Step 3 At the best linkage criteria, determine the SSE % Change of each *k* number of clusters and select the best *k* value from the consideration of the first *k* value that has passed both conditions: the SSE value is less than the average SSE and the change value is less than the threshold, *T*.

Then, the method get *k* value which is the number of clusters suitable for the segmentation of the dataset.

IV. EXPERIMENT

The learning style dataset has been collected and analyzed for Agglomerative Hierarchical clustering data by adjusting k from 2 to 10 with two distance functions and varying linkage criteria. Next, the resulted clusters are evaluated the Error Sum of Squares (SSE). Whole SSE values are adopted to construct a comparison table to examine the change rate of error sum of square (% Change). Next, the average of the Error Sum of Squares (SSE) in each clustering parameters will be considered to the appropriate clustering parameters, such as the number of clusters (k), distance function, and linkage criteria will be determined.

In the clustering, different linkage criteria yield different resulted clusters with different SSE values. Tables 3 and 4 show that SSE values and change rate of each the varied k and the average of SSE with Euclidean Distance and varying Linkage criteria. Fig. 1 demonstrates the graph of SSE values of the clustering with Euclidean distance and varying Linkage criteria. The graph shows the relationship between the number of clusters (k) and their SSE values. The defined cutoff threshold of SSE % Change, T is 30%. Based on the proposed method for selecting an appropriate number of clusters, the best choice parameters are 7 clusters and Ward's linkage criteria. The total change rate of error sum of square (% Change) is 26.1653% in cluster k=7, which is the segmentation method that best suits the dataset.

Likewise, Figure 2 demonstrates the graph of SSE values of the clustering with Manhattan distance and varying Linkage criteria. Therefore, this research chooses 7 clusters and Ward's linkage criteria and Euclidean distance because of the smallest average of SSE values and SSE % Change more than 30 %.

TABLE III. EXPLAIN THE CHANGE RATE OF THE ERROR SUM OF SQUARE SSE WITH THE HIERARCHICAL CLUSTERING BY ADJUSTING THE NUMBER OF THE CLUSTER FROM 2 TO 10 AND EUCLIDEAN DISTANCE METHOD AND VARIOUS LINKAGE CRITERIA

Number of Cluster	Euclidean Distance					
	Average		Complete		Single	
	SSE	*%Change*	*SSE*	*%Change*	*SSE*	*%Change*
K=2	2628.99	-	2548.23	-	3055.61	-
K=3	2286.97	14.96	2181.25	16.82	2901.9	5.30
K=4	1871.89	22.17	1552.80	40.47	2771.18	4.72
K=5	1722.89	8.65	1243.45	24.88	1819.25	52.36
K=6	1094.45	57.42	1094.45	13.61	1248.17	45.75
K=7	998.57	9.60	569.46	92.19	1106.29	12.82
K=8	473.58	110.86	473.58	20.24	728.50	51.86
K=9	397.92	19.02	386.89	22.41	632.31	15.21
K=10	324.72	22.54	311.22	24.31	559.11	13.09
Average SSE	1311.11	-	1151.26	-	1646.93	-

TABLE IV. EXPLAIN THE CHANGE RATE OF THE ERROR SUM OF SQUARE SSE WITH THE HIERARCHICAL CLUSTERING BY ADJUSTING THE NUMBER OF THE CLUSTER FROM 2 TO 10 AND EUCLIDEAN DISTANCE METHOD AND VARIOUS LINKAGE CRITERIA

Number of Cluster	Euclidean Distance			
	Ward		Median	
	SSE	*%Change*	*SSE*	*%Change*
K=2	2508.85	-	3055.61	-
K=3	1983.86	26.46	2924.74	4.47
K=4	1394.80	42.23	2771.18	5.54
K=5	1084.71	28.59	2132.75	29.93
K=6	718.46	50.98	1926.84	10.69
K=7	569.46	**26.17**	1675.11	15.03
K=8	482.76	17.96	885.63	89.14
K=9	386.89	24.78	859.13	3.08
K=10	313.69	23.34	785.93	9.31
Average SSE	1049.27	-	1890.77	-

