การเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีการจัดกลุ่มด้วยเทคนิคการเรียนรู้ของเครื่อง สำหรับการแบ่งกลุ่มลูกค้าธุรกิจร้านอาหาร

นายธัญวัฒน์ แก้วประดิษฐ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม
บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2568
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

การเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีการจัดกลุ่มด้วยเทคนิคการเรียนรู้ของเครื่อง สำหรับการแบ่งกลุ่มลูกค้าธุรกิจร้านอาหาร

นายธัญวัฒน์ แก้วประดิษฐ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาหลักสูตร
วิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม
บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2568
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ



ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง การเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีการจัดกลุ่มด้วยเทคนิคการเรียนรู้ของเครื่อง สำหรับการแบ่งกลุ่มลูกค้าธุรกิจร้านอาหาร

โดย นายธัญวัฒน์ แก้วประดิษฐ์

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม

คณบดีบัณฑิตวิทยาลัย

(ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ จันทร์วิพัฒน์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธาน

(ผู้ช่วยศาสตราจารย์ ดร.มธุรส ผ่านเมือง)

..... อาจารย์ที่ปรึกษา

(รองศาสตราจารย์ ดร.นลินภัสร์ บำเพ็ญเพียร)

..... อาจารย์ที่ปรึกษาร่วม

(ดร.กาญจนา วิริยะพันธ์)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ผุสดี บุญรอด)

ชื่อ : นายธัญวัฒน์ แก้วประดิษฐ์
ชื่อวิทยานิพนธ์ : การเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีการจัดกลุ่มด้วยเทคนิคการเรียนรู้ของเครื่อง สำหรับการแบ่งกลุ่มลูกค้าธุรกิจร้านอาหาร
สาขาวิชา : วิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก : รองศาสตราจารย์ ดร.นลินภัสร์ บำเพ็ญเพียร
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม : ดร.กาญจนา วิริยะพันธ์
ปีการศึกษา : 2568

บทคัดย่อ

การศึกษานี้เปรียบเทียบอัลกอริทึมการจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม (QML) กับอัลกอริทึมแบบคลาสสิก (K-Means DBSCAN GMM) ในการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน โดยใช้ข้อมูล 6,025 ราย จากกรุงเทพฯ เป็นเวลา 3 เดือน ประเมินผลด้วย Silhouette Score Davies-Bouldin Index Cluster Cohesion และเวลาประมวลผล ผลการศึกษาพบว่า DBSCAN ให้ค่าเมตริกสูงสุดแต่สร้าง 54 กลุ่มที่กระจายเกินไป QML มีข้อจำกัดทั้งประสิทธิภาพและคุณภาพต่ำ (Silhouette Score = 0.4090) ส่วน GMM เป็นวิธีที่ดีที่สุดด้วย Silhouette Score สูงสุด (0.4704) Davies-Bouldin Index ต่ำสุด (0.9014) Cluster Cohesion ดีกว่า (2.2171) และสร้าง 10 กลุ่มที่เหมาะสมต่อการใช้งานจริง มีความสมดุลที่ดีระหว่างคุณภาพการจัดกลุ่ม ประสิทธิภาพการคำนวณ และประโยชน์ทางธุรกิจ

(มีจำนวนทั้งสิ้น 77 หน้า)

คำสำคัญ : ควอนตัมแมชชีนเลิร์นนิง อัลกอริทึมการจัดกลุ่ม การแบ่งกลุ่มลูกค้า การเปรียบเทียบประสิทธิภาพ การวิเคราะห์ธุรกิจร้านอาหาร

_____ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Name : Mr. Tanyavat Kaewpradit

Thesis Title : Performance Comparison of Clustering Algorithms using Machine Learning Techniques for Segmenting Customers in Restaurant Businesses

Major Field : Data Science for Innovation
King Mongkut's University of Technology North Bangkok

Thesis Advisor : Associate Professor Dr. Nalinpat Bhumpenpein

Co-Advisor : Dr. Kanchana Viriyapant

Academic Year : 2025

ABSTRACT

This study compares quantum machine learning (QML) clustering algorithms with classical algorithms (K-Means DBSCAN GMM) for customer segmentation of nighttime restaurants using data from 6,025 customers collected over 3 months in Bangkok. Performance was evaluated using Silhouette Score Davies-Bouldin Index Cluster Cohesion and processing time. Results show that DBSCAN achieved the highest metric scores but generated 54 overly fragmented clusters unsuitable for practical use. QML demonstrated limitations in both computational performance and clustering quality (Silhouette Score = 0.4090) reflecting current quantum computing constraints. GMM emerged as the best method with the highest Silhouette Score (0.4704) lowest Davies-Bouldin Index (0.9014), superior Cluster Cohesion (2.2171), and generated 10 practical clusters, offering the best balance between clustering quality computational efficiency and business applicability.

(Total 77 Pages)

Keywords : Quantum Machine Learning Clustering Algorithms Customer Segmentation Performance Comparison Restaurant Analytics

_____ Advisor

กิตติกรรมประกาศ

การทำวิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยความช่วยเหลืออันมีค่ายิ่งจากบุคคลหลายท่าน ผู้วิจัยขอขอบพระคุณเป็นอย่างสูงต่อรองศาสตราจารย์ ดร.นลินภัสร์ บำเพ็ญเพียร อาจารย์ที่ปรึกษาหลัก และ ดร.กาญจนา วิริยะพันธ์ อาจารย์ที่ปรึกษาร่วม ที่ได้ให้คำแนะนำ คำปรึกษา และถ่ายทอดความรู้อันทรงคุณค่า ตลอดจนให้กำลังใจและสนับสนุนอย่างดียิ่ง ตลอดระยะเวลาการทำวิจัยทำให้การวิจัยครั้งนี้สำเร็จลุล่วงด้วยดี

ขอขอบพระคุณคณาจารย์ทุกท่านในสาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า ที่ได้ถ่ายทอดความรู้และประสบการณ์อันมีค่ามาตลอดหลักสูตร

ขอขอบพระคุณผู้เชี่ยวชาญทั้งสามท่าน ได้แก่ ดร.วสุพล ตรีโสภากุล คุณนฤมล สิมะขจรบุญ และคุณนิธิรุจน์ สวัสดิ์พรปิติ ที่กรุณาสะสางเวลาตรวจสอบ และให้ข้อเสนอแนะอันมีค่า สำหรับแบบสอบถามที่ใช้ในการวิจัยครั้งนี้ ทำให้เครื่องมือวิจัย มีคุณภาพและความเที่ยงตรงยิ่งขึ้น

ขอขอบพระคุณเจ้าของร้านอาหารกลางคืนที่ให้ความอนุเคราะห์สถานที่ ในการเก็บรวบรวมข้อมูล และลูกค้าทุกท่านที่กรุณาสละเวลาในการตอบแบบสอบถาม ซึ่งเป็นส่วนสำคัญที่ทำให้การวิจัยครั้งนี้สำเร็จลุล่วงได้

สุดท้ายนี้ ขอขอบคุณครอบครัวและเพื่อนๆ ทุกคนที่คอยให้กำลังใจ คำปรึกษา และความเข้าใจตลอดระยะเวลาการศึกษา ความสำเร็จของวิทยานิพนธ์ฉบับนี้ ผู้วิจัย ขอมอบเป็นเกียรติแก่บิดา มารดา และผู้มีพระคุณทุกท่านที่สนับสนุนและส่งเสริมการศึกษา เสมอมา

หากวิทยานิพนธ์ฉบับนี้มีข้อผิดพลาดประการใด ผู้วิจัยขอน้อมรับผิดชอบ แต่เพียงผู้เดียว และหวังเป็นอย่างยิ่งว่าวิทยานิพนธ์ฉบับนี้จะเป็นประโยชน์ต่อผู้ที่สนใจศึกษา ค้นคว้าในอนาคต

ธัญวัฒน์ แก้วประดิษฐ์

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฌ
สารบัญรูปภาพ ..... ญ
บทที่ 1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์การวิจัย 2
1.3 ขอบเขตของการวิจัย 2
1.4 ประโยชน์ที่ได้รับจากงานวิจัย 3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4
2.1 การแบ่งกลุ่มลูกค้าในการตลาด (Customer Segmentation in Marketing) 4
2.2 วิธีการจัดกลุ่มข้อมูลแบบดั้งเดิม (Traditional Clustering Methods) 5
2.3 หลักการของการเรียนรู้ของเครื่องแบบควอนตัม (Fundamentals of Quantum Machine Learning) 6
2.4 อัลกอริทึมการจัดกลุ่มข้อมูลด้วยควอนตัม (Quantum Clustering Algorithms) 7
2.5 การศึกษาเปรียบเทียบผลลัพธ์ต่างๆ (Performance Comparison) 8
2.6 การวิเคราะห์ลูกค้าในอุตสาหกรรมร้านอาหาร (Restaurant Industry Customer Analytics) 9
บทที่ 3 วิธีการดำเนินงานวิจัย 11
3.1 โมเดลการวิจัย 11
3.2 ขั้นตอนการดำเนินงานวิจัย 12
3.3 ประชากรและกลุ่มตัวอย่าง 13
3.4 เครื่องมือที่ใช้ในการวิจัย 14
3.5 การประยุกต์ใช้เทคนิคการแบ่งกลุ่ม 15
3.6 สถิติที่ใช้ในการวิเคราะห์ข้อมูล 19

สารบัญ (ต่อ)

บทที่ 4 ผลการดำเนินงานวิจัย 20
4.1 การวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง 21
4.2 การวิเคราะห์พฤติกรรมการใช้บริการของกลุ่มตัวอย่าง 23
4.3 การวิเคราะห์ประสิทธิภาพด้าน Silhouette Score 26
4.4 การวิเคราะห์ประสิทธิภาพด้าน Davies-Bouldin Index 28
4.5 การวิเคราะห์ประสิทธิภาพด้าน Cluster Cohesion 31
4.6 การวิเคราะห์ประสิทธิภาพด้าน Execution Time 34
4.7 การเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี 36
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ 47
5.1 สรุปผลการวิจัย 47
5.2 อภิปรายผลการวิจัย 50
5.3 การบูรณาการทฤษฎีกับกรอบการวิจัย 51
5.4 ข้อเสนอแนะการวิจัย 52
บรรณานุกรม 54
ภาคผนวก ก แบบสอบถามการวิจัย 60
ภาคผนวก ข แบบฟอร์มที่ให้ผู้เชี่ยวชาญแสดงความคิดเห็น 63
ประวัติผู้เขียน 77

สารบัญตาราง

ตารางที่ หน้า

4-1 จำนวนและร้อยละของข้อมูลทั่วไปที่เกี่ยวกับลักษณะกลุ่มตัวอย่าง.....21
4-2 จำนวนและร้อยละของพฤติกรรมการใช้บริการของกลุ่มตัวอย่าง.....23
4-3 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Silhouette Score ทุกกลุ่มตัวอย่าง.....26
4-4 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Silhouette Score ตามขนาดกลุ่มตัวอย่าง.....27
4-5 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Davies-Bouldin Index ทุกกลุ่มตัวอย่าง.....29
4-6 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Davies-Bouldin Index ตามขนาดกลุ่มตัวอย่าง.....29
4-7 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Cluster Cohesion ทุกกลุ่มตัวอย่าง.....31
4-8 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Cluster Cohesion ตามขนาดกลุ่มตัวอย่าง.....32
4-9 ประสิทธิภาพด้านเวลาของอัลกอริทึมการจัดกลุ่ม ทุกกลุ่มตัวอย่าง.....34
4-10 ประสิทธิภาพด้านเวลาของอัลกอริทึมการจัดกลุ่ม ตามขนาดกลุ่มตัวอย่าง.....35
4-11 ผลการเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี.....37

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

สารบัญรูปภาพ

ภาพที่ หน้า

2-1 QML Framework Matrix.....7
3-1 โมเดลแนวคิดการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน.....12

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในยุคดิจิทัล ข้อมูลเชิงลึกเกี่ยวกับลูกค้าที่ขับเคลื่อนด้วยข้อมูลได้กลายเป็นสิ่งสำคัญต่อความสำเร็จทางธุรกิจ โดยเฉพาะอย่างยิ่งในอุตสาหกรรมร้านอาหาร [1] ร้านอาหารที่เปิดให้บริการในช่วงกลางคืนต้องเผชิญกับความท้าทายที่เป็นเอกลักษณ์ในการทำความเข้าใจพฤติกรรมลูกค้าที่หลากหลาย โดยสร้างรูปแบบที่ซับซ้อนซึ่งวิธีการวิเคราะห์แบบดั้งเดิมไม่สามารถจับภาพได้อย่างมีประสิทธิภาพ [2] การแบ่งกลุ่มลูกค้า [3] ช่วยให้ธุรกิจสามารถจัดกลุ่มลูกค้าตามลักษณะและพฤติกรรมที่คล้ายคลึงกัน [4] อย่างไรก็ตาม เทคนิคการจัดกลุ่มแบบคลาสสิก เช่น K-means, DBSCAN และ Gaussian Mixture Models เผชิญกับข้อจำกัดที่สำคัญเมื่อต้องจัดการกับชุดข้อมูลขนาดใหญ่และซับซ้อนจากแพลตฟอร์มออนไลน์และธุรกรรมบนโซเชียลมีเดีย [5]

การเรียนรู้ของเครื่องเชิงควอนตัม (Quantum Machine Learning, QML) แสดงให้เห็นถึงขอบเขตที่มีแนวโน้มดีในการวิเคราะห์ข้อมูล โดยใช้ประโยชน์จากหลักการควอนตัม เช่น การทับซ้อน (superposition) และการพัวพัน (entanglement) การประยุกต์ใช้ QML สามารถแบ่งออกเป็นสี่ประเภท ได้แก่ CC (Classical Data-Classical Computing), CQ (Classical Data-Quantum Computing), QC (Quantum Data-Classical Computing) และ QQ (Quantum Data-Quantum Computing) [6] งานวิจัยนี้มุ่งเน้นไปที่แนวทาง CQ ซึ่งประมวลผลข้อมูลลูกค้าแบบคลาสสิกโดยใช้อัลกอริทึมควอนตัม ซึ่งเป็นหมวดหมู่ที่มีการวิจัยอย่างแข็งขันมากที่สุด พร้อมการประยุกต์ใช้ทางธุรกิจในทางปฏิบัติ

การจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม นำเสนอข้อได้เปรียบเชิงทฤษฎี รวมถึงความขนานแบบควอนตัมและการค้นพบความสัมพันธ์ที่ซับซ้อนซึ่งวิธีการแบบคลาสสิกไม่สามารถรับรู้ได้ อย่างไรก็ตาม การตรวจสอบเชิงประจักษ์ของอัลกอริทึมการจัดกลุ่มการเรียนรู้ของเครื่องเชิงควอนตัม ยังคงมีอยู่อย่างจำกัด คำถามสำคัญยังคงมีอยู่เกี่ยวกับข้อได้เปรียบด้านประสิทธิภาพที่แท้จริงของอัลกอริทึมเหล่านี้เมื่อเทียบกับวิธีการแบบคลาสสิกในการประยุกต์ใช้งานจริง โดยพิจารณาจากภาระการคำนวณ ความสามารถในการตีความ และข้อจำกัดด้านความสามารถในการขยายขนาด [7]

งานวิจัยนี้ได้ทำการวิเคราะห์เปรียบเทียบอย่างครอบคลุมระหว่างการจัดกลุ่มด้วย การเรียนรู้ของเครื่องเชิงควอนตัม กับวิธีการแบบดั้งเดิมสามวิธี ได้แก่ K-Means, DBSCAN และ Gaussian Mixture Model โดยใช้ข้อมูลลูกค้าจริงจากร้านอาหารที่เปิดให้บริการในช่วงกลางคืน [8]

การศึกษานี้ใช้เมตริกซ์การประเมินมาตรฐาน (Silhouette Score, Davies-Bouldin Index และ Cluster Cohesion) [9] ผลการวิจัยจะระบุแนวทางการจัดกลุ่มที่มีประสิทธิภาพมากที่สุดสำหรับการพัฒนากลยุทธ์การตลาดเฉพาะกลุ่มและการปรับปรุงการดำเนินงาน นอกจากนี้ งานวิจัยนี้ยังมีส่วนช่วยให้หลักฐานเชิงประจักษ์เกี่ยวกับการนำเทคโนโลยีควอนตัมมาใช้ในการวิเคราะห์เชิงพาณิชย์ ซึ่งช่วยเพิ่มพัฒนาความเข้าใจของเราเกี่ยวกับการประยุกต์ใช้ระบบคอมพิวเตอร์ควอนตัมในด้านธุรกิจอัจฉริยะ

