

แบบจำลองการเรียนรู้ผลการสอบภาษาอังกฤษโดยเทคนิคการจัดกลุ่ม K-means
ของนักศึกษามหาวิทยาลัยราชภัฏอุดรธานี
THE ENGLISH ACHEIVEMENT LEARNING MODEL USING K-MEANS CLUSTERING
TECHNIQUE OF UDON THANI RAJABHAT UNIVERSITY STUDENTS

นัยรัตน์ ปิ่นคำ^{1*} อรอนงค์ ทองหล่อ ทะกอง² ชลิดา เสน่ห์เมือง³ และ มงคล ทะกอง⁴

¹สาขาวิทยาการข้อมูลและเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏอุดรธานี

²สำนักวิชาศึกษาทั่วไป มหาวิทยาลัยราชภัฏอุดรธานี

³ศูนย์ภาษา สำนักวิทยบริการและเทคโนโลยีสารสนเทศ

⁴ศูนย์วิเคราะห์ข้อมูลและปัญญาประดิษฐ์ มหาวิทยาลัยราชภัฏอุดรธานี

Naiyarat Pinkum^{1*} Onaong Thonglor Thakong² Chalida Sanemueang³ and Mongkhon Thakong⁴

¹Data Science and Information Technology, Udon Thani Rajabhat University,

²School of General Education, Udon Thani Rajabhat University

³Language Center, Office of Academic Resources and Information Technology

⁴Data Analysis and Artificial Intelligence Center, Udon Thani Rajabhat University

(Received: January 9, 2023; Revised: April 15, 2023; Accepted: April 20, 2023)

*ผู้ประสานงาน: นัยรัตน์ ปิ่นคำ อีเมล: 64120604105@udru.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองการเรียนรู้การจัดกลุ่มของนักศึกษามหาวิทยาลัยราชภัฏอุดรธานีด้วยวิธีการเทคนิค K-Means ภายใต้เงื่อนไขของการนำผลสัมฤทธิ์การสอบวัดผลทักษะภาษาอังกฤษ ประกอบด้วย ทักษะด้านการอ่าน ทักษะการฟัง และทักษะการใช้ไวยากรณ์ จากโปรแกรม English Discoveries ของนักศึกษามหาวิทยาลัยราชภัฏอุดรธานี ปีการศึกษา พ.ศ. 2560 - 2562 ของทั้งหมด 7 คณะ จำนวนข้อมูล 3,596 รายการ ผลการศึกษาพบว่าแบบจำลองสามารถจัดกลุ่มนักศึกษาที่มีระดับคะแนนใกล้เคียงและจำนวนกลุ่มที่เหมาะสม เท่ากับ 8 กลุ่ม และทดสอบประสิทธิภาพของแบบจำลองการวัดคุณภาพของการจัดกลุ่มที่เหมาะสมด้วยวิธี Davies-Bouldin Index ซึ่งทำให้ผู้วิจัยเข้าใจศักยภาพทักษะภาษาอังกฤษของนักศึกษาแต่ละคณะ เพื่อนำไปสู่การจัดเตรียมโครงการที่เหมาะสมต่อการส่งเสริมการเรียนรู้พัฒนาหลักสูตรทักษะภาษาอังกฤษของนักศึกษาอย่างมีประสิทธิภาพต่อไป

คำสำคัญ: การจัดกลุ่มเคมีน, การเรียนรู้เครื่องจักร, ทักษะภาษาอังกฤษ

ABSTRACT

The objective of this research was to develop an English achievement learning model using K-Means Clustering Technique of Udon Thani Rajabhat University Students. The study data were collected from 3,596 items drawn from the English Discoveries program. It consisted of English achievement results in reading, listening, and grammar from students of 7 faculties in the academic year of 2017 to 2019.

The experimental results revealed that the learning model was able to make groups of similar students' scores and classify them into the appropriate number of groups which were 8 groups. Moreover, the effectiveness of the model was measured and the quality of appropriate clustering measuring model was examined by the Davies-Bouldin Index approach. Consequently, the researcher was able to comprehend students' English capabilities in order to design suitable courses for developing and expanding the students' competencies.

Keywords: K-Means Clustering, Machine Learning, English Skill.