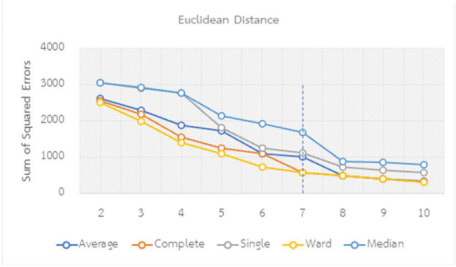


Fig. 1. The relationship between Error Sum of Square and the Cluster (k) of the Hierarchical Clustering with the Euclidean distance

TABLE V. EXPLAIN THE CHANGE RATE OF THE ERROR SUM OF SQUARE SSE WITH THE HIERARCHICAL CLUSTERING BY ADJUSTING THE NUMBER OF THE CLUSTER FROM 2 TO 10 AND EUCLIDEAN DISTANCE METHOD AND VARIOUS LINKAGE CRITERIA

Number of Cluster	Manhattan Distance					
	Average		Complete		Single	
	SSE	*%Change*	*SSE*	*%Change*	*SSE*	*%Change*
K=2	2628.99	-	2899.62	-	3055.61	-
K=3	2286.97	14.96	2032.61	42.66	2924.74	4.47
K=4	2137.97	6.97	1657.92	22.60	2771.18	5.54
K=5	1722.89	24.09	1489.67	11.29	1819.25	52.33
K=6	1094.46	57.42	1139.37	30.74	1666.31	9.18
K=7	998.57	9.60	614.38	85.45	1288.51	29.32
K=8	473.58	110.85	492.63	24.71	1193.87	7.93
K=9	397.92	19.02	423.83	16.23	632.31	88.81
K=10	324.72	22.54	348.17	21.73	371.42	70.24
Average SSE	1340.67	-	1233.13	-	1747.02	-

TABLE VI. EXPLAIN THE CHANGE RATE OF THE ERROR SUM OF SQUARE SSE WITH THE HIERARCHICAL CLUSTERING BY ADJUSTING THE NUMBER OF THE CLUSTER FROM 2 TO 10 AND MANHATTAN DISTANCE METHOD AND VARIOUS LINKAGE CRITERIA

Number of Cluster	Manhattan Distance			
	Ward		Median	
	SSE	*%Change*	*SSE*	*%Change*
K=2	2508.85	-	3055.61	-
K=3	1974.47	27.06	2053.53	48.8
K=4	1449.48	36.22	1445.82	42.03
K=5	1103.47	31.36	1331.01	8.63
K=6	737.22	49.68	966.24	37.75
K=7	588.22	25.33	870.37	11.02
K=8	502.74	17	624.88	39.29
K=9	407.16	23.48	467.28	33.73
K=10	314.91	29.29	371.42	25.81
Average SSE	1065.17	-	1242.91	-

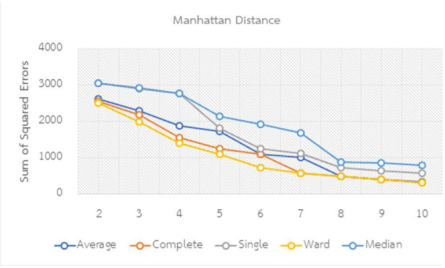


Fig. 2. The relationship between Error Sum of Square and the Cluster (k) of the Hierarchical Clustering with the Manhattan distance.

Firstly, after selecting the most suitable parameters and groups (k) of students dataset. Explanation of learning styles within a group (k) using Cluster Centroid, which is the mean of variables for observation, to explain the learning styles of each group. Table 7 shows the Cluster Centroid of all 7 groups of clusters and the 8 features of learning styles, which are used to describe the characteristics of each group's learning styles.

Secondly, consider the inside of the cluster with features of learning style using the Cluster Centroid as a leaning descriptor of the learning pattern within the cluster.