1.2 วัตถุประสงค์การวิจัย

เพื่อเปรียบเทียบประสิทธิภาพการแบ่งกลุ่มลูกค้า ระหว่างเทคนิค การจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม (Quantum Machine Learning Clustering) กับเทคนิคการแบ่งกลุ่มแบบดั้งเดิม (K-Means, DBSCAN และ Gaussian Mixture Model) โดยใช้เกณฑ์การประเมินมาตรฐาน (Silhouette Score, Davies-Bouldin Index และ Cluster Cohesion) ในธุรกิจร้านอาหารกลางคืน

1.3 ขอบเขตของการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของการแบ่งกลุ่มลูกค้าระหว่างการจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม และการแบ่งกลุ่มแบบดั้งเดิม สำหรับธุรกิจร้านอาหารกลางคืน โดยมีขอบเขตการวิจัยดังนี้

1.3.1 ขอบเขตด้านเนื้อหา

ในงานวิจัยนี้ ประการแรกจะทำการศึกษา และประยุกต์ใช้การจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม สำหรับการแบ่งกลุ่มลูกค้า ซึ่งเป็นเทคนิคที่นำหลักการของ การคำนวณเชิงควอนตัม (Quantum Computing) มาประยุกต์ใช้กับการวิเคราะห์ข้อมูล ประการที่สองจะทำการเปรียบเทียบประสิทธิภาพกับเทคนิคการแบ่งกลุ่มแบบดั้งเดิม 3 วิธีได้แก่ K-Means, DBSCAN และ Gaussian Mixture Model โดยใช้เกณฑ์การประเมินมาตราฐาน ได้แก่ Silhouette Score, Davies-Bouldin Index และ Cluster Cohesion ประการที่สามจะทำการวิเคราะห์ลักษณะเฉพาะของลูกค้าในแต่ละกลุ่มที่ได้จากการแบ่งกลุ่ม และประการสุดท้ายจะพัฒนาข้อเสนอแนะเชิงกลยุทธ์การตลาดเฉพาะกลุ่มเพื่อให้ธุรกิจสามารถนำไปปรับใช้ได้จริง

1.3.2 ขอบเขตด้านพื้นที่

งานวิจัยนี้ทำการเก็บรวบรวมข้อมูลจากธุรกิจร้านอาหารกลางคืน 1 ร้านในเขตกรุงเทพมหานคร โดยเลือกร้านอาหารที่มีลักษณะการดำเนินงานในช่วงเวลากลางคืน มีการให้บริการที่เปิดในช่วงเวลา 19.00-02.00 น. และมีฐานลูกค้าที่หลากหลาย

1.3.3 ขอบเขตด้านระยะเวลา

ระยะเวลาในการเก็บข้อมูล 3 เดือน ตั้งแต่เดือนกันยายน ถึงเดือนพฤศจิกายน พ.ศ. 2568 โดยจะทำการเก็บข้อมูลพฤติกรรมการใช้บริการของลูกค้าอย่างต่อเนื่องตลอดช่วงเวลาดังกล่าว

1.3.4 ขอบเขตด้านกลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัยนี้ประกอบด้วยลูกค้าจำนวน 6,025 คน ที่มาใช้บริการร้านอาหารที่เปิดในช่วงเวลา 19.00-02.00 น.

1.4 ประโยชน์ที่ได้รับจากงานวิจัย

ช่วยให้ธุรกิจร้านอาหารกลางคืนเข้าใจลูกค้าได้ลึกซึ้งมากขึ้น และสนับสนุนการพัฒนากลยุทธ์การตลาดที่เฉพาะเจาะจงและมีประสิทธิภาพ ซึ่งจะนำไปสู่การเพิ่มความได้เปรียบทางการแข่งขันในธุรกิจบริการ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การแบ่งกลุ่มลูกค้าเป็นกลยุทธ์สำคัญในการตลาดสมัยใหม่ที่องค์กรต่าง ๆ นำมาใช้เพื่อเข้าใจและตอบสนองความต้องการของลูกค้าได้อย่างมีประสิทธิภาพมากยิ่งขึ้น การศึกษาวิจัยในช่วงทศวรรษที่ผ่านมาแสดงให้เห็นถึงการพัฒนาเทคนิคการแบ่งกลุ่มลูกค้าที่หลากหลาย โดยเฉพาะการประยุกต์ใช้เทคโนโลยีการเรียนรู้ของเครื่อง (Machine Learning) และเทคนิคการทำเหมืองข้อมูล (Data Mining) เพื่อให้ได้ผลลัพธ์ที่แม่นยำและเป็นประโยชน์มากขึ้น ผู้วิจัยได้แบ่งการศึกษางานวิจัยออกเป็น 6 หัวข้อหลักดังนี้

2.1 การแบ่งกลุ่มลูกค้าในการตลาด (Customer Segmentation and Marketing)

2.2 วิธีการจัดกลุ่มข้อมูลแบบดั้งเดิม (Traditional Clustering Methods)

2.3 หลักการของการเรียนรู้ของเครื่องแบบควอนตัม (Fundamentals of Quantum Machine Learning)

2.4 อัลกอริทึมการจัดกลุ่มข้อมูลด้วยควอนตัม (Quantum Clustering Algorithms)

2.5 การศึกษาเปรียบเทียบผลลัพธ์ต่างๆ (Performance Comparison)

2.6 การวิเคราะห์ลูกค้าในอุตสาหกรรมร้านอาหาร (Restaurant Industry Customer Analytics)

2.1 การแบ่งกลุ่มลูกค้าในการตลาด (Customer Segmentation in Marketing)

แบบจำลอง RFM ซึ่งย่อมาจาก ความใหม่ (Recency), ความถี่ (Frequency) และ มูลค่า (Monetary) ยังคงเป็นวิธีที่นิยมใช้กันอย่างแพร่หลาย โดยมีการนำไปประยุกต์ใช้ในยุคปัจจุบันที่ให้ผลลัพธ์ที่น่าประทับใจ คือสามารถเพิ่มอัตราการเปลี่ยนใจของลูกค้าจาก 5% ในช่วงแรกเป็น 41% ภายในหกเดือน [10] [11] จากการศึกษาพบว่าการใช้โมเดล RFM ร่วมกับ K-means สามารถแบ่งข้อมูลออกเป็น 10 กลุ่มที่แตกต่างกัน ได้แก่ กลุ่มที่มีศักยภาพและภักดี, กลุ่มเสี่ยง, กลุ่มที่ห้ามพลาด, ลูกค้าภักดี, ลูกค้าใหม่, กลุ่มที่กำลังจะหลับ, กลุ่มที่พักตัว, กลุ่มที่มีแนวโน้มดี, กลุ่มที่ต้องดูแล และกลุ่มแชมป์เปี้ยน โดยการประเมินประสิทธิภาพด้วย Silhouette index พบว่า RFM ให้คะแนนที่ 0.637 ซึ่งดีกว่า K-means แบบธรรมดาที่ได้เพียง 0.559 แสดงให้เห็นว่าโมเดล RFM ช่วยปรับปรุงการแบ่งกลุ่มลูกค้าให้มีความชัดเจนและแยกจากกันมากขึ้นอย่างมีประสิทธิภาพ [11] การผสานรวม Customer Lifetime Value หรือมูลค่าตลอดชีวิตของลูกค้ากับการแบ่งกลุ่มช่วยให้สามารถสร้างภาพลักษณ์ของลูกค้าได้อย่างมีประสิทธิภาพผ่าน Decision Tree Models

โดยโมเดลที่นำเสนอมีประสิทธิภาพค่อนข้างแข็งแกร่งด้วยค่า R-squared ที่ 0.25 และ RMSE ที่ 956.04 ซึ่งการวิเคราะห์ระบุคุณลักษณะสำคัญ เช่น ยอดคงเหลือในบัญชีจำนอง ยอดคงเหลือบัตรเครดิต ยอดคงเหลือบัญชีออมทรัพย์ และจำนวนบัญชีที่เปิดใช้งาน เป็นตัวแปรสำคัญในการแบ่งแยกลูกค้า [12]

2.2 วิธีการจัดกลุ่มข้อมูลแบบดั้งเดิม (Traditional Clustering Methods)

K-means เป็นวิธีที่ได้รับความนิยมสูงสุดเนื่องจากมีความเรียบง่ายและให้ผลลัพธ์ที่รวดเร็ว สามารถแบ่งลูกค้าออกเป็นกลุ่มต่างๆ เช่น ลูกค้ามูลค่าสูงและลูกค้าที่มีความภักดี ถึงแม้จะมีข้อจำกัดคือจำนวนกลุ่มต้องกำหนดล่วงหน้าและมีความไวต่อการเลือกจุดเริ่มต้นของจุดศูนย์กลางกลุ่ม แต่ปัญหานี้ก็สามารถแก้ไขได้ด้วยเทคนิค K-means++ และวิธี Elbow Method มีการศึกษาที่ชี้ให้เห็นว่าการนำ K-means มาใช้ร่วมกับ อัลกอริทึมแอฟฟินิตีพรอพาเกชัน (Affinity Propagation) ให้ผลลัพธ์ที่ดีเยี่ยมในการแบ่งกลุ่มลูกค้าเพื่อวัตถุประสงค์ทางการตลาด โดยใช้เครื่องมือวิเคราะห์อย่าง Silhouette score ช่วยในการกำหนดจำนวนกลุ่มที่เหมาะสม การวิจัยได้เปรียบเทียบวิธีการจัดกลุ่มต่างๆ เช่น K-Means และ DBSCAN โดยกล่าวถึงประสิทธิภาพและความแม่นยำของ K-Means พร้อมทั้งแนะนำให้มีการวิจัยเพิ่มเติมในการใช้อัลกอริทึมอื่นๆ เพื่อเอาชนะความท้าทายที่เกิดจากชุดข้อมูลขนาดใหญ่ การประยุกต์ใช้ RFM ร่วมกับ K-Means ช่วยปรับปรุงการแบ่งกลุ่มลูกค้าในระบบ CRM ซึ่งการแบ่งกลุ่มลูกค้ามีความสำคัญต่อการโฆษณาแบบกำหนดเป้าหมายและการปรับปรุงความภักดีของลูกค้า อัลกอริทึมการจัดกลุ่มที่ได้รับการปรับปรุงช่วยอำนวยความสะดวกในการแบ่งกลุ่มลูกค้าโทรคมนาคมที่ดีขึ้นโดยหลีกเลี่ยงข้อผิดพลาดของ K-Means และ DBSCAN และสนับสนุนการตลาดที่มุ่งเน้นและการเพิ่มประสิทธิภาพทรัพยากร [13]

DBSCAN นั้นมีความโดดเด่นในการจัดการข้อมูลที่มีรูปทรงไม่ปกติและสามารถตรวจจับค่าผิดปกติได้ดี ต้านทานต่อสัญญาณรบกวนและสามารถค้นหากลุ่มที่มีรูปร่างต่างๆ ได้ ให้ค่า DB Index (Davies-Bouldin Index) ที่ต่ำที่สุดและมีความเหนียวแน่นของกลุ่มสูงสุด แม้ว่าจะมีความยากในการตั้งค่าพารามิเตอร์ต่างๆ [14] Fuzzy C-Means ให้ผลลัพธ์ที่ดีมากในแง่ของ Silhouette Score สูงสุดโดยมีกลุ่มที่กำหนดไว้อย่างดี แสดงให้เห็นว่าจุดข้อมูลภายในกลุ่มเดียวกันมีความคล้ายคลึงกันเมื่อเทียบกับกลุ่มอื่น อนุญาตให้จุดข้อมูลสามารถเป็นสมาชิกของหลายกลุ่มได้ อย่างไรก็ตาม ผลลัพธ์สำหรับ Fuzzy Clustering ไม่ดีนักเกี่ยวกับการแยกกลุ่ม ซึ่งสะท้อนให้เห็นจากค่า Davies-Bouldin Index ที่โดดเด่นและความเหนียวแน่น [15] ในขณะที่การจัดกลุ่มแบบไดนามิกสำหรับข้อมูลที่เป็นช่วงให้ผลลัพธ์ที่ดีกว่าวิธี K-means ทั่วไปในแง่ของความแม่นยำ [16]

Gaussian Mixture Models หรือ GMM นั้นแสดงให้เห็นประสิทธิภาพที่ดีด้วยค่า Silhouette score ที่ 0.435 และ Davies-Bouldin index ที่ 0.7 ซึ่งบ่งบอกว่า GMM

มีความแม่นยำโดยรวมที่สูงและสามารถจัดการกับข้อมูลที่มีโครงสร้างซับซ้อนได้ดี พร้อมทั้งสามารถคำนวณค่าความน่าจะเป็นได้อย่างมีประสิทธิภาพ [15] Agglomerative clustering ก็แสดงให้เห็นถึงความแม่นยำที่สูงเช่นกัน โดยค่าผิดปกติต่างๆ สามารถมองเห็นได้ในผลการจัดกลุ่ม ซึ่งไม่ได้ถูกรวมเข้าไปในกลุ่มใดๆ โดยวิธีนี้จะสร้างกลุ่มขึ้นมาแบบค่อยเป็นค่อยไปจากล่างขึ้นบน

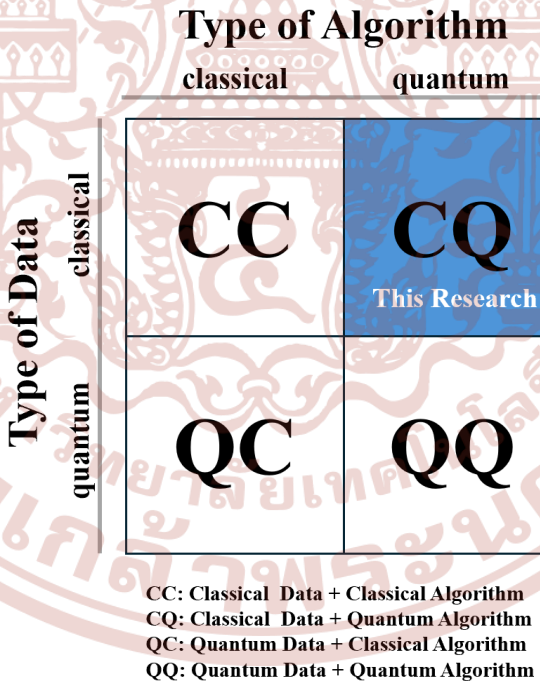
2.3 หลักการของการเรียนรู้ของเครื่องแบบควอนตัม (Fundamentals of Quantum Machine Learning)

การเรียนรู้ของเครื่องเชิงควอนตัม เป็นการผสมผสานระหว่างคอมพิวเตอร์ควอนตัมกับการเรียนรู้ของเครื่อง โดยอาศัยหลักการของการทับซ้อน (superposition) และการพัวพัน (entanglement) ซึ่งเป็นหลักการพื้นฐานของกลศาสตร์ควอนตัมที่เพิ่มความสามารถในการขยายตัวทางการคำนวณ ทำให้เกิดพลังการคำนวณที่เพิ่มขึ้นแบบทวีคูณ ซึ่งแตกต่างจากคอมพิวเตอร์แบบคลาสสิกที่ใช้ 0 และ 1 เพื่อกำหนดบิตแบบไบนารีคิวบิต สามารถมีหลายสถานะพร้อมกันได้ ทำให้พลังการคำนวณถูกขยายแบบทวีคูณ [17, 18] เกตควอนตัม (Quantum gates) และอัลกอริทึมทำให้คอมพิวเตอร์ควอนตัมสามารถแก้ปัญหาที่ซับซ้อนได้หลายประเภท เช่น การเพิ่มประสิทธิภาพ การเข้ารหัส หรือการจำลองโครงสร้างโมเลกุล อัลกอริทึมควอนตัมเป็นหัวใจสำคัญของการเรียนรู้ของเครื่องเชิงควอนตัม เนื่องจากใช้ประโยชน์จากการคำนวณเชิงควอนตัมและสร้างกระบวนการคำนวณใหม่ๆ