1. บทนำ

การพัฒนาทักษะภาษาอังกฤษ มีทั้งนโยบาย เกณฑ์มาตรฐานวิชาชีพ จัดหลักสูตรการเรียนรู้ทุกระดับชั้น เพื่อให้ผู้เรียนได้มีทักษะภาษาอังกฤษตามเกณฑ์มาตรฐานกำหนด สำนักงานคณะกรรมการการอุดมศึกษาได้ออกประกาศ เรื่องนโยบายการยกระดับมาตรฐานภาษาอังกฤษในสถาบันอุดมศึกษาโดยระบุให้สถาบันอุดมศึกษาพิจารณาจัดให้นักศึกษาทุกคนได้สอบวัดทักษะภาษาอังกฤษ และนักศึกษาควรผ่านเกณฑ์ประเมินผลมาตรฐานวิชาชีพไม่ต่ำกว่าระดับ B1

สถาบันการศึกษา ทั้งหน่วยงานภาครัฐและเอกชนได้คิดแนวทางการพัฒนาภาษาอังกฤษด้วยการจัดโครงการอบรมเพิ่มทักษะและสมรรถนะให้แก่ผู้เรียน เพื่อส่งเสริมให้ผู้เรียนได้เรียนรู้ตามอัธยาศัย การพัฒนาสื่อการเรียนรู้รูปแบบอิเล็กทรอนิกส์ทั้งรูปแบบออนไลน์และออนไซต์ เป็นแบบชุดการเรียนรู้ หรือการทบทวนแนวข้อสอบสำหรับผู้เรียนเหมือนกันทุกคน แต่ศักยภาพและความสามารถของผู้เรียนแต่ละคนจะมีผลลัพธ์การเรียนรู้ที่แตกต่างกันไปตามคุณลักษณะของแต่ละคน จึงทำให้กระบวนการเรียนรู้ความเข้าใจของผู้เรียนไม่เท่ากัน [1] ดังนั้นการจัดกลุ่มของผู้เรียนที่มี

ความรู้ความสามารถและศักยภาพที่ใกล้เคียงกันให้อยู่กลุ่มเดียวกันจึงเป็นสิ่งสำคัญ เพื่อช่วยส่งเสริมผู้เรียนได้เรียนรู้ตรงประเด็นกับศักยภาพผู้เรียนและนำไปใช้ในการบริหารจัดการตามวัตถุประสงค์โครงการได้อย่างมีประสิทธิภาพ จากปัญหาดังกล่าวผู้วิจัยเห็นแนวทางในการวิเคราะห์พฤติกรรมตามคุณลักษณะและคุณสมบัติผู้เรียนที่มีความใกล้เคียงกัน โดยการประยุกต์ใช้เทคนิคการจัดกลุ่มเพื่อพัฒนาแบบจำลองด้วยวิธีการ เทคนิค K-means เพื่อแนะนำหลักสูตรอบรมทักษะภาษาอังกฤษประกอบไปด้วยทักษะด้านการอ่าน การฟัง และการใช้ไวยากรณ์ให้เหมาะสมและตรงตามศักยภาพผู้เรียนเป็นสำคัญ

2. แนวคิดทฤษฎีที่เกี่ยวข้อง

2.1 การเรียนรู้ของเครื่องจักร (Machine Learning)

การเรียนรู้ของเครื่องจักรที่มุ่งเน้นการสอนให้ระบบคอมพิวเตอร์เรียนรู้ข้อมูล วิเคราะห์ข้อมูล รวมถึงเรียนรู้ความสัมพันธ์ของข้อมูลด้วยตัวระบบเอง จากการรับข้อมูลและชุดคำสั่ง ผ่านกระบวนการเรียนรู้ของเครื่อง สามารถจำแนก จัดกลุ่มแยกแยะวัตถุต่าง ๆ ได้ เพิ่มประสิทธิภาพการทำงานด้วยความแม่นยำมากขึ้น และยังสามารถวิวัฒนาการเรียนรู้ด้วยระบบตนเองจากการเรียนรู้ข้อมูลที่มีพื้นฐานได้