Finally, Considering all 8 features of learning styles with all 7 clusters to discover clusters that have distinctive characteristics in each type of learning and select cluster with a high Centroid Cluster is indicative cluster is similar to the type of learning style

TABLE VII. COMPARES CLUSTER CENTROID OF THE CLUSTER WITH TYPE OF LEARNING STYLE

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Active	8.0	7.32	5.25	5.5	6.5	0	2.0
Reflective	0.0	0	1.75	0.0	0.88	6.25	3.4
Sensing	0.0	6.95	8.00	5.5	0	7.38	0.0
Intuitive	8.5	0	0.00	0.0	7	0	7.6
Visual	0.0	7.89	8.75	0.0	7.75	7.63	7.0
Verbal	8.5	0	0.00	5.5	0	0	0.0
Sequential	0.0	7.11	0.00	2.5	7.25	7.75	0.0
Global	7.5	0	7.25	2.5	0	0	6.2

V. CONCLUSION

This paper presents the method for creating a dataset of learning styles for students based on considering the 4 dimensions of student preference of learning styles (separated into 8 features). This dataset is input to the process of clustering students in the class according to their learning styles. The result is clusters of students with different learning styles of students. The paper presents the experiment for finding parameters related to a cluster that enables the clustering of data to be appropriate with the learning dataset. In addition, the research evaluates and compares the efficiency of clustering methods that are appropriate to the dataset.

The results of this research will develop a decision support system that the teacher knows about student learning styles and group learning styles in the classroom. for the development of classroom learning styles and better student achievement

REFERENCES

[1] RM Felder, LK Silverman, “Learning and Teaching Styles In Engineering Education,” Engineering education, 1988

[2] Rokach, Lior, and Oded Maimon. "Clustering methods." Data mining and knowledge discovery handbook. Springer US, 2005. 321-352.

[3] Frank Nielsen (2016). "Chapter 8: Hierarchical Clustering". Introduction to HPC with MPI for Data Science. Springer.

[4] "The DISTANCE Procedure: Proximity Measures". SAS/STAT 9.2 Users Guide. SAS Institute. Retrieved 2009-04-26.

[5] "The CLUSTER Procedure: Clustering Methods". SAS/STAT 9.2 Users Guide. SAS Institute. Retrieved 2009-04-26.

[6] Székely, G. J.; Rizzo, M. L. (2005). "Hierarchical clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method". Journal of Classification. 22 (2): 151–183. doi:10.1007/s00357-005-0012-9.

[7] Abdelhakim Abdelhadi, “Group Technology Application to Investigate Learning/Teaching Style of Engineering Students,” 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 2017

[8] Tippaya Thinsungnoena, Nuntawut Kaoungkub, Pongsakorn Durongdumronchaib, Kittisak Kerdprasopb, Nittaya Kerdprasop, “The Clustering Validity with Silhouette and Sum of Squared Errors,” Proceedings of the 3rd International Conference on Industrial Application Engineering, 2015

ประวัติผู้วิจัย

ชื่อ-สกุล	ณรงค์ศักดิ์ โยตมาน
เกิดเมื่อ	04 ตุลาคม พ.ศ. 2538
ประวัติการศึกษา	พ.ศ. 2561 สำเร็จการศึกษาปริญญา วิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยแม่โจ้
ประวัติการทำงาน	พ.ศ. 2561 - 2562: Programmer บริษัท ไอเดียเคิล จำกัด พ.ศ. 2563 - 2565: เจ้าหน้าที่ศูนย์ USO Net โรงเรียนชลประทานผาแตก พ.ศ. 2565 - ปัจจุบัน: นักวิชาการคอมพิวเตอร์ปฏิบัติการ ศูนย์ข้อมูลการศึกษา กระทรวงศึกษาธิการ, ศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร สำนักงานปลัดกระทรวงศึกษาธิการ กระทรวงศึกษาธิการ