อัลกอริทึมควอนตัมที่มีชื่อเสียงได้แก่ Grover's Algorithm และ Shor's Algorithm โดย Grover's Algorithm ใช้สำหรับแก้ปัญหาการค้นหาที่ไม่มีโครงสร้าง จำนวนรอบการวนซ้ำที่ต้องใช้ในการค้นหารายการที่ทำเครื่องหมายในฐานข้อมูลที่ไม่เรียงลำดับ ส่วน Shor's Algorithm นั้น เป็นอัลกอริทึมควอนตัมที่มีชื่อเสียงอีกตัวหนึ่งที่สามารถแยกตัวประกอบจำนวนเต็มขนาดใหญ่ได้เร็วกว่าอัลกอริทึมคลาสสิกที่รู้จักกันดีที่สุดแบบทวีคูณ Shor's Algorithm สามารถแยกตัวประกอบจำนวนเต็ม N-bit ใน $O((\log N)^3)$ เวลาบนคอมพิวเตอร์ควอนตัม ในขณะที่อัลกอริทึมคลาสสิกที่รู้จักกันดีที่สุดจะใช้เวลามากกว่ามาก อัลกอริทึมเหล่านี้ร่วมกับ การแปลงฟูเรียร์เชิงควอนตัม (Quantum Fourier Transform), อัลกอริทึมการหาค่าเหมาะสมแบบประมาณเชิงควอนตัม (Quantum Approximation Optimization Algorithm), ตัวแก้ค่าเจาะจงเชิงควอนตัมแบบแปรผัน (Variational Quantum Eigen Solver) และ การประมาณเฟสเชิงควอนตัม (Quantum Phase Estimation) ทำหน้าที่เป็นรากฐานสำหรับแนวทางการเรียนรู้ของเครื่องเชิงควอนตัม หลายอย่างเปิดเส้นทางใหม่สำหรับการเรียนรู้ของเครื่อง และการวิเคราะห์ข้อมูล [18]

การเรียนรู้ของเครื่องเชิงควอนตัมสามารถแบ่งออกเป็นสี่ประเภท ได้แก่ CC คือการใช้ข้อมูลและการคำนวณแบบคลาสสิก, CQ คือการนำข้อมูลคลาสสิกมาประมวลผลด้วยคอมพิวเตอร์ควอนตัม

ซึ่งเป็นประเภทที่มีการวิจัยมากที่สุดโดยใช้การแมปคุณลักษณะเชิงควอนตัม, QC คือการใช้ข้อมูลควอนตัมมาประมวลผลด้วยคอมพิวเตอร์คลาสสิก และ QQ คือการใช้ทั้งข้อมูลและการคำนวณแบบควอนตัม [6, 19, 20] อัลกอริทึมการเรียนรู้ของเครื่องเชิงควอนตัม ให้ความสามารถที่ปรับปรุงแล้วสำหรับปัญหาการจำแนก การถดถอย และการจัดกลุ่มไปจนถึง การจัดกลุ่มแบบ K-Means เชิงควอนตัม (Quantum K-Means Clustering) อัลกอริทึมเหล่านี้มีศักยภาพในการแก้ปัญหาในโลกจริงด้วยความแม่นยำและประสิทธิภาพที่สูงขึ้นมาก นอกจากนี้การเรียนรู้ของเครื่องเชิงควอนตัมยังมีการประยุกต์ใช้ในหลากหลายสาขา รวมถึงการเข้ารหัส การค้นพบยา การเงิน เป็นต้น อัลกอริทึมกลศาสตร์ควอนตัมเร่งความเร็วในการค้นหาปัญหาการเพิ่มประสิทธิภาพที่ยากและวิธีเคมีควอนตัมให้การเข้าถึงการจำลองโครงสร้างโมเลกุลและพารามิเตอร์ควอนตัมที่แม่นยำอย่างน่าทึ่ง ช่วยให้สามารถสร้างยาและวัสดุใหม่ๆ ตัวอย่างเช่น ระบบการเข้ารหัสแบบควอนตัมให้เส้นทางการสื่อสารที่ปลอดภัยจากการดักฟัง ปรับปรุงความปลอดภัยทางไซเบอร์ซึ่งกำลังกลายเป็นอันตรายมากขึ้นอย่างรวดเร็ว [21, 22]



ภาพที่ 2-1 QML Framework Matrix [6]

2.4 อัลกอริทึมการจัดกลุ่มข้อมูลด้วยควอนตัม (Quantum Clustering Algorithms)

การพัฒนาการจัดกลุ่มแบบ K-Means เชิงควอนตัม ในระดับที่ก้าวหน้ามีหลายรูปแบบ เช่น การใช้ อัลกอริทึมเชิงพันธุกรรมเชิงควอนตัม (Quantum Genetic Algorithm)

ซึ่งช่วยในการเลือกจุดศูนย์กลางของกลุ่มได้ดีขึ้นพร้อมทั้งลดสัญญาณรบกวน โดยอัลกอริทึมเชิงพันธุกรรมเชิงควอนตัม ช่วยลดความไวของการกำหนดจุดศูนย์กลางเริ่มต้นและเพิ่มความเข้มข้นของอัลกอริทึมต่อสัญญาณรบกวนผ่านการใช้เครื่องมือที่เรียกว่า ออโตเอนโคเดอร์แบบลดสัญญาณรบกวน (Noise reduction self-encoder) ส่งผลให้โมเดลที่พัฒนาขึ้นสามารถทำงานได้ดีแม้จะมีสัญญาณรบกวนอยู่ โดยโมเดล ทั้งสองแสดงให้เห็นจำนวนรอบการทำงานที่น้อยลง มีโอกาสน้อยที่จะติดอยู่ในคำตอบที่ดีเฉพาะบริเวณหนึ่ง และทนทานต่อสัญญาณรบกวนได้ดีขึ้นหลังจากที่มีการปรับปรุงจุดศูนย์กลางของกลุ่ม [23]

CVQK-means ซึ่งเป็นวิธีที่ใช้ สถานะเกาส์เซียน (Gaussian states) บนฮาร์ดแวร์แบบโฟตอนิกส์ ให้ผลลัพธ์ความแม่นยำในการจัดกลุ่มที่ใกล้เคียงกับ K-means แบบเดิม และมีค่าผลรวมของระยะทางกำลังสอง (Sum of Square Distances) ที่ใกล้เคียงกัน แม้ว่าจะมีปัญหาเรื่องสัญญาณรบกวนในฮาร์ดแวร์ควอนตัมที่อาจส่งผลกระทบต่อความแม่นยำของการทำงาน [24] QCSWAPK-means ให้ผลลัพธ์ที่ดีมากด้วยคะแนน Silhouette สูงสุดที่ 0.491 และ Davies-Bouldin index ที่ 0.739 เมื่อเทียบกับ K-means ธรรมดาที่ได้ 0.451 แสดงให้เห็นว่ากลุ่มที่ได้มีความชัดเจนและแยกจากกันมากกว่า การประยุกต์ใช้การแมปฟีเจอร์ (Feature mapping) ในวงจรแปรผันเพื่อให้ได้ประสิทธิภาพที่ดีขึ้นในพื้นที่สถานะควอนตัม โดยใช้โปรเซสเซอร์ควอนตัม 5-qubit ibmqx2 และ 14-qubit ibmq_16_melbourne ในการนำอัลกอริทึมไปใช้ ซึ่งโปรเซสเซอร์ 5-qubit มีประสิทธิภาพดีกว่า 14-qubit เนื่องจากมี Fredkin หรือ controlled-SWAP (CSWAP) gates น้อยกว่า [25] นอกจากนี้ยังมีการพัฒนาเพิ่มเติม เช่น การใช้ asymmetric Parzen estimators และ quantum fidelity metrics โดย Quantum Fidelity Based Fuzzy C-Means หรือ QFFCM ได้นำหลักการของการคำนวณเชิงควอนตัม มาผสมผสานกับระบบ fuzzy ที่มีโครงสร้างเพื่อคำนวณระยะทางระหว่างจุดข้อมูลในแบบควอนตัม [26, 27]

2.5 การศึกษาเปรียบเทียบผลลัพธ์ต่างๆ (Performance Comparison)

จากการศึกษาเปรียบเทียบพบว่าผลลัพธ์มีความหลากหลาย โดย QSVM หรือ Quantum Support Vector Machine สามารถให้ความแม่นยำถึง 100% ในชุดข้อมูล Seeds เมื่อเทียบกับวิธีแบบดั้งเดิมที่ได้ 94.97% การศึกษายังแสดงให้เห็นว่า QSVM สามารถจำแนกข้อมูลคลาสสิกได้อย่างมีประสิทธิภาพโดยใช้แผนที่คุณลักษณะ ZZ เพื่อเข้ารหัสเป็นสถานะควอนตัม โดย ZZFeatureMap ถูกสร้างขึ้นด้วยมิติคุณลักษณะ 13 มิติตามชุดข้อมูล Cleveland Heart Disease และการทำซ้ำครั้งเดียว สร้างวงจรควอนตัมที่มีพารามิเตอร์ ผลการทดลองแสดงให้เห็นว่า QSVC ทำงานได้ดีกว่า VQC ในทุกเมตริกซ์ประสิทธิภาพ QSVC ใช้อย่างมีประสิทธิภาพหลักการกลศาสตร์ควอนตัมในการจำแนกข้อมูล เมื่อประยุกต์ใช้กับชุดข้อมูล Cleveland Heart Disease

QSVC บรรลุความแม่นยำ 0.75 ในขณะที่ VQC ให้ความแม่นยำเพียง 0.62 แสดงให้เห็นชัดเจนว่า QSVC ทำงานได้ดีกว่าในการจำแนกโรคหัวใจ [28, 29]

เรื่องของเวลาในการฝึกสอนเป็นความท้าทายที่สำคัญ โดย QSVC ต้องใช้เวลาเพิ่มขึ้นแบบทวีคูณเกินกว่า 35 ชั่วโมงสำหรับข้อมูล 400-600 รายการ จากการศึกษาเบื้องต้นพบว่า QSVC ใช้เวลานานกว่าอัลกอริทึมการเรียนรู้ของเครื่องเชิงควอนตัม และอัลกอริทึมแบบเดิมอื่นๆ อย่างมาก เมื่อขนาดของข้อมูลเข้าใกล้ 400 ถึงมากกว่า 600 รายการ เวลาที่ใช้ก็จะเพิ่มขึ้นอย่างรวดเร็วมาก จากกราฟที่แสดงผลจะเห็นได้ว่าการเพิ่มขึ้นของเวลามีลักษณะแบบทวีคูณ ซึ่งเป็นข้อกังวลเกี่ยวกับความสามารถในการตอบสนองของ QSVC เมื่อข้อมูลมีขนาดใหญ่ขึ้น [10, 29] การจัดกลุ่มแบบ K-Means เชิงควอนตัม แสดงให้เห็นข้อได้เปรียบด้วยค่าเฉลี่ย WCSS ที่ 2,654,322 เมื่อเทียบกับแบบคลาสสิกที่ได้ 2,937,604 และมี Silhouette coefficient ที่สูงกว่าคือ 0.2541 เทียบกับ 0.1970 การทดลองกับชุดข้อมูล MNIST พบว่าอัลกอริทึมการจัดกลุ่มแบบ K-Means เชิงควอนตัม ที่ผ่านการปรับปรุงให้ผลการจัดกลุ่มที่ดีกว่าแบบเดิมเมื่อมีสัญญาณรบกวน [23, 30]

2.6 การวิเคราะห์ลูกค้าในอุตสาหกรรมร้านอาหาร (Restaurant Industry Customer Analytics)

ร้านอาหารในยุคปัจจุบันได้หันมาใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) และการเรียนรู้ของเครื่องเพื่อเพิ่มประสิทธิภาพในการทำงาน ตัวอย่างที่โดดเด่นคือระบบ SmartDine ซึ่งเป็นระบบที่ครอบคลุมและออกแบบมาเพื่อเพิ่มความพึงพอใจของลูกค้าและเพิ่มประสิทธิภาพการดำเนินงานภายในอุตสาหกรรมร้านอาหาร วิธีการของระบบมุ่งเน้นไปที่การประเมินเมตริกซ์สำคัญ เช่น เวลารอการรับบริการของลูกค้า การโต้ตอบระหว่างพนักงานกับลูกค้า การมีส่วนร่วมกับโต๊ะ การทำนายอัตราการหมุนเวียนของโต๊ะ และการจัดสรรทรัพยากรแบบไดนามิก โดยใช้ประโยชน์จากการเรียนรู้เชิงลึก การประมวลผลภาพ และเทคนิคการเรียนรู้ของเครื่อง SmartDine ให้ข้อมูลเชิงลึกที่สามารถนำไปปฏิบัติได้เพื่อปรับปรุงประสบการณ์ของลูกค้าและประสิทธิภาพการดำเนินงาน [31]

แนวทางนี้ช่วยให้สามารถตรวจสอบและวิเคราะห์เมตริกซ์สำคัญได้แบบเรียลไทม์ ซึ่งแตกต่างจากวิธีการแบบดั้งเดิม SmartDine ใช้ภาพจากกล้องวงจรปิดและโมเดลการเรียนรู้ของเครื่องที่ล้ำสมัย รวมถึงโมเดลการตรวจจับวัตถุ YOLO และโมเดล LSTM ที่กำหนดเองเพื่อระบุสัญญาณพฤติกรรมและทำนายความต้องการของลูกค้าแบบไดนามิก การวิเคราะห์เชิงรุกและแบบเรียลไทม์นี้ ช่วยให้พนักงานร้านอาหารสามารถจัดการกับปัญหาที่อาจเกิดขึ้นก่อนที่จะส่งผลกระทบต่อประสบการณ์ของลูกค้า ปรับปรุงคุณภาพการบริการโดยรวมและความพึงพอใจของลูกค้า SmartDine ถูกสร้างขึ้นจากหกองค์ประกอบหลัก

โดยแต่ละองค์ประกอบกำหนดเป้าหมายไปที่แง่มุมสำคัญ ของประสบการณ์การรับประทานอาหารและประสิทธิภาพการดำเนินงาน [31]

เทคโนโลยีปัญญาประดิษฐ์ ช่วยในการเพิ่มประสิทธิภาพตัวเลือกเมนูตามแนวโน้มและความชอบของผู้บริโภค อัลกอริทึมเทคโนโลยีปัญญาประดิษฐ์ สามารถแนะนำการปรับเปลี่ยนเมนูที่จะเพิ่มความพึงพอใจและรายได้ของลูกค้าโดยการวิเคราะห์ข้อมูลว่าผู้คนสั่งอะไร ความถี่ในการสั่งและเวลาที่สั่ง นอกจากนี้เทคโนโลยีปัญญาประดิษฐ์ ยังช่วยในการปรับเมนูให้เหมาะสม ในขณะที่แชทบอทและผู้ช่วยเสมือนที่ขับเคลื่อนด้วยเทคโนโลยีปัญญาประดิษฐ์ ใช้เพื่อรับคำสั่งซื้อ ตอบคำถามของผู้บริโภค และให้คำแนะนำที่ปรับแต่งได้ ส่งผลให้ผู้บริโภคสามารถเพลิดเพลิน กับประสบการณ์การรับประทานอาหารที่ราบรื่นและน่าพอใจมากขึ้น ในขณะเดียวกัน ได้ช่วยประหยัดเวลาและเงินในค่าแรงงาน การวิเคราะห์เชิงคาดการณ์ ซึ่งอาจคาดการณ์ความต้องการสำหรับรายการเฉพาะเป็นอีกวิธีหนึ่งที่เทคโนโลยีปัญญาประดิษฐ์ ถูกใช้ในร้านอาหาร [32]