การเรียนรู้ของเครื่องจักรนั้นจำแนกออกได้เป็น 3 รูปแบบคือ 1) การเรียนรู้โดยมีผู้สอน (Supervised Learning) เป็นการเรียนรู้ที่เครื่องจักรจะต้องอาศัยข้อมูลใน การฝึกฝน เปรียบเสมือนกับการเรียนการสอนของเด็กเล็ก โดยจำเป็นจะต้องอาศัยชุดข้อมูลต่าง ๆ ซึ่งประกอบไปด้วยชุดของข้อมูลและชุดผลลัพธ์ของข้อมูลที่ต้องการ โดยผลที่ได้จากการเรียนรู้คือ Machine Learning สามารถคาดคะเนผลลัพธ์ที่เกิดขึ้นจากการได้รับข้อมูล 2) การเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning) เป็นการเรียนรู้ที่ให้เครื่องจักรนั้นสามารถเรียนรู้ได้ด้วยตนเอง โดยไม่ต้องมีค่าเป้าหมายของแต่ละข้อมูล ซึ่งวิธีการคือมนุษย์จะเป็นผู้ใส่ข้อมูลต่าง ๆ และกำหนดสิ่งที่ต้องการจากข้อมูลเหล่านั้น โดยให้เครื่องจักรวิเคราะห์จากการจำแนกและสร้างแบบแผนจากข้อมูลที่ได้รับมา 3) การเรียนแบบเสริมกำลัง (Reinforcement Learning) เป็นการเรียนรู้สิ่งต่าง ๆ จากการลองผิดลองถูกภายใต้แนวคิดที่ว่าจะเลือกกระทำสิ่งที่ทำให้ได้ผลลัพธ์มากที่สุด โดยทำการเรียนรู้จากการลองผิดลองถูกในสถานการณ์ในอดีตหรือระบบจำลองและพยายามที่จะพัฒนาระบบการตัดสินใจของตัวเองให้ดีขึ้นเรื่อย ๆ โดยที่อาจจะพัฒนาด้วยการพยายามสร้างแบบจำลองสถานการณ์ต่าง ๆ

2.2 อัลกอริทึมจัดกลุ่ม (Clustering Algorithms)

การวิเคราะห์ข้อมูลแบบกลุ่ม การเรียนรู้ของเครื่องจักรรูปแบบ Unsupervised Learning ที่วิเคราะห์ข้อมูลโดยไม่มีการกำหนดเป้าหมาย ปัจจุบันมีเทคนิคการจัดกลุ่มหลายวิธีเช่น k-means, Fuzzy C-Means, Self-Organizing Map (SOM), Expectation-Maximization (EM), Hierarchical เป็นต้น [2-4]

2.2.1 เทคนิคการจัดกลุ่มข้อมูลเคมีน (K-Means) [5,6] คืออัลกอริทึมที่ใช้ในการจัดกลุ่มข้อมูล โดยเป็นวิธีการ clustering ที่ใช้งานได้ง่ายและได้ผลลัพธ์ที่มีประสิทธิภาพสูงในการจัดกลุ่มข้อมูลที่มีลักษณะเป็นจุดข้อมูล (data points) โดยการทำงานของ K-Means จะแบ่งข้อมูลเป็นกลุ่มๆ โดยกลุ่มที่ได้จะมีความคล้ายคลึงกันที่สุดภายในกลุ่มเอง และแตกต่างจากกลุ่มอื่น ๆ โดยอัลกอริทึม K-Means ทำงานโดยการกำหนดจำนวนกลุ่ม (K) ที่ต้องการแบ่งข้อมูล และทำการคำนวณหาจุดศูนย์กลาง (Centroid) ของแต่ละกลุ่มโดยอ้างอิงจากระยะทาง (distance) ของจุดข้อมูลจากจุดศูนย์กลาง

การคำนวณระยะทางจะคำนวณระยะทางแบบยุคลิด โดยกระบวนการจะเริ่มจากค่ากลางตั้งต้นชุดหนึ่ง จากนั้นจะจัดกลุ่มข้อมูลโดยพิจารณาจากระยะทางจากศูนย์กลาง หลังจากนั้นจะมีการคำนวณค่ากลางอีกชุดหนึ่งขึ้นมา ซึ่งจะมีการเปลี่ยนข้อมูลในกลุ่มโดยพิจารณาจากระยะทางจากศูนย์กลางค่าใหม่ หากซ้ำจนกระทั่งค่ากลางของกลุ่มไม่มีการเปลี่ยนแปลงแล้ววิธีการวิเคราะห์กลุ่มส่วนใหญ่ขึ้นอยู่กับความคล้ายคลึงกันของข้อมูลโดยคำนวณระยะทางระหว่างแต่ละคู่ การคำนวณระยะทางแบบยุคลิด สามารถคำนวณได้จากสมการที่ (1)

$$\begin{aligned}d(p, q) &= \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \\&= \sqrt{\sum_{i=1}^n (p_i - q_i)^2}\end{aligned}\quad (1)$$

ระยะทางแบบยุคลิดระหว่างจุดสองจุด p และ q = ความยาวของส่วนของเส้นตรง pq ถ้า $p = (p_1, p_2, \dots, p_n)$ และ $q = (q_1, q_2, \dots, q_n)$ n = จำนวนคุณลักษณะของข้อมูลแต่ละชุด

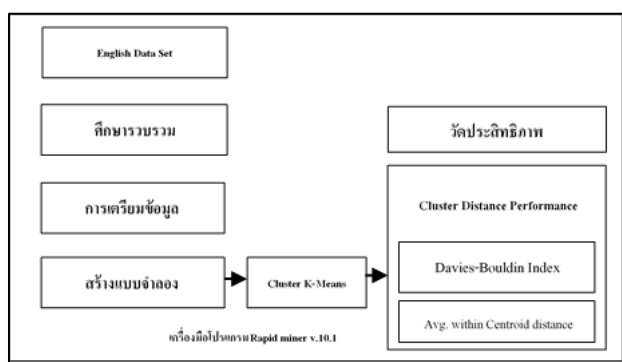
2.1.2 Davies-Bouldin index คือตัวชี้วัดเพื่อประเมินประสิทธิภาพของการจัดกลุ่ม Davies-Bouldin index คำนวณค่าเฉลี่ยของแต่ละรายการในชุดข้อมูล ค่าของแต่ละจุดจะคำนวณเป็นผลรวมหารด้วยระยะทางระหว่างจุดศูนย์กลางทั้งสองของกลุ่ม [7] ดังจากสมการที่ (2)

$$\frac{1}{K} \sum_{k=1}^K \max_{k \neq i} \left\{ \frac{S(U_k) + S(U_i)}{d(U_k, U_i)} \right\} \quad (2)$$

โดย ค่า $S(U_k) + S(U_i)$ = ระยะห่างของข้อมูลภายในกลุ่ม k และกลุ่ม i ค่า $d(U_k, U_i)$ = ระยะห่างระหว่างจุดกึ่งกลางกลุ่ม k กับกลุ่ม i ดังนั้น หากค่าของ Davies-Bouldin index มีค่าน้อยที่สุดจะทำให้ได้การแบ่งแยกของกลุ่มที่ดีที่สุด

3. วิธีดำเนินงานวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงประยุกต์ (Applied Research) ได้นำผลสัมฤทธิ์การสอบภาษาอังกฤษของ English discovery มาประยุกต์ใช้กับการเรียนรู้เครื่องจักร ซึ่งได้พัฒนาแบบจำลองการจัดกลุ่มนักศึกษาที่มีผลสัมฤทธิ์ทักษะภาษาอังกฤษ 3 ด้านคือทักษะการอ่าน ทักษะการฟังและทักษะการใช้ไวยากรณ์ โดยมีขั้นตอนการดำเนินงาน 4 ขั้นตอน ดังนี้ ขั้นตอนที่ 1 ผู้วิจัยศึกษาทำความเข้าใจและรวบรวมข้อมูล ขั้นตอนที่ 2 การเตรียมข้อมูล ขั้นตอนที่ 3 การสร้างแบบจำลอง ขั้นตอนที่ 4 การวัดประสิทธิภาพแบบจำลอง โดยแสดงภาพรวมขั้นตอนการดำเนินงานวิจัยดังรูปที่ 1



รูปที่ 1 ขั้นตอนการดำเนินงาน

3.1 ผู้วิจัยได้ทำการศึกษาและรวบรวมข้อมูล โดยนำผลสัมฤทธิ์การเรียนรู้ของนักศึกษาที่สอบวัดผลจากโปรแกรม English Discoveries ย้อนหลังของนักศึกษามหาวิทยาลัยราชภัฏอุดรธานี ปี พ.ศ.2560 - 2562 ทั้งหมด 7 คณะ จำนวน 3,596 ตัวอย่าง 5 คุณลักษณะ ประกอบด้วย Group, Branch, Cefr_R, Cefr_L และ Cefr_G ดังรายละเอียดตารางที่ 1 และตารางที่ 2 แสดงรายละเอียดความหมายของระดับคะแนน

ตารางที่ 1 คุณลักษณะของข้อมูลการสอบวัดทักษะภาษาอังกฤษ

คุณลักษณะ	ความหมาย
Group	คณะครุศาสตร์ = FE ,คณะมนุษยศาสตร์และสังคมศาสตร์ = FH , คณะเทคโนโลยี = FT, คณะวิทยาการจัดการ = FM ,คณะวิทยาศาสตร์ = FOS, บัณฑิตวิทยาลัย = GS, สำนักวิชาศึกษาทั่วไป = SGE
Branch	สาขาวิชา
Cefr_R	CEFR ทักษะการอ่าน
Cefr_L	CEFR ทักษะการฟัง
Cefr_G	CEFR ทักษะการใช้แกรมมา