การพัฒนาที่สำคัญอื่นๆ ในระบบอัตโนมัติร้านอาหารรวมถึงการใช้ระบบสั่งและชำระเงินอัตโนมัติที่ลดเวลารอและปรับปรุงความแม่นยำของคำสั่งซื้ออย่างมีนัยสำคัญ ผู้ช่วยการทำอาหารด้วยหุ่นยนต์นำระบบอัตโนมัติมาสู่การเตรียมอาหารและสร้างความสม่ำเสมอและความแม่นยำในการทำอาหาร นวัตกรรมนี้ไม่เพียงแต่ปรับปรุงการดำเนินงานครัวเท่านั้น แต่ยังเพิ่มความน่าสนใจแบบอนาคตให้กับประสบการณ์การรับประทานอาหาร ร้านอาหารใช้ประโยชน์จากพลังของการวิเคราะห์ข้อมูลเพื่อรับข้อมูลเชิงลึกเกี่ยวกับความชอบของลูกค้า ปรับปรุงการจัดการสินค้าคงคลังและเพิ่มประสิทธิภาพห่วงโซ่อุปทาน การแพร่ระบาดของ COVID-19 ได้เร่งการนำโซลูชันแบบไม่สัมผัสมาใช้ เช่น เมนูที่ใช้ QR code และการชำระเงินผ่านมือถือ เพื่อความสะอาดและความปลอดภัย [33]

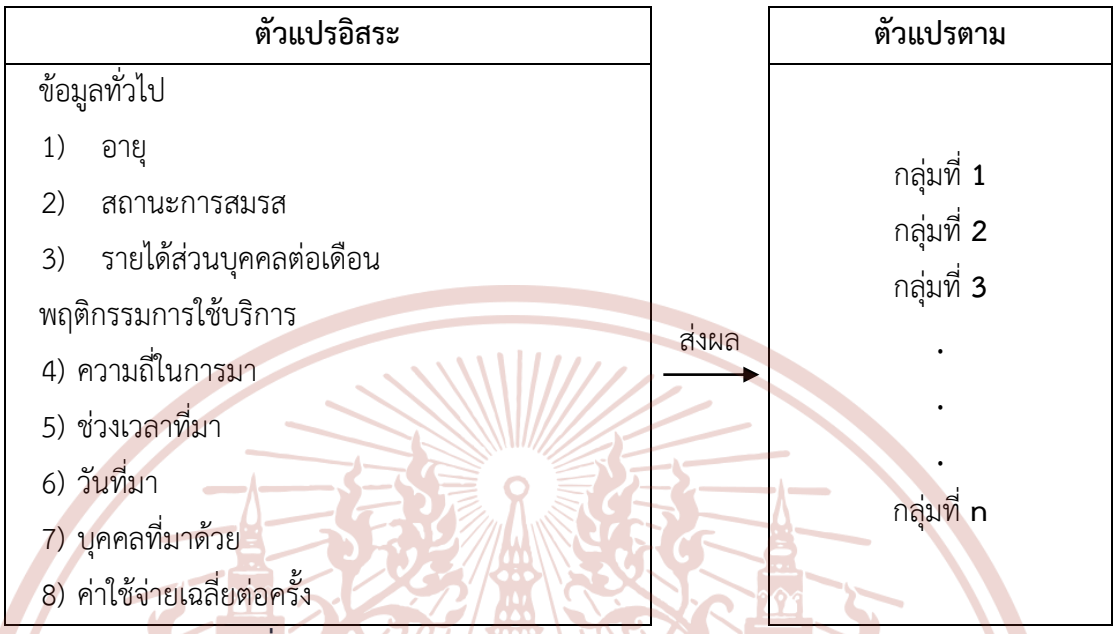
บทที่ 3
วิธีการดำเนินงานวิจัย

การวิจัยนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Research) ที่ใช้วิธีการวิเคราะห์ข้อมูลเปรียบเทียบ (Comparative Analysis) ด้วยเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเทคนิค Quantum Machine Learning Clustering กับเทคนิคการแบ่งกลุ่มแบบดั้งเดิม ได้แก่ K-Means, DBSCAN และ Gaussian Mixture Model สำหรับการแบ่งกลุ่มลูกค้าธุรกิจร้านอาหารกลางคืน การวิจัยนี้จะดำเนินการโดยการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างที่เป็นลูกค้าร้านอาหารกลางคืนจริง จากนั้นนำข้อมูลมาวิเคราะห์ด้วยเทคนิคการแบ่งกลุ่มต่างๆ และประเมินประสิทธิภาพโดยใช้เกณฑ์การประเมินมาตรฐาน โดยงานวิจัยนี้มีขั้นตอนการวิจัยดังนี้

3.1 โมเดลการวิจัย
3.2 ขั้นตอนการดำเนินงานวิจัย
3.3 ประชากรและกลุ่มตัวอย่าง
3.4 เครื่องมือที่ใช้ในการวิจัย
3.5 การประยุกต์ใช้เทคนิคการแบ่งกลุ่ม
3.6 สถิติที่ใช้ในการวิเคราะห์ข้อมูล

3.1 โมเดลการวิจัย

การวิจัยนี้ศึกษาข้อมูลลูกค้าและการแบ่งกลุ่มลูกค้า โดยตัวแปรอิสระในการวิจัยประกอบด้วยข้อมูลลูกค้าทั้งหมดที่แบ่งเป็น 2 กลุ่มหลัก คือ ข้อมูลทั่วไป (Demographic Data) ซึ่งประกอบด้วย อายุ สถานะการสมรสและรายได้ และพฤติกรรมการใช้บริการ (Behavioral Data) ซึ่งประกอบด้วย ความถี่ในการมาใช้บริการ ช่วงเวลาที่มาใช้บริการ วันที่มาใช้บริการ บุคคลที่มาด้วย และค่าใช้จ่ายเฉลี่ยต่อครั้ง ส่วนตัวแปรตามคือกลุ่มลูกค้าที่ได้จากการแบ่งกลุ่ม (Customer Segments) ซึ่งจะได้กลุ่มลูกค้าจำนวน n กลุ่ม ขึ้นอยู่กับลักษณะของข้อมูลและผลลัพธ์จากการวิเคราะห์ แต่ละกลุ่มจะมีลักษณะเฉพาะที่แตกต่างกันในด้านข้อมูลทั่วไปและพฤติกรรมการใช้บริการ



ภาพที่ 3-1 โมเดลแนวคิดการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน

3.2 ขั้นตอนการดำเนินงานวิจัย

3.2.1 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเป็นขั้นตอนแรกและเป็นขั้นตอนที่สำคัญของการวิจัย โดยผู้วิจัยได้ดำเนินการเก็บข้อมูลเป็น 3 ขั้นตอน ขั้นตอนแรกคือการเก็บข้อมูล ผู้วิจัยได้ประสานงานกับเจ้าของร้านอาหารเพื่อขอความร่วมมือในการเก็บข้อมูล พร้อมทั้งอธิบายวัตถุประสงค์และประโยชน์ของการวิจัย จากนั้นจัดทำแบบสอบถามออนไลน์ในรูปแบบของ Microsoft Forms พร้อมสร้าง QR Code เพื่อให้ลูกค้าสแกนและกรอกแบบสอบถามได้สะดวก นอกจากนี้ยังได้เตรียมเอกสารขอความยินยอม (Consent Form) เพื่อให้ผู้เข้าร่วมวิจัยทราบถึงวัตถุประสงค์ การใช้ข้อมูล และสิทธิ ในการปกป้องข้อมูลส่วนบุคคล

ขั้นตอนที่สองคือการเก็บรวบรวมข้อมูล ผู้วิจัยได้จัดวาง QR Code ของแบบสอบถามไว้บนโต๊ะทุกโต๊ะในร้านอาหาร พร้อมทั้งแจ้งให้พนักงานเสิร์ฟช่วยแนะนำและชักจูงให้ลูกค้ากรอกแบบสอบถาม โดยมีการให้สิทธิพิเศษหรือส่วนลดเล็กน้อยเป็นแรงจูงใจ การเก็บข้อมูลดำเนินการอย่างต่อเนื่อง เป็นระยะเวลา 3 เดือน โดยผู้วิจัยได้ตรวจสอบความครบถ้วนและคุณภาพของข้อมูลที่ได้รับเป็นประจำทุกสัปดาห์ เพื่อให้มั่นใจว่าจะได้ข้อมูลที่มีคุณภาพและเพียงพอสำหรับการวิเคราะห์

ขั้นตอนที่สามคือการจัดเก็บข้อมูล เมื่อเก็บรวบรวมข้อมูลจากแบบสอบถามออนไลน์ได้ครบตามจำนวนที่กำหนดแล้ว ผู้วิจัยได้ทำการ Export ข้อมูลออกมาในรูปแบบไฟล์ CSV ซึ่งเป็นรูปแบบที่สะดวกต่อการนำไปประมวลผลด้วยโปรแกรมต่างๆ พร้อมทั้งสำรองข้อมูลไว้หลายชุด

เพื่อป้องกันการสูญหายของข้อมูล โดยเก็บไว้ทั้งใน Local Storage, Cloud Storage และ External Hard Drive

3.2.2 การเตรียมข้อมูล

การเตรียมข้อมูลเป็นขั้นตอนที่สำคัญที่จะทำให้ข้อมูลมีคุณภาพและพร้อมสำหรับการวิเคราะห์ ผู้วิจัยได้ดำเนินการเตรียมข้อมูลเป็น 5 ขั้นตอนหลัก

ขั้นตอนแรกคือการทำความสะอาดข้อมูล (Data Cleaning) ผู้วิจัยได้ตรวจสอบข้อมูลทั้งหมดเพื่อหาและลบข้อมูลที่ซ้ำซ้อน (Duplicate data) ซึ่งอาจเกิดจากลูกค้าคนเดียวกันกรอกแบบสอบถามหลายครั้ง และตรวจสอบความถูกต้องของข้อมูลตามเกณฑ์ที่กำหนด เช่น ค่าใช้จ่ายต้องเป็นตัวเลขบวก อายุต้องอยู่ในช่วงที่สมเหตุสมผล

ขั้นตอนที่สองคือการแปลงข้อมูล (Data Transformation) ผู้วิจัยได้แปลงตัวแปรเชิงคุณภาพให้เป็นเชิงปริมาณเพื่อให้สามารถนำไปใช้ในการวิเคราะห์ได้ โดยใช้เทคนิค Encoding ซึ่งแบ่งเป็น 2 ประเภท คือ Label Encoding สำหรับตัวแปรที่มีลำดับ (Ordinal) เช่น อายุ รายได้ต่อเดือน ความถี่ในการมา ค่าใช้จ่ายเฉลี่ยต่อครั้ง ซึ่งจะแปลงเป็นตัวเลข 0, 1, 2, 3 ตามลำดับ และ One-Hot Encoding สำหรับตัวแปรที่ไม่มีลำดับ (Nominal) เช่น สถานะการสมรส ช่วงเวลาที่มา วันที่มา บุคคลที่มาด้วย ซึ่งจะแปลงเป็น Binary columns หลายคอลัมน์ นอกจากนี้ยังได้ปรับมาตราส่วนข้อมูล (Normalization) โดยใช้เทคนิค Min-Max Scaling เพื่อแปลงค่าให้อยู่ในช่วง 0 ถึง 1 หรือใช้ Standardization (Z-score normalization) เพื่อแปลงให้มีค่าเฉลี่ยอยู่ที่ 0 และส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 1 ขึ้นอยู่กับลักษณะของข้อมูลและความเหมาะสมกับอัลกอริทึมที่จะใช้

ขั้นตอนที่สามคือการคัดเลือกคุณลักษณะที่สำคัญ (Feature Selection) ผู้วิจัยได้ใช้เทคนิค Feature Importance จาก Machine Learning models เพื่อหาตัวแปรที่มีความสำคัญต่อการแบ่งกลุ่ม จากนั้นเลือกเฉพาะตัวแปรที่มีความสำคัญและเกี่ยวข้องกับวัตถุประสงค์การวิจัยมาใช้ในการวิเคราะห์

3.3 ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัยครั้งนี้คือลูกค้าที่มาใช้บริการร้านอาหารกลางคืนในเขตกรุงเทพมหานคร ที่เปิดให้บริการในช่วงเวลา 19.00-02.00 น. สำหรับกลุ่มตัวอย่างในการวิจัยนี้จำนวน 6,025 คน ซึ่งเป็นลูกค้าที่มาใช้บริการร้านอาหารกลางคืน 1 ร้านในเขตกรุงเทพมหานคร โดยเก็บข้อมูลในช่วงระยะเวลา 3 เดือน ตั้งแต่เดือนกันยายน ถึงเดือนพฤศจิกายน พ.ศ. 2568 การเลือกกลุ่มตัวอย่างใช้เกณฑ์การคัดเข้า (Inclusion Criteria) ได้แก่ ต้องเป็นลูกค้าที่มาใช้บริการร้านอาหารในช่วงเวลา 19.00-02.00 น. ยินยอมให้ใช้ข้อมูลเพื่อการวิจัย และมีข้อมูลครบถ้วน

ตามที่กำหนด ส่วนเกณฑ์การคัดออก (Exclusion Criteria) ได้แก่ ข้อมูลไม่ครบถ้วนหรือมีข้อมูลผิดพลาดชัดเจน และไม่ยินยอมให้ใช้ข้อมูลเพื่อการวิจัย

3.4 เครื่องมือที่ใช้ในการวิจัย

3.4.1 เครื่องมือในการเก็บรวบรวมข้อมูล

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลคือแบบสอบถามออนไลน์ (Online Questionnaire) ที่ผู้วิจัยสร้างขึ้นเพื่อเก็บข้อมูลพฤติกรรมการใช้บริการของลูกค้าร้านอาหารกลางคืน โดยแบบสอบถามแบ่งออกเป็น 2 ส่วนหลัก ส่วนแรกเป็นข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ได้แก่ เพศ อายุ สถานะการสมรส อาชีพ และรายได้ส่วนบุคคลต่อเดือน ส่วนที่สองเป็นพฤติกรรมการใช้บริการ ประกอบด้วยคำถามเกี่ยวกับความถี่ในการมาใช้บริการ ช่วงเวลาและวันที่มาใช้บริการ บุคคลที่มาด้วยและจำนวนคน ค่าใช้จ่ายเฉลี่ยต่อครั้ง เคยใช้บริการร้านอื่นหรือไม่ และช่วงเวลาที่ต้องการรับข้อมูลจากทางร้าน

3.4.2 การตรวจสอบคุณภาพเครื่องมือ

การตรวจสอบคุณภาพของเครื่องมือที่ใช้ในการวิจัยเป็นขั้นตอนที่สำคัญเพื่อให้มั่นใจว่าแบบสอบถามที่สร้างขึ้นมีคุณภาพและสามารถเก็บข้อมูลได้อย่างถูกต้องและน่าเชื่อถือ ผู้วิจัยได้ดำเนินการตรวจสอบคุณภาพของเครื่องมือในด้าน ความเที่ยงตรง (Content Validity) สำหรับความเที่ยงตรงเชิงเนื้อหา (Content Validity) ผู้วิจัยได้นำแบบสอบถามที่สร้างขึ้นเสนอต่อผู้เชี่ยวชาญจำนวน 3 ท่าน ซึ่งประกอบด้วยผู้เชี่ยวชาญด้านการตลาด ผู้เชี่ยวชาญด้านธุรกิจร้านอาหาร และผู้เชี่ยวชาญด้านการวิเคราะห์ข้อมูล เพื่อตรวจสอบความเที่ยงตรงเชิงเนื้อหา โดยพิจารณาจากความสอดคล้องของข้อคำถามกับวัตถุประสงค์การวิจัย ความครอบคลุมของเนื้อหา ความเหมาะสมของภาษาที่ใช้ และความชัดเจนของคำถาม จากนั้นคำนวณค่า IOC (Index of Item-Objective Congruence) เพื่อหาค่าความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ โดยผู้เชี่ยวชาญ แต่ละท่านจะให้คะแนนแต่ละข้อคำถามเป็น +1 หากแน่ใจว่าข้อคำถามสอดคล้องกับวัตถุประสงค์ 0 หากไม่แน่ใจว่าข้อคำถามสอดคล้องกับวัตถุประสงค์ และ -1 หากแน่ใจว่าข้อคำถามไม่สอดคล้อง กับวัตถุประสงค์ จากนั้นนำคะแนนที่ได้มาคำนวณหาค่า IOC ซึ่งข้อคำถามที่มีค่า IOC มากกว่า หรือเท่ากับ 0.5 จะถือว่ามีความเที่ยงตรงเชิงเนื้อหาและสามารถนำไปใช้ได้ ส่วนข้อคำถามที่มีค่า IOC ต่ำกว่า 0.5 จะต้องปรับปรุงหรือตัดออก