ตารางที่ 2 ความหมายของระดับคะแนน

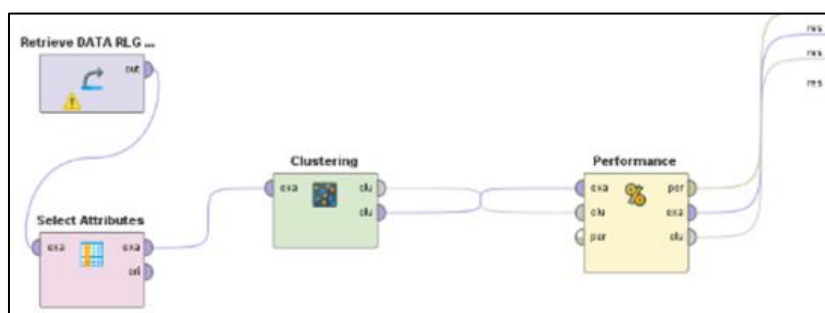
คะแนน	English Discoveries	CEFR
0	First Discoveries	A1
1	Basic 1	A1
2	Basic 2	A2
3	Basic 3	A2
4	Intermediate 1	B1
5	Intermediate 2	B1
6	Intermediate 3	B2
7	Advanced 1	B2
8	Advanced 2	C1
9	Advanced 3	C1

3.2 ผู้วิจัยทำการเตรียมข้อมูล ให้พร้อมต่อการวิเคราะห์ข้อมูล ตรวจสอบข้อมูลที่สุ่มหา และทำการแปลงข้อมูล จากข้อมูลผลสัมฤทธิ์สอบภาษาอังกฤษ จำนวน 3,596 ชุดข้อมูล ดังตัวอย่าง ข้อมูลตารางที่ 3 และผู้วิจัยได้พิจารณาคุณลักษณะเพื่อจะนำไปวิเคราะห์ร่วมกับผู้เชี่ยวชาญเลือก คุณลักษณะ 5 คุณลักษณะเพื่อนำไปวิเคราะห์ข้อมูล

ตารางที่ 3 ข้อมูลผลการสอบวัดทักษะภาษาอังกฤษ

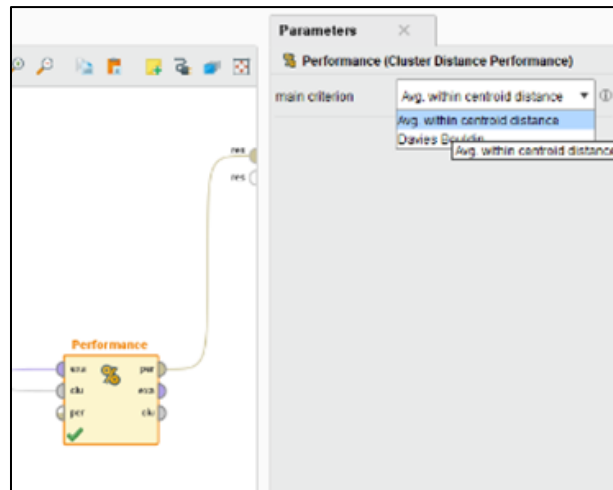
Group	Branch	Cefr_R	Cefr_L	Cefr_G
FE	THA	5	2	2
FH	ENG	1	2	1
FT	ENC	4	1	1
FM	BUC	5	4	2

3.3 การสร้างแบบจำลอง ผู้วิจัยได้ใช้เครื่องมือโปรแกรม Rapid miner v.10.1 ในการวิเคราะห์ข้อมูล และประยุกต์ใช้เทคนิคการจัดกลุ่มข้อมูลด้วยวิธีการ k-means ซึ่งผู้วิจัยกำหนดค่า $k = 3$ ถึง $k=10$ เพื่อจัดกลุ่มตามคุณลักษณะของนักศึกษาที่ใกล้เคียงกัน ภายใต้ข้อกำหนดจากทักษะภาษาอังกฤษ 3 ด้าน คือทักษะการฟัง การพูด ไวยากรณ์ ดังรูปที่ 2 การสร้างแบบจำลอง k-means



รูปที่ 2 การสร้างแบบจำลอง k-means

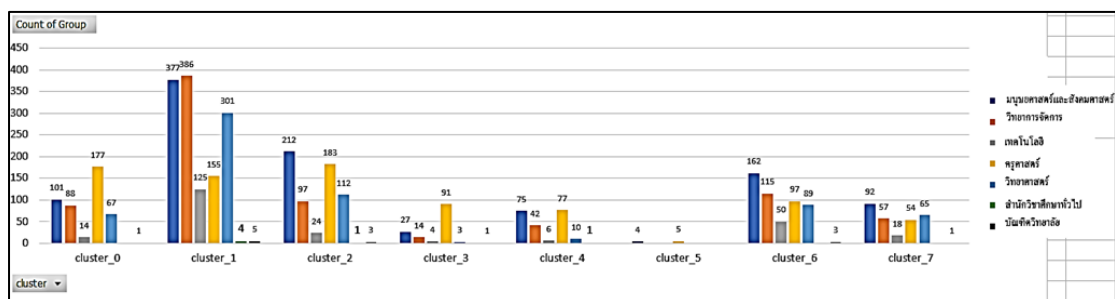
3.4 การวัดประสิทธิภาพแบบจำลองการจัดกลุ่ม [8,9] โดยใช้วิธีวัดคุณภาพการจัดกลุ่มที่ใช้ในการวิเคราะห์ เพื่อการแบ่งกลุ่มข้อมูล ผู้วิจัยได้ใช้พารามิเตอร์ด้วยวิธี Davies-Bouldin Index ซึ่งเป็นการคำนวณเปรียบเทียบ Centroid ที่ใกล้เคียงที่สุด และการวัดความคลาดเคลื่อนเฉลี่ยของ Centroid และวัดอัตราความถูกต้องพร้อมกันโดยใช้ Operators Cluster Distance Performance ในโปรแกรม Rapid miner ดังรูปที่ 3



รูปที่ 3 พารามิเตอร์ Cluster Distance Performance

4. ผลการวิจัย

ผลลัพธ์จากการทดลอง ผู้วิจัยได้วิเคราะห์การจัดกลุ่มจากคะแนนสอบภาษาอังกฤษ โดยเทคนิคการเรียนรู้เครื่องจักรด้วยวิธีเคมีนซึ่งได้ผลการจัดกลุ่มจำนวน 8 กลุ่ม จำแนกตามคณะ และจำนวนนักศึกษา Cluster 0 จำนวน 448 คน Cluster 1 จำนวน 1353 คน Cluster 2 จำนวน 632 คน Cluster 3 จำนวน 140 คน Cluster 4 จำนวน 211 คน Cluster 5 จำนวน 9 คน Cluster 6 จำนวน 516 คน Cluster 7 จำนวน 287 ดังรูปที่ 4



รูปที่ 4 ผลลัพธ์การจัดกลุ่มเคมีนจำแนกตามคณะ

ตารางที่ 4 การจัดกลุ่มคะแนนสอบภาษาอังกฤษด้วย K-means

กลุ่ม	Centriod			คำอธิบาย (เฉพาะทักษะที่มีระดับคะแนนน้อยกว่าที่กำหนด)
	Cefr_L	Cefr_R	Cefr_G	
Cluster_0	2.26	4.67	1.96	กลุ่ม L _{Basic2} G _{Basic2}
Cluster_1	1.00	1.25	1.00	กลุ่ม L _{Basic1} R _{Basic1} G _{Basic1}
Cluster_2	1.47	3.02	1.50	กลุ่ม L _{Basic2} R _{Basic3} G _{Basic2}
Cluster_3	4.65	4.20	2.61	กลุ่ม R _{Intermediate1} G _{Basic3}
Cluster_4	3.12	2.80	1.86	กลุ่ม L _{Basic3} R _{Basic3} G _{Basic2}
Cluster_5	3.33	4.33	5.00	กลุ่ม L _{Basic3} R _{Intermediate1}
Cluster_6	2.13	1.50	1.23	กลุ่ม L _{Basic2} R _{Basic2} G _{Basic1}
Cluster_7	1.01	1.54	2.03	กลุ่ม L _{Basic1} R _{Basic2} G _{Basic2}

จากตารางที่ 4 แสดงถึงผลลัพธ์การจัดกลุ่มคะแนนสอบภาษาอังกฤษด้วยวิธีการ K-means จำนวน 8 กลุ่ม ที่แสดงถึงระดับคะแนนทักษะที่มีคะแนนน้อยกว่าที่กำหนด เป็นทักษะที่ควรจะต้องได้รับการส่งเสริมและสนับสนุนการเรียนรู้มากยิ่งขึ้น ดังนี้