3.4.3 เครื่องมือในการวิเคราะห์ข้อมูล

เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลเป็นซอฟต์แวร์และไลบรารีต่างๆ ที่ใช้ในการประมวลผลข้อมูลและพัฒนาโมเดลการแบ่งกลุ่ม โดยใช้ภาษา Python เวอร์ชัน 3.8 ขึ้นไป เป็นภาษาโปรแกรมหลักในการวิเคราะห์ข้อมูล สำหรับการพัฒนาเทคนิค การจัดกลุ่มด้วยการเรียนรู้

ของเครื่องเชิงควอนตัม ผู้วิจัยใช้ไลบรารี Qiskit ซึ่งเป็นเครื่องมือสำหรับการพัฒนาและจำลองการทำงานของอัลกอริทึมควอนตัม ส่วนการพัฒนาเทคนิคการแบ่งกลุ่มแบบดั้งเดิมใช้ไลบรารี Scikit-learn ซึ่งมีอัลกอริทึมการแบ่งกลุ่มต่างๆ เช่น K-Means, DBSCAN และ Gaussian Mixture Model พร้อมใช้งาน สำหรับการจัดการและประมวลผลข้อมูลใช้ไลบรารี Pandas และสำหรับการคำนวณ ทางคณิตศาสตร์ใช้ไลบรารี NumPy และ SciPy ส่วนการแสดงผลข้อมูลและสร้างกราฟต่างๆ ใช้ไลบรารี Matplotlib และ Seaborn

3.5 การประยุกต์ใช้เทคนิคการแบ่งกลุ่ม

หลังจากที่ข้อมูลได้รับการเตรียมและพร้อมสำหรับการวิเคราะห์แล้ว ผู้วิจัยได้ดำเนินการประยุกต์ใช้เทคนิคการแบ่งกลุ่มทั้ง 4 เทคนิค เพื่อเปรียบเทียบประสิทธิภาพ ได้แก่ K-Means Clustering, DBSCAN, Gaussian Mixture Model และ การจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม

3.5.1 K-Means Clustering

K-Means เป็นเทคนิคที่เหมาะสมกับข้อมูลที่กลุ่มมีรูปร่างกลมๆ และมีขนาดใกล้เคียงกัน เนื่องจากอัลกอริทึมใช้การวัดระยะทาง Euclidean และคำนวณจุดศูนย์กลางจากค่าเฉลี่ย ทำให้เหมาะกับกลุ่มที่มีการกระจายตัวแบบสมมาตร จุดเด่นของเทคนิคนี้คือความเร็วในการประมวลผลและความง่ายในการใช้งาน สามารถนำไปประยุกต์ใช้ได้ทันทีโดยไม่ต้องปรับแต่งพารามิเตอร์มากนัก อีกทั้งผลลัพธ์ที่ได้ยังสามารถตีความได้ง่ายเพราะแต่ละจุดข้อมูลจะถูกจัดเข้ากลุ่มใดกลุ่มหนึ่งอย่างชัดเจน

อย่างไรก็ตาม K-Means มีข้อจำกัดสำคัญหลายประการ ประการแรกคือต้องกำหนดจำนวนกลุ่ม k ล่วงหน้า ซึ่งในความเป็นจริงเราอาจไม่ทราบว่าควรมีกี่กลุ่ม แม้จะใช้ Elbow Method ช่วยก็ยังต้องอาศัยการตีความจากผู้วิจัย ประการที่สองคือไม่สามารถหากลุ่มที่มีรูปร่างแปลกๆ ได้ เช่น กลุ่มที่เป็นรูปวงรี รูปโค้ง หรือรูปตัวเอส เพราะอัลกอริทึมจะพยายามสร้างกลุ่มที่เป็นวงกลมเสมอ ประการที่สามคือมีความอ่อนไหวต่อจุดแปลกปลอม (Outliers) เพราะการคำนวณค่าเฉลี่ยจะถูกดึงไปหาจุดแปลกปลอมเหล่านั้น ทำให้จุดศูนย์กลางเคลื่อนไปจากตำแหน่งที่เหมาะสม และประการสุดท้ายคือผลลัพธ์อาจแตกต่างกันในแต่ละครั้งที่รันเพราะการเลือกจุดศูนย์กลางเริ่มต้นแบบสุ่ม แม้ว่าการใช้ k-means++ จะช่วยลดปัญหานี้ได้บ้าง

3.5.2 DBSCAN

DBSCAN เป็นเทคนิคที่โดดเด่นในการจัดการกับข้อมูลที่มีลักษณะพิเศษ เหมาะสมกับข้อมูลที่กลุ่มมีรูปร่างไม่สม่ำเสมอและมีจุดแปลกปลอมปะปนอยู่ ความสามารถในการค้นหากลุ่มที่มีรูปร่างใดก็ได้ทำให้เทคนิคนี้มีความยืดหยุ่นสูงกว่า K-Means มาก ไม่ว่าจะเป็น

กลุ่มรูปโค้ง รูปวงแหวน หรือรูปร่างซับซ้อนอื่นๆ DBSCAN สามารถจับกลุ่มเหล่านี้ได้อย่างแม่นยำ นอกจากนี้ความสามารถ ในการระบุและแยกจุดแปลกปลอม (Noise) ออกจากกลุ่มหลักเป็นคุณสมบัติที่มีค่ามาก ในการวิเคราะห์ข้อมูลจริง เพราะข้อมูลในโลกจริงมักมีข้อมูลผิดปกติหรือข้อผิดพลาดปะปนอยู่เสมอ ข้อดีอีกประการคือไม่ต้องกำหนดจำนวนกลุ่มล่วงหน้า อัลกอริทึมจะค้นหาจำนวนกลุ่มเองตามความหนาแน่นของข้อมูล

แต่ข้อจำกัดของ DBSCAN ก็มีความสำคัญไม่แพ้กัน การตั้งค่าพารามิเตอร์ epsilon และ MinPts เป็นเรื่องที่ค่อนข้างยากและต้องอาศัยประสบการณ์ ต้องลองหลายค่าแล้วดูผลลัพธ์ว่าเหมาะสมหรือไม่ ไม่มีสูตรตายตัวในการหาค่าที่เหมาะสม นอกจากนี้ DBSCAN ไม่เหมาะกับข้อมูลที่แต่ละกลุ่มมีความหนาแน่นแตกต่างกันมาก เพราะค่า epsilon ที่ตั้งไว้อาจเหมาะกับกลุ่มหนึ่งแต่ไม่เหมาะกับอีกกลุ่ม ส่งผลให้บางกลุ่มอาจถูกแบ่งย่อยเกินไปหรือรวมกันเกินไป อีกทั้งในกรณีที่ข้อมูลมีมิติสูง ประสิทธิภาพของ DBSCAN จะลดลงเพราะแนวคิดของระยะห่างและความหนาแน่น จะไม่ชัดเจนเท่าในข้อมูลมิติต่ำ

3.5.3 Gaussian Mixture Model (GMM)

GMM เป็นเทคนิคที่มีความซับซ้อนและมีความยืดหยุ่นสูง เหมาะสมกับข้อมูลที่ต้องการทราบความน่าจะเป็นในการเป็นสมาชิกของแต่ละกลุ่มมากกว่าการจัดกลุ่มแบบตายตัว ความสามารถในการบอกว่าจุดข้อมูลหนึ่งมีโอกาสเป็นสมาชิกของหลายกลุ่มพร้อมกันเป็นคุณสมบัติที่มีค่ามากในการวิเคราะห์ข้อมูลที่มีความซ้อนทับกันระหว่างกลุ่ม เช่น ลูกค้าที่มีพฤติกรรมผสมผสานระหว่างหลายกลุ่ม GMM สามารถจับลักษณะนี้ได้ดีกว่า K-Means ที่บังคับให้แต่ละจุดอยู่ในกลุ่มเดียว นอกจากนี้ GMM ยังมีความยืดหยุ่นสูงเพราะสามารถปรับรูปร่างและขนาดของกลุ่มได้ผ่านพารามิเตอร์ covariance matrix ทำให้สามารถจัดการกับกลุ่มที่มีรูปร่างเป็นวงรีหรือมีการกระจายตัวที่ไม่เท่ากันในแต่ละมิติได้ดี ผลลัพธ์ที่เป็นความน่าจะเป็นยังช่วยให้เราเข้าใจความไม่แน่นอนในการจัดกลุ่มได้ดีขึ้น

ข้อจำกัดของ GMM มีหลายประการที่ต้องคำนึงถึง ประการแรกคือความซับซ้อนในการคำนวณที่สูงกว่าเทคนิคอื่น ทำให้ใช้เวลาในการประมวลผลนานกว่ามาก โดยเฉพาะเมื่อข้อมูลมีขนาดใหญ่ หรือมีมิติสูง ประการที่สองคือ EM Algorithm ที่ใช้อาจลู่เข้าสู่ local optimum ได้ทำให้ผลลัพธ์ไม่ใช่คำตอบที่ดีที่สุดเสมอไป อาจต้องรันหลายครั้งด้วยค่าเริ่มต้นที่ต่างกันแล้วเลือกผลลัพธ์ที่ดีที่สุด ประการที่สามคือยังคงต้องกำหนดจำนวนกลุ่มล่วงหน้าเหมือน K-Means แม้ว่าจะใช้ BIC หรือ AIC ช่วยก็ตาม แต่การหาจำนวนที่เหมาะสมก็ยังใช้เวลาและทรัพยากรในการคำนวณมาก และประการสุดท้ายคือ GMM สมมติว่าข้อมูลแต่ละกลุ่มมีการแจกแจงแบบ Gaussian ถ้าข้อมูลจริงไม่เป็นไปตามสมมติฐานนี้ ผลลัพธ์อาจไม่ดีเท่าที่ควร

3.5.4 Quantum Machine Learning Clustering

การจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม (Quantum Machine Learning Clustering) เป็นเทคนิคที่นำเสนอในงานวิจัยนี้ มีศักยภาพสูงในการจับความสัมพันธ์ที่ซับซ้อนของข้อมูลได้ดีกว่าเทคนิคแบบคลาสสิก การใช้ quantum distance metrics เช่น quantum fidelity-based distance ทำให้สามารถวัดความคล้ายคลึงกันของข้อมูลในมิติที่ลึกซึ้งกว่าการวัดระยะทาง Euclidean แบบธรรมดา แนวคิดของ quantum superposition ช่วยให้อัลกอริทึมสามารถสำรวจหลายทางเลือกพร้อมกันในการหาจุดศูนย์กลางที่เหมาะสม แทนที่จะต้องลองทีละทางเหมือนเทคนิคคลาสสิก ทำให้มีโอกาสหาคำตอบที่ดีกว่าได้ การใช้ quantum genetic algorithm ในการหาตำแหน่งจุดศูนย์กลางที่เหมาะสมยังช่วยลดปัญหาการติด local optimum ได้ดีกว่าวิธีการแบบดั้งเดิม เทคนิคนี้มีความน่าสนใจเป็นพิเศษสำหรับข้อมูลที่มีความซับซ้อนสูง มีมิติมาก หรือมีโครงสร้างที่ไม่สามารถจับได้ด้วยเทคนิคแบบคลาสสิก

อย่างไรก็ตามการจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม มีข้อจำกัดที่ต้องพิจารณาอย่างรอบคอบ ประการแรกคือความซับซ้อนในการพัฒนาและ implement อัลกอริทึม ต้องอาศัยความเข้าใจในหลักการของการคำนวณเชิงควอนตัม และต้องใช้ไลบรารีพิเศษ Qiskit ซึ่งต้องใช้ทักษะเฉพาะทาง ประการที่สองคือเวลาในการประมวลผลที่นานกว่าเทคนิคอื่นๆ มาก โดยเฉพาะเมื่อข้อมูลมีขนาดใหญ่ เพราะต้องทำการจำลองระบบควอนตัมซึ่งใช้ทรัพยากรในการคำนวณสูง ประการที่สามคือยังเป็นเทคนิคใหม่ที่อยู่ในระหว่างการพัฒนา ยังไม่มีการยอมรับอย่างกว้างขวางในแวดวงวิชาการและอุตสาหกรรมเท่ากับเทคนิคคลาสสิก อาจมีข้อจำกัดในการนำไปใช้งานจริงหรือการเปรียบเทียบกับงานวิจัยอื่นๆ และประการสุดท้ายคือความเข้าใจและการตีความผลลัพธ์อาจทำได้ยากกว่าเทคนิคแบบดั้งเดิม เพราะกระบวนการทำงานมีความซับซ้อนและใช้แนวคิดที่ไม่คุ้นเคยสำหรับผู้ใช้ทั่วไป

3.5.5 การประเมินประสิทธิภาพ

การประเมินประสิทธิภาพของแต่ละเทคนิคการแบ่งกลุ่มเป็นขั้นตอนที่สำคัญเพื่อเปรียบเทียบว่าเทคนิคใดให้ผลลัพธ์ที่ดีที่สุด ผู้วิจัยใช้เกณฑ์การประเมิน 3 ตัวชี้วัดหลัก ได้แก่ Silhouette Score, Davies-Bouldin Index และ Cluster Cohesion

Silhouette Score เป็นเมตริกที่วัดคุณภาพของการจัดกลุ่มโดยเปรียบเทียบความเหนียวแน่นภายในกลุ่มกับการแยกออกระหว่างกลุ่มโดยรวม ค่าที่ได้จะอยู่ระหว่าง -1 ถึง 1 โดยค่าใกล้ 1 แสดงว่ากลุ่มแยกจากกันได้ดีและจุดข้อมูลอยู่ในกลุ่มที่เหมาะสม ค่าใกล้ 0 แสดงว่าจุดข้อมูลอยู่บนขอบเขตระหว่างสองกลุ่ม และค่าใกล้ -1 แสดงว่าจุดข้อมูลอาจถูกจัดกลุ่มผิด ผู้วิจัยได้คำนวณ Silhouette Score สำหรับผลลัพธ์จากแต่ละเทคนิคและนำมาเปรียบเทียบกัน

Davies-Bouldin Index เป็นเมตริกที่ประเมินอัตราส่วนความคล้ายคลึงเฉลี่ยระหว่างแต่ละกลุ่มกับคู่ที่ใกล้ที่สุด โดยคำนวณจากระยะทางเฉลี่ยภายในกลุ่ม (S) และระยะห่างระหว่างจุดศูนย์กลางของแต่ละคู่กลุ่ม (M) จากนั้นหาอัตราส่วน R สำหรับทุกคู่กลุ่มและเลือกค่าสูงสุดสำหรับแต่ละกลุ่ม สุดท้ายหาค่าเฉลี่ยของทุกกลุ่มเป็น Davies-Bouldin Index ค่าที่ได้จะเป็นค่าบวก โดยค่าต่ำกว่าแสดงว่าการแบ่งกลุ่มดีกว่า และค่าใกล้ 0 แสดงว่ามีการแยกกลุ่มสูงสุด ผู้วิจัยได้คำนวณ Davies-Bouldin Index สำหรับผลลัพธ์จากแต่ละเทคนิคเพื่อประเมินคุณภาพการแบ่งกลุ่ม

Cluster Cohesion เป็นเมตริกที่วัดความกะทัดรัดหรือความแน่นหนาของกลุ่ม โดยคำนวณจากผลรวมของระยะทางยกกำลังสองระหว่างจุดข้อมูลแต่ละจุดภายในกลุ่มกับจุดศูนย์กลางของกลุ่ม จากนั้นรวมค่าของทุกกลุ่มเป็น Total Cluster Cohesion ค่าที่ได้จะเป็นค่าบวก โดยค่าต่ำแสดงว่า มีความกะทัดรัดสูง กล่าวคือกลุ่มแน่นหนาและจุดข้อมูลอยู่ใกล้จุดศูนย์กลาง ส่วนค่าสูงแสดงว่ามีการกระจายตัวมาก ผู้วิจัยได้คำนวณ Cluster Cohesion สำหรับผลลัพธ์จากแต่ละเทคนิคเพื่อเปรียบเทียบความกะทัดรัดของกลุ่มที่ได้

นอกจากการคำนวณเมตริกทั้ง 3 ตัวแล้ว ผู้วิจัยยังได้บันทึกเวลาที่ใช้ในการประมวลผล (Computation Time) ของแต่ละเทคนิคเพื่อประกอบการพิจารณาด้วย ในการนำไปใช้งานจริงนอกจากความแม่นยำแล้ว ความเร็วในการประมวลผลก็เป็นปัจจัยสำคัญเช่นกัน