Cluster_0 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic2 และด้านการใช้ไวยากรณ์ เท่ากับ Basic2

Cluster_1 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic1 ด้านการอ่าน เท่ากับ Basic1 และด้านการใช้ไวยากรณ์ เท่ากับ Basic1

Cluster_2 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic2 ด้านการอ่าน เท่ากับ Basic3 และด้านการใช้ไวยากรณ์ เท่ากับ Basic2

Cluster_3 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการอ่าน เท่ากับ Intermediate1 และด้านการใช้ไวยากรณ์ เท่ากับ Basic3

Cluster_4 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic3 ด้านการอ่าน เท่ากับ Basic3 และด้านการใช้ไวยากรณ์ เท่ากับ Basic2

Cluster_5 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic3 และด้านการอ่าน เท่ากับ Intermediate1

Cluster_6 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic2 ด้านการอ่าน เท่ากับ Basic2 และด้านการใช้ไวยากรณ์ เท่ากับ Basic1

Cluster_7 คือกลุ่มของระดับคะแนนภาษาอังกฤษของนักศึกษาที่มีระดับทักษะด้านการฟัง เท่ากับ Basic1 ด้านการอ่าน เท่ากับ Basic2 และด้านการใช้ไวยากรณ์ เท่ากับ Basic2

ผู้วิจัยได้วัดประสิทธิภาพแบบจำลองด้วยวิธีการ Davies-Bouldin Index และทดลอง กำหนดค่า k ที่แตกต่างกันจนกระทั่งได้ค่า k ที่เหมาะสมคือ k เท่ากับ 8 ดังรูปที่ 4 ส่งผลกระทบต่อ คุณภาพการจัดกลุ่มการกระจายตัวของข้อมูล และได้กลุ่มที่เหมาะสม ที่มี Davies Bouldin เท่ากับ 0.854 ซึ่งเป็นค่าน้อยที่สุด ซึ่งแสดงให้เห็นว่าเป็นการกำหนดกลุ่มที่ดีที่สุด [10] ดังตารางที่ 5

ตารางที่ 5 แสดงค่า Davies Bouldin index

จำนวนคลัสเตอร์	Davies Bouldin index
3	0.996
4	0.950
5	0.982
6	0.963
7	1.113
8	0.854
9	0.959
10	0.987

5. สรุปผลและอภิปราย

บทความนี้มีวัตถุประสงค์เพื่อนำเสนอการจัดกลุ่มนักศึกษาโดยพิจารณาจากผลคะแนนการ สอบทักษะภาษาอังกฤษประกอบด้วย ทักษะการอ่าน ทักษะการเขียน และทักษะการใช้ไวยากรณ์ ของนักศึกษามหาวิทยาลัยราชภัฏอุดรธานี จำนวน 3,596 รายการ 5 คุณลักษณะ ทดสอบการจัดกลุ่มด้วย ขั้นตอนวิธี K-Means ด้วยการกำหนดจำนวนกลุ่มที่แตกต่างกันตั้งแต่ 3 กลุ่ม ถึง 10 กลุ่ม ผลการจัด กลุ่มของนักศึกษาเป็นจำนวน 8 กลุ่ม โดยวัดจากค่า Davies Bouldin index คือ 0.854 ซึ่งสอดคล้อง กับ [10,11] ที่ทดลองจัดกลุ่มและวัดประสิทธิภาพของการจัดกลุ่มด้วยค่า Davies Bouldin index ซึ่งพบว่าค่า Davies Bouldin index ต่ำสุดจะให้ประสิทธิภาพของการจัดกลุ่มดีที่สุด ผลจากการวิจัย ทำให้พบว่า ค่าเฉลี่ยผลสอบของนักศึกษาทั้งหมดอยู่ระหว่างคะแนน 1-5 เทียบเท่าการวัดระดับ มาตรฐานต่ำกว่า B1 และบ่งบอกคุณลักษณะทักษะภาษาอังกฤษทั้ง 3 ด้านของนักศึกษาเป็น รายบุคคลที่อยู่ในแต่ละกลุ่ม นำไปสู่การออกแบบการพัฒนาหลักสูตรให้เหมาะสมกับผู้เรียนตาม ศักยภาพ และเพิ่มประสิทธิภาพการบริหารจัดการโครงการพัฒนาหลักสูตรโดยเน้นผู้เรียนเป็นสำคัญ