3.5.6 การวิเคราะห์ผลลัพธ์

หลังจากได้ผลลัพธ์จากการประเมินประสิทธิภาพของทุกเทคนิคแล้ว ผู้วิจัยได้ดำเนินการวิเคราะห์ผลลัพธ์เพื่อหาข้อสรุปและข้อเสนอแนะ โดยแบ่งการวิเคราะห์ออกเป็น 2 ส่วนหลัก

ส่วนแรกคือการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิค ผู้วิจัยได้จัดทำตารางสรุปค่าเมตริกทั้ง 3 ตัว (Silhouette Score, Davies-Bouldin Index, Cluster Cohesion) พร้อมทั้งเวลาที่ใช้ในการประมวลผลของแต่ละเทคนิค เพื่อแสดงประสิทธิภาพโดยรวมของแต่ละเทคนิคในทุกมิติ จากนั้นจัดอันดับเทคนิคต่างๆ ตามแต่ละเมตริกและคำนวณคะแนนรวมโดยอาจกำหนดน้ำหนักให้กับแต่ละเมตริกตามความสำคัญ เพื่อหาเทคนิคที่ให้ผลลัพธ์ดีที่สุดโดยรวม

ส่วนที่สองคือการวิเคราะห์ลักษณะเฉพาะของกลุ่มลูกค้า ผู้วิจัยได้เลือกเทคนิคที่ให้ผลลัพธ์ดีที่สุดจากการเปรียบเทียบในส่วนแรก จากนั้นนำผลลัพธ์การแบ่งกลุ่มจากเทคนิคนั้นมาวิเคราะห์ลักษณะเฉพาะของแต่ละกลุ่มลูกค้าอย่างละเอียด โดยวิเคราะห์ทั้งด้านข้อมูลทั่วไป เช่น การแจกแจงของเพศ ช่วงอายุเฉลี่ย อาชีพที่พบบ่อย และระดับรายได้เฉลี่ย และวิเคราะห์ด้านพฤติกรรม (Behavioral Patterns) เช่น ความถี่ในการมาใช้บริการ ช่วงเวลาที่มาใช้บริการ วันที่มาใช้บริการ รูปแบบการมา และค่าใช้จ่ายเฉลี่ยต่อครั้ง

3.6 สถิติที่ใช้ในการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยนี้ใช้สถิติ 2 ประเภท คือ สถิติเชิงพรรณนา (Descriptive Statistics) และสถิติเชิงอนุมาน (Inferential Statistics)

3.6.1 สถิติเชิงพรรณนา ผู้วิจัยใช้ความถี่และร้อยละ (Frequency and Percentage) ในการอธิบายลักษณะของข้อมูลในส่วนที่ 1 ข้อมูลทั่วไป ได้แก่ เพศ อายุ สถานะ อาชีพ และรายได้ต่อเดือน รวมถึงส่วนที่ 2 พฤติกรรมการใช้บริการ / การรับรู้สื่อ ได้แก่ ความถี่ในการมาใช้บริการ ช่วงเวลาที่มาใช้บริการ วันที่มาใช้บริการ บุคคลที่มาด้วย ค่าใช้จ่ายเฉลี่ยต่อครั้ง ร้านอื่นที่เคยใช้บริการ และช่วงเวลาที่ต้องการรับข้อมูลจากทางร้าน

3.6.2 สถิติเชิงอนุมาน ผู้วิจัยใช้เทคนิคการแบ่งกลุ่ม (Clustering Techniques) ทั้ง 4 วิธี ได้แก่ K-Means Clustering, DBSCAN, Gaussian Mixture Model และ การจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม เพื่อจัดกลุ่มลูกค้าตามพฤติกรรมและลักษณะประชากรศาสตร์ โดยใช้การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรต่างๆ เพื่อระบุรูปแบบและลักษณะเฉพาะของแต่ละกลุ่มลูกค้า จากนั้นประเมินประสิทธิภาพของแต่ละเทคนิคด้วยดัชนีวัดประสิทธิภาพ ได้แก่ Silhouette Score ซึ่งวัดความกระชับของกลุ่มและความแยกจากกลุ่มอื่น, Davies-Bouldin Index ซึ่งวัดความคล้ายคลึงระหว่างกลุ่ม และ Cluster Cohesion ซึ่งวัดความสอดคล้องภายในกลุ่ม เพื่อเลือกเทคนิคที่ให้ผลลัพธ์ดีที่สุดในการสร้างโปรไฟล์ลูกค้าและพัฒนากลยุทธ์การตลาดแบบกำหนดเป้าหมาย

บทที่ 4
ผลการดำเนินงานวิจัย

บทนี้นำเสนอผลการวิเคราะห์ข้อมูลจากการศึกษาเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มลูกค้าทั้ง 4 วิธี ได้แก่ K-Means, DBSCAN, Gaussian Mixture Model (GMM) และการจัดกลุ่มด้วยการเรียนรู้ของเครื่องเชิงควอนตัม (Quantum Machine Learning Clustering) สำหรับธุรกิจร้านอาหารกลางคืน โดยใช้ข้อมูลจากลูกค้าจริงจำนวน 6,025 ราย ที่เก็บรวบรวมอย่างต่อเนื่องตลอดระยะเวลา 3 เดือน ตั้งแต่เดือนกันยายน ถึงเดือนพฤศจิกายน พ.ศ. 2568

การวิจัยครั้งนี้มุ่งเน้นการประเมินประสิทธิภาพของอัลกอริทึมการจัดกลุ่มในหลายมิติ ทั้งด้านคุณภาพการจัดกลุ่ม ความแม่นยำ ความเร็วในการประมวลผล และความเหมาะสมในการนำไปใช้งานจริง โดยใช้เกณฑ์การประเมินมาตรฐาน 4 ตัวชี้วัดหลัก ได้แก่ Silhouette Score, Davies-Bouldin Index, Cluster Cohesion และ Execution Time ซึ่งเป็นมาตรวัดที่ได้รับการยอมรับในวงการวิจัยด้านการเรียนรู้ของเครื่องและการวิเคราะห์ข้อมูล การนำเสนอผลการวิเคราะห์ข้อมูลในบทนี้ได้รับการจัดระเบียบอย่างเป็นระบบ แบ่งออกเป็น 7 ส่วนหลัก เพื่อให้ผู้อ่านสามารถติดตามและทำความเข้าใจผลการวิจัยได้อย่างชัดเจนและครบถ้วน ดังนี้

4.1 การวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

4.2 การวิเคราะห์พฤติกรรมการใช้บริการของกลุ่มตัวอย่าง

4.3 การวิเคราะห์ประสิทธิภาพด้าน Silhouette Score

4.4 การวิเคราะห์ประสิทธิภาพด้าน Davies-Bouldin Index

4.5 การวิเคราะห์ประสิทธิภาพด้าน Cluster Cohesion

4.6 การวิเคราะห์ประสิทธิภาพด้าน Execution Time

4.7 การเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี

การนำเสนอผลในบทนี้จะใช้ทั้งตารางสรุปข้อมูล กราฟแสดงการเปรียบเทียบ และการอธิบายเชิงคุณภาพ เพื่อให้ผู้อ่านสามารถเข้าใจผลการวิจัยได้อย่างครบถ้วนและลึกซึ้ง พร้อมทั้งสามารถนำไปประยุกต์ใช้ในบริบทของธุรกิจร้านอาหารหรือธุรกิจบริการอื่นๆ ได้อย่างมีประสิทธิภาพ

4.1 การวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่าง

ตารางที่ 4-1 จำนวนและร้อยละของข้อมูลทั่วไปที่เกี่ยวกับลักษณะกลุ่มตัวอย่าง

ข้อมูลทั่วไป	กลุ่มตัวอย่าง (n=6,025)	
	จำนวน (คน)	ร้อยละ
1. เพศ		
ชาย	2,169	36.00%
หญิง	3,647	60.53%
LGBTQ+	209	3.47%
รวม	**6,025**	**100.00%**
2. อายุ		
18 – 22 ปี	1,323	21.96%
23 - 30 ปี	4,182	69.41%
31 – 40 ปี	475	7.88%
มากกว่า 40 ปี	45	0.75%
รวม	**6,025**	**100.00%**
3. สถานะภาพสมรส		
โสด	4,491	74.54%
มีแฟน / คู่สมรส	1,494	24.80%
หย่า / แยกทาง	40	0.66%
รวม	**6,025**	**100.0%**
4. อาชีพ		
นักเรียน/นักศึกษา	1,408	23.37%
พนักงานบริษัท	2,749	45.63%
ธุรกิจส่วนตัว	1,024	17.00%
ข้าราชการ	692	11.49%
อาชีพอิสระ	132	2.19%
ว่างงาน	20	0.33%
รวม	**6,025**	**100.0%**

ตารางที่ 4-1 (ต่อ)

ข้อมูลทั่วไป	กลุ่มตัวอย่าง (n=6,025)	
	จำนวน (คน)	ร้อยละ
5. รายได้ต่อเดือน		
น้อยกว่า 15,000 บาท	1,401	23.25%
15,000-30,000 บาท	3,397	56.38%
30,001 – 50,000 บาท	844	14.01%
มากกว่า 50,000 บาท	383	6.36%
รวม	**6,025**	**100.0%**

จากตารางที่ 4-1 พบว่า ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม จำแนกตามเพศ พบว่า ส่วนใหญ่เป็นเพศหญิง จำนวน 3,647 คน คิดเป็นร้อยละ 60.53 รองลงมาเป็นเพศชาย จำนวน 2,169 คน คิดเป็นร้อยละ 36.0 และน้อยที่สุดเป็น LGBTQ+ จำนวน 209 คน คิดเป็นร้อยละ 3.47 ตามลำดับ

จำแนกตามอายุ พบว่า ส่วนใหญ่มีอายุระหว่าง 23-30 ปี จำนวน 4,182 คน คิดเป็นร้อยละ 69.41 รองลงมามีอายุระหว่าง 18-22 ปี จำนวน 1,323 คน คิดเป็นร้อยละ 21.96 อันดับที่ 3 มีอายุระหว่าง 31-40 ปี จำนวน 475 คน คิดเป็นร้อยละ 7.88 และน้อยที่สุดมีอายุมากกว่า 40 ปี จำนวน 45 คน คิดเป็นร้อยละ 0.75 ตามลำดับ

จำแนกตามสถานภาพสมรส พบว่า ส่วนใหญ่มีสถานภาพโสด จำนวน 4,491 คน คิดเป็นร้อยละ 74.54 รองลงมามีแฟน/คู่สมรส จำนวน 1,494 คน คิดเป็นร้อยละ 24.80 และน้อยที่สุด หย่า/แยกทาง จำนวน 40 คน คิดเป็นร้อยละ 0.66 ตามลำดับ

จำแนกตามอาชีพ พบว่า ส่วนใหญ่ประกอบอาชีพพนักงานบริษัท จำนวน 2,749 คน คิดเป็นร้อยละ 45.63 รองลงมาเป็นนักเรียน/นักศึกษา จำนวน 1,408 คน คิดเป็นร้อยละ 23.37 อันดับที่ 3 ธุรกิจส่วนตัว จำนวน 1,024 คน คิดเป็นร้อยละ 17.00 อันดับที่ 4 ข้าราชการ จำนวน 692 คน คิดเป็นร้อยละ 11.49 อันดับที่ 5 อาชีพอิสระ จำนวน 132 คน คิดเป็นร้อยละ 2.19 และน้อยที่สุดว่างงาน จำนวน 20 คน คิดเป็นร้อยละ 0.33 ตามลำดับ

จำแนกตามรายได้ต่อเดือน พบว่า ส่วนใหญ่มีรายได้ระหว่าง 15,000-30,000 บาท จำนวน 3,397 คน คิดเป็นร้อยละ 56.38 รองลงมามีรายได้น้อยกว่า 15,000 บาท จำนวน 1,401 คน คิดเป็นร้อยละ 23.25 อันดับที่ 3 มีรายได้ระหว่าง 30,001-50,000 บาท จำนวน 844 คน คิดเป็นร้อยละ 14.01 และน้อยที่สุดมีรายได้มากกว่า 50,000 บาท จำนวน 383 คน คิดเป็นร้อยละ 6.36 ตามลำดับ

4.2 การวิเคราะห์พฤติกรรมการใช้บริการของกลุ่มตัวอย่าง

ตารางที่ 4-1 จำนวนและร้อยละของพฤติกรรมการใช้บริการของกลุ่มตัวอย่าง

พฤติกรรมการใช้บริการ	กลุ่มตัวอย่าง (n=6,025)	
	จำนวน (คน)	ร้อยละ
1. มาร้านนี้บ่อยแค่ไหน		
ครั้งแรก	1,115	18.51%
2 - 3 ครั้ง / เดือน	2,761	45.83%
1 ครั้ง/สัปดาห์	808	13.41%
2-3 ครั้ง/สัปดาห์	845	14.02%
เกือบทุกวัน	349	5.79%
มากกว่า 1 ครั้ง / เดือน	147	2.44%
รวม	**6,025**	**100.0%**
2. ช่วงเวลาที่มาบ่อยที่สุด		
17:00-19:00	56	0.93%
19:00-21:00	1,744	28.95%
21:00-23:00	3,769	62.56%
23:00-01:00	456	7.57%
รวม	**6,025**	**100.0%**
3. วันที่มาบ่อยที่สุด		
วันธรรมดา	3,676	61.01%
วันเสาร์อาทิตย์	2,349	38.99%
รวม	**6,025**	**100.0%**

ตารางที่ 4-2 (ต่อ)

พฤติกรรมการใช้บริการ	กลุ่มตัวอย่าง (n=6,025)	
	จำนวน (คน)	ร้อยละ
4. มาที่ร้านกับใคร		
คนเดียว	196	3.25%
เพื่อน 2-3 คน	2,964	49.20%
กลุ่มเพื่อน มากกว่า 4 คน	2,657	44.10%
แฟน/คู่รัก	130	2.16%
เพื่อนร่วมงาน	78	1.29%
รวม	**6,025**	**100.0%**
5. ค่าใช้จ่ายเฉลี่ยต่อครั้ง		
น้อยกว่า 500 บาท	431	7.15%
500-1,000 บาท	2,005	33.28%
1,001-2,000 บาท	1,839	30.52%
2,001-3,000 บาท	1,068	17.73%
> 3,000 บาท	682	11.32%
รวม	**6,025**	**100.0%**
6. เคยใช้บริการ ร้านอื่นหรือไม่		
ไม่เคย	1,328	22.04%
1 ร้าน	876	14.54%
2-3 ร้าน	2,432	40.37%
4-5 ร้าน	531	8.81%
มากกว่า 5 ร้าน	858	14.24%
รวม	**6,025**	**100.0%**
7. ช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร		
06.00 - 12.00	436	7.24%
12.01 – 18.00	1,573	26.11%
18.01 – 24.00	2,939	48.78%
ไม่ต้องการ	1,077	17.88%
รวม	**6,025**	**100.0%**

จากตารางที่ 4-2 พบว่า พฤติกรรมการใช้บริการ จำแนกตามความถี่ในการมาร้าน พบว่า ส่วนใหญ่มาร้านนี้ 2-3 ครั้ง/เดือน จำนวน 2,761 คน คิดเป็นร้อยละ 45.83 รองลงมามาครั้งแรก จำนวน 1,115 คน คิดเป็นร้อยละ 18.51 อันดับที่ 3 มา 2-3 ครั้ง/สัปดาห์ จำนวน 845 คน คิดเป็นร้อยละ 14.02 อันดับที่ 4 มา 1 ครั้ง/สัปดาห์ จำนวน 808 คน คิดเป็นร้อยละ 13.41 อันดับที่ 5 มาเกือบทุกวัน จำนวน 349 คน คิดเป็นร้อยละ 5.79 และน้อยที่สุด มามากกว่า 1 ครั้ง/เดือน จำนวน 147 คน คิดเป็นร้อยละ 2.44 ตามลำดับ

จำแนกตามช่วงเวลาที่มาบ่อยที่สุด พบว่า ส่วนใหญ่มาช่วงเวลา 21:00-23:00 น. จำนวน 3,769 คน คิดเป็นร้อยละ 62.56 รองลงมามาช่วงเวลา 19:00-21:00 น. จำนวน 1,744 คน คิดเป็นร้อยละ 28.95 อันดับที่ 3 มาช่วงเวลา 23:00-01:00 น. จำนวน 456 คน คิดเป็นร้อยละ 7.57 และน้อยที่สุดมาช่วงเวลา 17:00-19:00 น. จำนวน 56 คน คิดเป็นร้อยละ 0.93 ตามลำดับ

จำแนกตามวันที่มาบ่อยที่สุด พบว่า ส่วนใหญ่มาในวันธรรมดา จำนวน 3,676 คน คิดเป็นร้อยละ 61.01 รองลงมามาในวันเสาร์อาทิตย์ จำนวน 2,349 คน คิดเป็นร้อยละ 38.99 ตามลำดับ

จำแนกตามการมาที่ร้านกับใคร พบว่า ส่วนใหญ่มากับเพื่อน 2-3 คน จำนวน 2,964 คน คิดเป็นร้อยละ 49.20 รองลงมามากับกลุ่มเพื่อนมากกว่า 4 คน จำนวน 2,657 คน คิดเป็นร้อยละ 44.10 อันดับที่ 3 มาคนเดียว จำนวน 196 คน คิดเป็นร้อยละ 3.25 อันดับที่ 4 มากับแฟน/คู่รัก จำนวน 130 คน คิดเป็นร้อยละ 2.16 และน้อยที่สุดมากับเพื่อนร่วมงาน จำนวน 78 คน คิดเป็นร้อยละ 1.29 ตามลำดับ

จำแนกตามค่าใช้จ่ายเฉลี่ยต่อครั้ง พบว่า ส่วนใหญ่มีค่าใช้จ่ายระหว่าง 500-1,000 บาท จำนวน 2,005 คน คิดเป็นร้อยละ 33.28 รองลงมามีค่าใช้จ่ายระหว่าง 1,001-2,000 บาท จำนวน 1,839 คน คิดเป็นร้อยละ 30.52 อันดับที่ 3 มีค่าใช้จ่ายระหว่าง 2,001-3,000 บาท จำนวน 1,068 คน คิดเป็นร้อยละ 17.73 อันดับที่ 4 มีค่าใช้จ่ายมากกว่า 3,000 บาท จำนวน 682 คน คิดเป็นร้อยละ 11.32 และน้อยที่สุดมีค่าใช้จ่ายน้อยกว่า 500 บาท จำนวน 431 คน คิดเป็นร้อยละ 7.15 ตามลำดับ

จำแนกตามการเคยใช้บริการร้านอื่นหรือไม่ พบว่า ส่วนใหญ่เคยใช้บริการ 2-3 ร้าน จำนวน 2,432 คน คิดเป็นร้อยละ 40.37 รองลงมาไม่เคยใช้บริการร้านอื่น จำนวน 1,328 คน คิดเป็นร้อยละ 22.04 อันดับที่ 3 เคยใช้บริการ 1 ร้าน จำนวน 876 คน คิดเป็นร้อยละ 14.54 อันดับที่ 4 เคยใช้บริการมากกว่า 5 ร้าน จำนวน 858 คน คิดเป็นร้อยละ 14.24 และน้อยที่สุดเคยใช้บริการ 4-5 ร้าน จำนวน 531 คน คิดเป็นร้อยละ 8.81 ตามลำดับ

จำแนกตามช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร พบว่า ส่วนใหญ่ต้องการรับข้อมูลช่วงเวลา 18:01-24:00 น. จำนวน 2,939 คน คิดเป็นร้อยละ 48.78 รองลงมาต้องการรับข้อมูลช่วงเวลา

12:01-18:00 น. จำนวน 1,573 คน คิดเป็นร้อยละ 26.11 อันดับที่ 3 ไม่ต้องการรับข้อมูล จำนวน 1,077 คน คิดเป็นร้อยละ 17.88 และน้อยที่สุดต้องการรับข้อมูลช่วงเวลา 06:00-12:00 น. จำนวน 436 คน คิดเป็นร้อยละ 7.24 ตามลำดับ

4.3 การวิเคราะห์ประสิทธิภาพด้าน Silhouette Score

4.3.1 การวิเคราะห์ประสิทธิภาพด้าน Silhouette Score ทุกกลุ่มตัวอย่าง การวิเคราะห์คะแนน Silhouette ประเมินคุณภาพการจัดกลุ่มผ่านความสอดคล้องภายในกลุ่มและการแยกระหว่างกลุ่มของลูกค้าร้านอาหารในช่วงเวลากลางคืนจำนวน 6,025 ราย DBSCAN ได้คะแนน ที่สมบูรณ์แบบ (1.0) โดยมีการแยกที่สมบูรณ์และความสอดคล้องสูงสุด บ่งชี้ว่ามีกลุ่มที่แตกต่างกันอย่างชัดเจนโดยไม่มีการจัดกลุ่มผิดหรือความซ้อนทับ GMM ได้คะแนน (0.4704) แสดงให้เห็นคุณภาพที่ดีพร้อมขอบเขตที่ชัดเจนและความคล้ายคลึงภายในกลุ่มที่แข็งแกร่งผ่านการกำหนดแบบความน่าจะเป็น K-Means ได้คะแนนปานกลาง (0.4543) โดยมีการแยกที่ยอมรับได้ แต่มีความซ้อนทับบางส่วนระหว่างกลุ่มที่อยู่ติดกัน บ่งชี้ว่ามีกรณีที่อยู่บนขอบเขตซึ่งเชื่อมโยงหลายกลุ่มเข้าด้วยกัน QML แสดงประสิทธิภาพที่แย่ที่สุด (0.4090) โดยมีความซ้อนทับที่มากและการแยก ที่ไม่ชัดเจน ซึ่งจุดบนขอบเขตมักจะอยู่ใกล้กับกลุ่มข้างเคียงมากกว่าจุดศูนย์กลางที่ถูกกำหนดให้สะท้อนให้เห็นถึงความท้าทายในแนวทางการจัดกลุ่มแบบควอนตัม ดังตารางที่ 4-3

ตารางที่ 4-2 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Silhouette Score ทุกกลุ่มตัวอย่าง

เทคนิค	K-Means	GMM	DBSCAN	QML
Silhouette Score	0.4543	0.4704	1.0000	0.4090

4.3.2 การวิเคราะห์ประสิทธิภาพด้าน Silhouette Score ตามขนาดกลุ่มตัวอย่าง เพื่อประเมินความมั่นคงและความสามารถในการปรับขนาดของแต่ละเทคนิคการแบ่งกลุ่ม ผู้วิจัยได้ทำการทดสอบประสิทธิภาพกับข้อมูลขนาดต่างๆ 8 ระดับ ได้แก่ 300, 500, 1,000, 2,000, 3,000, 4,000, 5,000 และ 6,025 ราย โดยวัดด้วย Silhouette Score ซึ่งแสดงผลดังตารางที่ 4-4

ตารางที่ 4-3 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Silhouette Score ตามขนาดกลุ่มตัวอย่าง

Sample Size	K-Means	GMM	DBSCAN	QML
300	**0.652**	**0.6416**	1.0000	**0.4849**
500	0.5289	0.4971	1.0000	0.5021
1,000	0.4433	0.4828	1.0000	0.4936
2,000	0.4372	0.4417	1.0000	0.4547
3,000	0.4285	0.5087	1.0000	0.3868
4,000	0.4195	0.4145	1.0000	0.3703
5,000	0.4106	0.4106	1.0000	0.4408
6,025	**0.4543**	**0.4704**	1.0000	**0.4090**

จากตารางที่ 4-4 พบว่า K-Means และ GMM มีประสิทธิภาพสูงสุดอย่างสม่ำเสมอ ในทุกขนาดของกลุ่มตัวอย่าง โดย K-Means ให้คะแนนสูงสุดที่ 0.652 เมื่อใช้ข้อมูล 300 ราย และลดลงเหลือ 0.4543 เมื่อเพิ่มขนาดข้อมูลเป็น 6,025 ราย ในทำนองเดียวกัน GMM ให้คะแนน 0.6416 ที่ขนาดข้อมูล 300 ราย และลดลงเหลือ 0.4704 ที่ขนาดข้อมูล 6,025 ราย การลดลง ของคะแนนเมื่อข้อมูลเพิ่มขึ้นนี้สะท้อนให้เห็นถึงความซับซ้อนที่เพิ่มขึ้นของข้อมูลและความทับซ้อนระหว่างกลุ่มที่มากขึ้น

สำหรับเทคนิคแบบ QML แสดงประสิทธิภาพที่ต่ำกว่าเทคนิคแบบคลาสสิกอย่างชัดเจน โดย QML ให้คะแนน 0.4849 ที่ 300 ราย และลดลงเหลือ 0.4090 ที่ 6,025 ราย QML แสดงความไม่มั่นคงในการจัดกลุ่มที่เพิ่มขึ้นเมื่อขนาดข้อมูลเพิ่มขึ้น สะท้อนให้เห็นถึงความท้าทายในการปรับขนาดของอัลกอริทึมควอนตัมบนระบบจำลองแบบคลาสสิก

การวิเคราะห์แนวโน้มพบว่าเทคนิคทั้งหมดแสดงการลดลงของคะแนน Silhouette Score เมื่อขนาดข้อมูลเพิ่มขึ้นจาก 300 เป็น 6,025 ราย การลดลงนี้เป็นธรรมชาติของข้อมูลที่มีความซับซ้อนและความหลากหลายมากขึ้นเมื่อขนาดเพิ่มขึ้น ทำให้ขอบเขตระหว่างกลุ่มไม่ชัดเจนเท่าเดิม

ผลการทดสอบนี้ยืนยันว่า K-Means และ GMM เป็นตัวเลือกที่ดีที่สุดสำหรับ การแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน โดยให้ผลลัพธ์คุณภาพสูงอย่างสม่ำเสมอทั้งในข้อมูลขนาดเล็กและใหญ่ ส่วนเทคนิคแบบควอนตัมยังไม่พร้อมสำหรับการใช้งานจริงในปัจจุบัน เนื่องจากให้คุณภาพการจัดกลุ่มที่ต่ำกว่าและมีความไม่มั่นคงสูงกว่าเทคนิคแบบคลาสสิก

4.4 การวิเคราะห์ประสิทธิภาพด้าน Davies-Bouldin Index

4.4.1 การวิเคราะห์ประสิทธิภาพด้าน Davies-Bouldin Index ทุกกลุ่มตัวอย่าง การวิเคราะห์ดัชนี Davies-Bouldin (DBI) ประเมินคุณภาพการจัดกลุ่มผ่านอัตราส่วนความคล้ายคลึงเฉลี่ย ระหว่างกลุ่มต่างๆ กับกลุ่มใกล้เคียงที่สุด โดยพิจารณาจากการกระจายภายในกลุ่มและการแยกกันระหว่างกลุ่ม ค่าที่ต่ำกว่าบ่งชี้ประสิทธิภาพที่ดีกว่า DBSCAN ได้ DBI ที่เหมาะสมที่สุด คือ 0 ซึ่งแสดงถึงการแยกที่สมบูรณ์แบบโดยไม่มีความคล้ายคลึงระหว่างกลุ่ม - มีระยะห่างสูงสุดระหว่างกลุ่มพร้อมกับการกระจายภายในที่น้อยที่สุด สร้างส่วนที่แยกออกจากกันอย่างสมบูรณ์โดยไม่มีความคลุมเครือของขอบเขต K-Means ได้ผลลัพธ์ (0.9008) บ่งชี้ว่ามีคุณภาพปานกลาง โดยกลุ่มต่างๆ ยังคงแยกจากกันได้อย่างสมเหตุสมผล แต่บางส่วนแสดงลักษณะที่คล้ายคลึงกับกลุ่มใกล้เคียง สะท้อนให้เห็นถึงความท้าทายของการสมมติฐานแบบทรงกลมกับรูปแบบพฤติกรรมที่ซับซ้อน GMM บันทึกได้ (0.9014) บ่งชี้ประสิทธิภาพที่ดี

มีกลุ่มที่แยกจากกันอย่างชัดเจนและมีความคล้ายคลึงกับกลุ่มใกล้เคียงที่จำกัดรักษาขอบเขตที่ชัดเจนอย่างมีประสิทธิภาพในขณะที่จัดการกับความซ้อนทับแบบความน่าจะเป็นอย่างเหมาะสม QML แสดงประสิทธิภาพที่แย่ที่สุด (1.1491) บ่งชี้ความซ้อนทับที่มากและการแยกที่ไม่ดี โดยมีการกระจายภายในกลุ่มสูงเมื่อเทียบกับระยะห่างระหว่างกลุ่ม ส่งผลให้ขอบเขตคลุมเครือ ดังตารางที่ 4-5

ตารางที่ 4-4 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Davies-Bouldin Index ทุกกลุ่มตัวอย่าง

เทคนิค	K-Means	GMM	DBSCAN	QML
Davies-Bouldin Index	0.9008	0.9014	0.0000	1.1491

4.4.2 การวิเคราะห์ประสิทธิภาพด้าน Davies-Bouldin Index ตามขนาดกลุ่มตัวอย่าง เพื่อประเมินความสามารถในการแยกกลุ่มและความคล้ายคลึงระหว่างกลุ่มในข้อมูลขนาดต่างๆ ผู้วิจัยได้ทดสอบประสิทธิภาพด้าน Davies-Bouldin Index กับข้อมูล 8 ขนาด ได้แก่ 300, 500, 1,000, 2,000, 3,000, 4,000, 5,000 และ 6,025 ราย โดย Davies-Bouldin Index เป็นเมตริกที่ค่าต่ำ แสดงถึงการแยกกลุ่มที่ดีกว่า ผลการทดสอบแสดงดังตารางที่ 4-6

ตารางที่ 4-5 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Davies-Bouldin Index ตามขนาดกลุ่มตัวอย่าง

Sample Size	K-Means	GMM	DBSCAN	QML
300	**0.6507**	**0.7020**	0.0000	**0.8360**
500	0.7333	0.9001	0.0000	0.9931
1,000	0.8192	0.9513	0.0000	1.1352
2,000	0.8765	0.9641	0.0000	1.0952
3,000	0.8544	1.0346	0.0000	1.1136
4,000	0.9124	0.9095	0.0000	1.0316
5,000	0.8699	0.8699	0.0000	1.0062
6,025	**0.9008**	**0.9014**	0.0000	**1.1491**

จากตารางที่ 4-6 พบว่า K-Means และ GMM ให้ผลลัพธ์ที่ดีที่สุดอย่างสม่ำเสมอ ในทุกขนาดของกลุ่มตัวอย่าง โดยมีค่า Davies-Bouldin Index ที่ต่ำที่สุด ซึ่งบ่งชี้ว่ากลุ่มที่ได้มีการแยกจากกันได้ดีและมีความคล้ายคลึงภายในกลุ่มสูง K-Means ให้ค่า 0.6507 เมื่อใช้ข้อมูล 300 ราย และเพิ่มขึ้นเป็น 0.9008 เมื่อขนาดข้อมูลเพิ่มเป็น 6,025 ราย ในทำนองเดียวกัน GMM ให้ค่า 0.7020 ที่ขนาดข้อมูล 300 ราย และเพิ่มขึ้นเป็น 0.9014 ที่ขนาดข้อมูล 6,025 ราย การเพิ่มขึ้นของค่า Davies-Bouldin Index เมื่อข้อมูลเพิ่มขึ้นเป็นธรรมชาติของข้อมูลที่มีความซับซ้อนมากขึ้น ทำให้กลุ่มต่างๆ มีความทับซ้อนกันมากขึ้น การวิเคราะห์เปรียบเทียบระหว่าง K-Means และ GMM พบว่าทั้งสองเทคนิคให้ผลลัพธ์ที่ใกล้เคียงกันมาก โดย K-Means มีข้อได้เปรียบเล็กน้อยในข้อมูลขนาดเล็ก (300-2,000 ราย) โดยให้ค่า Davies-Bouldin Index ที่ต่ำกว่า GMM ประมาณ 0.05 แต่เมื่อขนาดข้อมูลเพิ่มขึ้นเป็น 6,025 ราย ทั้งสองเทคนิคให้ผลลัพธ์ที่ใกล้คียงกัน แสดงให้เห็นว่าทั้งสองเทคนิคมีความสามารถในการแยกกลุ่มที่ใกล้เคียงกันในข้อมูลขนาดใหญ่