ได้อย่างถูกต้องและมีประสิทธิภาพยิ่งขึ้น ในการวิจัยครั้งต่อไปผู้วิจัยจะทำทดสอบขั้นตอนวิธีการจัดกลุ่มแบบต่าง ๆ เพื่อหาประสิทธิภาพของขั้นตอนวิธีที่ดีที่สุด

6. กิตติกรรมประกาศ

การศึกษานี้สำเร็จสมบูรณ์ได้ด้วยความกรุณาและความช่วยเหลืออย่างดียิ่ง จากศูนย์ภาษา สำนักวิทยบริการและเทคโนโลยีสารสนเทศ และอาจารย์สาขาวิทยาการข้อมูลและเทคโนโลยีสารสนเทศทุกท่านที่ให้ข้อมูลและคำแนะนำ

เอกสารอ้างอิง

- [1] พรนภา ชุมเชื้อ. (2562). การใช้เทคนิคการจัดกลุ่มร่วมกับเทคนิคการจำแนกประเภทสำหรับการคัดกรองเด็กที่มีความบกพร่องทางการเรียนรู้จากบริบทพฤติกรรม. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิชาเทคโนโลยีสารสนเทศ, มหาวิทยาลัยทักษิณ.
- [2] ปริญญา สงวนสัตย์. (2562). **Artificial intelligence with machine learning AI สร้างได้ด้วยแมชชีนเลิร์นนิง**, นนทบุรี: โอดีซี พรีเมียร์.
- [3] สุพัตร์ จันทรโมตรี และศักดิ์ชาย ตั้งวรรณวิทย์. (2565). การวิเคราะห์ข้อมูลการขายผลิตภัณฑ์ให้แสงสว่างโดยใช้เทคนิคการจัดกลุ่ม. **The Eighteenth National Conference on Computing and Information Technology NCCIT2022 ครั้งที่ 18**, หน้า 152-157. กาญจนบุรี: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- [4] ณรรฐคุณ วิรุฬห์ศรี ลัทธพล โชครัตน์ประภา ณัฐนิชา ศรีสมาน, และพรทิพย์ เดชพิชัย. (2565). การวิเคราะห์แบ่งกลุ่มลูกค้าโดยใช้พฤติกรรมการซื้อเชิงลึก : กรณีศึกษาบริษัทผู้ผลิตอาหารสัตว์เลี้ยงแห่งหนึ่ง. **วารสารวิทยาศาสตร์ลาดกระบัง**, 31(1), 103 – 119.
- [5] Ikotun, A.M., Ezugwu, A.E., Abualigah, L., Abuhaija, B. and Heming. A. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. **Information Sciences**, 622, 178–210.
- [6] Angelo, L., Alberto, C. and Lucio, L. (2022). **k-Means Clusterization and Machine Learning Prediction of European Most Cited Scientific Publications**. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4195846.
- [7] Wijaya, Y.A., Kurniady, D.A., Wahdan Sanur Tarihoran, E.S., Rusmana, D. and Rahim, R. (2021). Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities. **TEM Journal**, 10(3), 1099-1103.

-
- [8] สุทธิพงศ์ ผ่องแผ้ว. (2561). การศึกษากระบวนการแบ่งกลุ่มเมตริกซ์ด้วยข้อมูลโครงสร้างของเมตริกซ์. สารนิพนธ์ วิทยาศาสตรมหาบัณฑิต. สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- [9] นางสาวสุชาวดี ปลั่งศรี. (2562). การจัดกลุ่มขนาดผลิตภัณฑ์โดยใช้ K-means Clustering เพื่อลดต้นทุนบรรจุภัณฑ์. วิชานิพนธ์ วิศวกรรมอุตสาหการและการจัดการบัณฑิต., วิทยาลัยมหาวิทยาลัยศิลปากร.
- [10] Baser, P. and Saini, J.R. (2014). An Optimum Cluster Size Identification for k-Means using Validity Index for Stock Market Data. **International Journal of Data Mining and Emerging Technologies**, 4(2), 107-110.
- [11] Thomas, J.C.R., Peñas, M.S. and Mora, M. (2013). New Version of Davies-Bouldin Index for Clustering Validation Based on Cylindrical Distance. In **Proceedings of 32nd International Conference of the Chilean Computer Science Society (SCCC)**, pp 49 – 53. Temuco, Chile: IEEE.