สำหรับ QML พบว่าแสดงประสิทธิภาพที่แย่ที่สุดอย่างชัดเจน โดยมีค่า Davies-Bouldin Index สูงที่สุดในทุกขนาดของกลุ่มตัวอย่าง QML ให้ค่า 0.8360 เมื่อใช้ข้อมูล 300 ราย และเพิ่มขึ้นเป็น 1.1491 ที่ขนาดข้อมูล 6,025 ราย ค่าที่สูงกว่า 1.0 บ่งชี้ว่ากลุ่มที่ได้มีความทับซ้อนกันอย่างมาก โดยมีการกระจายภายในกลุ่มที่มากกว่าระยะห่างระหว่างจุดศูนย์กลางของกลุ่ม ทำให้ขอบเขตระหว่างกลุ่มไม่ชัดเจน การเพิ่มขึ้นของค่า Davies-Bouldin Index ของ QML จาก 0.8360 เป็น 1.1491 (เพิ่มขึ้น 37%) สูงกว่าเทคนิคอื่นอย่างมีนัยสำคัญ แสดงให้เห็นถึงความไม่มั่นคงและความท้าทายในการจัดการกับข้อมูลขนาดใหญ่ของอัลกอริทึมควอนตัม

การวิเคราะห์แนวโน้มโดยรวมพบว่าเทคนิคทั้งหมดแสดงการเพิ่มขึ้นของค่า Davies-Bouldin Index เมื่อขนาดข้อมูลเพิ่มขึ้น โดยอัตราการเพิ่มขึ้นของ K-Means และ GMM อยู่ที่ประมาณ 28-38% ซึ่งถือว่าอยู่ในระดับที่ยอมรับได้และสะท้อนให้เห็นถึงความซับซ้อนที่เพิ่มขึ้นตามธรรมชาติของข้อมูลขนาดใหญ่ ในขณะที่ QML มีอัตราการเพิ่มขึ้นสูงถึง 37% และยังให้ค่าสูงเกินมาตรฐานที่ยอมรับได้ (>1.0) แสดงให้เห็นถึงข้อจำกัดที่สำคัญของเทคนิคนี้

ผลการทดสอบนี้ยืนยันว่า K-Means และ GMM เป็นตัวเลือกที่เหมาะสมที่สุดสำหรับการแบ่งกลุ่มลูกค้า โดยสามารถสร้างกลุ่มที่แยกจากกันได้ดีและมีความสอดคล้องภายในกลุ่มสูงทั้งในข้อมูลขนาดเล็กและใหญ่ ส่วน QML ยังไม่เหมาะสำหรับการใช้งานจริง เนื่องจากให้กลุ่มที่มีความทับซ้อนกันสูงและมีความไม่มั่นคงเมื่อขนาดข้อมูลเพิ่มขึ้น โดยเฉพาะในข้อมูลขนาดใหญ่ที่ค่า Davies-Bouldin Index สูงเกินกว่าเกณฑ์ที่ยอมรับได้สำหรับการนำไปใช้ในทางปฏิบัติ

4.5 การวิเคราะห์ประสิทธิภาพด้าน Cluster Cohesion

4.5.1 การวิเคราะห์ประสิทธิภาพด้าน Cluster Cohesion ทุกกลุ่มตัวอย่าง การวิเคราะห์ความสอดคล้องของกลุ่ม (Cluster Cohesion Analysis) วัดความกระชับของกลุ่มผ่านผลรวมของระยะทางกำลังสองระหว่างจุดข้อมูลกับจุดศูนย์กลาง ค่าที่ต่ำกว่าบ่งชี้ว่ากลุ่มมีความกระชับและแน่นหนามากขึ้นพร้อมคุณภาพที่ดีกว่า DBSCAN ได้ค่าความสอดคล้อง 0 ซึ่งสะท้อนให้เห็นถึงแนวทางที่ใช้ความหนาแน่นเป็นฐานของอัลกอริทึมที่ไม่ใช้จุดศูนย์กลางในการสร้างกลุ่มต่างจากวิธีที่ใช้จุดศูนย์กลางเป็นฐาน DBSCAN ระบุกลุ่มผ่านการเชื่อมต่อของความหนาแน่นทำให้การวัดความสอดคล้องแบบดั้งเดิมไม่สามารถนำมาใช้กับโครงสร้างการจัดกลุ่มของมันได้ GMM ได้ค่า Cluster Cohesion 2.2171 แสดงให้เห็นความกระชับที่ดีโดยจุดต่างๆ อยู่ใกล้กับจุดศูนย์กลางค่อนข้างมาก สร้างส่วนที่กระชับอย่างสมเหตุสมผลในขณะที่รองรับการกำหนดแบบความน่าจะเป็นและรักษาความยืดหยุ่นสำหรับลักษณะที่ซ้อนทับกัน K-Means ได้ค่า Cluster Cohesion 2.3393 บ่งชี้ว่ามีความกระชับปานกลางโดยมีการกระจายที่มากขึ้นรอบจุดศูนย์กลาง เมื่อเทียบกับ GMM และ DBSCAN สะท้อนให้เห็นถึงความท้าทายในการรักษาขอบเขตที่แน่นหนา เมื่อลักษณะของลูกค้าแสดงความแปรปรวนและความซ้อนทับตามธรรมชาติ QML แสดงความสอดคล้องที่แย่ที่สุดที่ 2.5421 บ่งชี้ว่ากลุ่มหลวมและไม่กระชับโดยมีการกระจายที่มาจากจุดศูนย์กลาง ขาดความสม่ำเสมอภายในโดยมีจุดที่กระจายอย่างกว้างขวาง ส่งผลให้ส่วนต่างๆ ไม่เป็นเนื้อเดียวกันมีโครงสร้างที่อ่อนแอดังตารางที่ 4-7

ตารางที่ 4-6 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Cluster Cohesion ทุกกลุ่มตัวอย่าง

เทคนิค	K-Means	GMM	DBSCAN	QML
Cluster Cohesion	2.3393	2.2171	0.0000	2.5421

4.5.2 การวิเคราะห์ประสิทธิภาพด้าน Cluster Cohesion ตามขนาดกลุ่มตัวอย่าง เพื่อประเมินความกระชับและความเป็นเนื้อเดียวกันของกลุ่มที่ได้จากแต่ละเทคนิคในข้อมูลขนาดต่างๆ ผู้วิจัยได้ทดสอบประสิทธิภาพด้าน Cluster Cohesion กับข้อมูล 8 ขนาด ได้แก่ 300, 500, 1,000, 2,000, 3,000, 4,000, 5,000 และ 6,025 ราย โดย Cluster Cohesion วัดผลรวมของระยะทางกำลังสอง

ระหว่างจุดข้อมูลกับจุดศูนย์กลางของกลุ่ม ซึ่งค่าที่ต่ำบ่งชี้ว่ากลุ่มมีความกระชับและแน่นหนามากกว่า แสดงดังตารางที่ 4-8

ตารางที่ 4-7 ประสิทธิภาพของอัลกอริทึมการจัดกลุ่มด้วยค่า Cluster Cohesion ตามขนาดกลุ่มตัวอย่าง

Sample Size	K-Means	GMM	DBSCAN	QML
300	**1.2309**	**1.2248**	0.0000	**2.1783**
500	1.9054	1.9607	0.0000	1.7540
1,000	2.2401	1.9021	0.0000	1.9753
2,000	2.3320	2.3367	0.0000	2.1295
3,000	2.2484	1.9917	0.0000	2.5327
4,000	2.4051	2.3058	0.0000	2.7113
5,000	2.3888	2.3888	0.0000	2.3752
6,025	**2.3393**	**2.2171**	0.0000	**2.5421**

จากตารางที่ 4-8 พบว่า K-Means และ GMM แสดงประสิทธิภาพที่โดดเด่น ในการสร้างกลุ่มที่กระชับ โดยเริ่มต้นด้วยค่า Cluster Cohesion ที่ต่ำมากในข้อมูลขนาดเล็ก K-Means ให้ค่า 1.2309 เมื่อใช้ข้อมูล 300 ราย และเพิ่มขึ้นเป็น 2.3393 เมื่อขนาดข้อมูลเพิ่มเป็น 6,025 ราย ในทำนองเดียวกัน GMM ให้ค่า 1.2248 ที่ขนาดข้อมูล 300 ราย และเพิ่มขึ้นเป็น 2.2171 ที่ขนาดข้อมูล 6,025 ราย การเพิ่มขึ้นของค่า Cluster Cohesion เมื่อข้อมูลเพิ่มขึ้นเป็นผลจาก ความหลากหลายและความซับซ้อนที่เพิ่มขึ้นของข้อมูล ทำให้จุดข้อมูลภายในกลุ่มมีการกระจายตัวมากขึ้นจากจุดศูนย์กลาง

การวิเคราะห์เปรียบเทียบระหว่าง K-Means และ GMM พบว่า GMM มีข้อได้เปรียบเล็กน้อยในการรักษาความกระชับของกลุ่ม โดยให้ค่า Cluster Cohesion ที่ต่ำกว่า K-Means อย่างสม่ำเสมอ

ในทุกขนาดของกลุ่มตัวอย่าง ที่ขนาดข้อมูล 6,025 ราย GMM ให้ค่า 2.2171 ในขณะที่ K-Means ให้ค่า 2.3393 ซึ่งต่ำกว่าประมาณ 5% ความแตกต่างนี้แสดงให้เห็นว่า GMM สามารถสร้างกลุ่มที่มีจุดข้อมูลอยู่ใกล้จุดศูนย์กลางมากกว่าเล็กน้อย เนื่องจากแนวทางแบบความน่าจะเป็นของ GMM ที่สามารถปรับรูปร่างและขนาดของกลุ่มได้ผ่าน covariance matrix

แนวโน้มการเปลี่ยนแปลงของ K-Means และ GMM แสดงให้เห็นว่าทั้งสองเทคนิคเริ่มต้นด้วยความกระชับสูงมากในข้อมูลขนาดเล็ก (ค่าประมาณ 1.2248-1.2309) แต่ค่อยๆ เพิ่มขึ้นเมื่อข้อมูลเพิ่มขึ้น โดยเฉพาะอย่างยิ่งในช่วงข้อมูล 2,000-6,025 ราย ซึ่งค่า Cluster Cohesion เพิ่มขึ้นอย่างเห็นได้ชัด แต่ยังคงอยู่ในระดับที่ยอมรับได้ (ต่ำกว่า 2.5) อัตราการเพิ่มขึ้นโดยรวมสำหรับทั้งสองเทคนิคอยู่ที่ประมาณ 80-90% เมื่อเทียบระหว่างข้อมูล 300 ราย กับ 6,025 ราย ซึ่งสะท้อนให้เห็นถึงผลกระทบของขนาดข้อมูลต่อการกระจายตัวภายในกลุ่ม

สำหรับ QML พบว่าแสดงประสิทธิภาพที่แย่ที่สุดในด้านความกระชับของกลุ่ม โดยมีค่า Cluster Cohesion สูงที่สุดในทุกขนาดของกลุ่มตัวอย่าง QML ให้ค่า 2.1783 เมื่อใช้ข้อมูล 300 ราย ซึ่งสูงกว่า K-Means และ GMM เกือบเท่าตัว และเพิ่มขึ้นเป็น 2.5421 ที่ขนาดข้อมูล 6,025 ราย ค่าที่สูงนี้บ่งชี้ว่ากลุ่มที่ได้จาก QML มีการกระจายตัวมาก โดยจุดข้อมูลอยู่ห่างจากจุดศูนย์กลางของกลุ่มมากกว่าเทคนิคอื่น ทำให้กลุ่มมีความหลวมและไม่เป็นเนื้อเดียวกัน

การวิเคราะห์ความผันผวนของ QML พบว่ามีการเปลี่ยนแปลงที่ไม่สม่ำเสมอ ตลอดช่วงขนาดข้อมูลต่างๆ โดยค่า Cluster Cohesion มีการขึ้นลงแบบไม่เป็นเส้นตรง แสดงให้เห็นถึงความไม่มั่นคงในการสร้างกลุ่ม อัตราการเพิ่มขึ้นโดยรวมของ QML จาก 2.1783 เป็น 2.5421 (เพิ่มขึ้น 16.5%) แม้จะดูน้อยกว่า K-Means และ GMM แต่เมื่อพิจารณาจากค่าเริ่มต้นที่สูงอยู่แล้ว การเพิ่มขึ้นนี้ทำให้ความกระชับของกลุ่มแย่ลงอย่างมีนัยสำคัญ โดยเฉพาะเมื่อเปรียบเทียบกับเทคนิคแบบคลาสสิก

ผลการทดสอบนี้ยืนยันว่า K-Means และ GMM สามารถสร้างกลุ่มที่มีความกระชับดีในทุกขนาดของข้อมูล โดย GMM มีข้อได้เปรียบเล็กน้อยในการรักษาความเป็นเนื้อเดียวกันของกลุ่ม แม้ว่าทั้งสองเทคนิคจะแสดงการเพิ่มขึ้นของค่า Cluster Cohesion เมื่อข้อมูลเพิ่มขึ้น แต่ยังคงอยู่ในระดับที่ยอมรับได้และสามารถนำไปใช้งานได้จริง ในทางตรงกันข้าม QML แสดงให้เห็นถึงข้อจำกัดที่สำคัญในการสร้างกลุ่มที่กระชับ โดยให้กลุ่มที่หลวมและมีการกระจายตัวสูงตั้งแต่ข้อมูลขนาดเล็ก และยังคงแย่ลงเมื่อข้อมูลเพิ่มขึ้น ทำให้ไม่เหมาะสมสำหรับการนำไปใช้ในการแบ่งกลุ่มลูกค้าที่ต้องการความชัดเจนและความเป็นเนื้อเดียวกันของกลุ่ม

4.6 การวิเคราะห์ประสิทธิภาพด้าน Execution Time

4.6.1 การวิเคราะห์ประสิทธิภาพด้าน Execution Time ทุกกลุ่มตัวอย่าง การวิเคราะห์เวลาในการประมวลผล วัดประสิทธิภาพการคำนวณสำหรับการจัดกลุ่มลูกค้า โดยให้ข้อมูลเชิงลึกเกี่ยวกับความเป็นไปได้ในทางปฏิบัติและความสามารถในการขยายขนาดสำหรับการประยุกต์ใช้ ทางธุรกิจ DBSCAN และ K-Means แสดงประสิทธิภาพที่โดดเด่นด้วยเวลาประมวลผลที่รวดเร็วและเกือบเท่ากัน เน้นย้ำถึงความเหมาะสมของทั้งสองวิธีสำหรับการแบ่งกลุ่มแบบเรียลไทม์และชุดข้อมูลขนาดใหญ่ด้วยภาระการคำนวณที่น้อยที่สุด ทำให้เหมาะสมอย่างยิ่งสำหรับสภาพแวดล้อมการดำเนินงานที่ต้องการการวิเคราะห์ที่รวดเร็วและข้อมูลเชิงลึกในการตัดสินใจทันที GMM บันทึกเวลาประมวลผลปานกลาง โดยทำงานช้ากว่าอัลกอริทึมแบบดั้งเดิมหลายสิบเท่า แต่ยังคงอยู่ในระดับที่ยอมรับได้สำหรับการใช้งานส่วนใหญ่

ต้นทุนการคำนวณที่เพิ่มขึ้นนี้สะท้อนให้เห็นถึงการสร้างแบบจำลองแบบความน่าจะเป็นที่ซับซ้อน ซึ่งคำนวณความน่าจะเป็นของการเป็นสมาชิกข้ามการแจกแจงแบบเกาส์เซียนหลายตัวผ่านกระบวนการ expectation-maximization แบบวนซ้ำ QML แสดงความซับซ้อนในการคำนวณอย่างมากด้วยเวลาประมวลผลที่ช้ากว่าอัลกอริทึมอื่นๆ หลายพันเท่า ภาระที่มากนี้เกิดจากการจำลองวงจรควอนตัม การเตรียมสถานะ การแมปคุณลักษณะ และความซับซ้อนโดยธรรมชาติของการจำลองอัลกอริทึมควอนตัมบนฮาร์ดแวร์แบบคลาสสิก ซึ่งนำเสนอข้อจำกัดทางปฏิบัติที่สำคัญสำหรับการนำไปใช้ในโลกแห่งความเป็นจริงที่ต้องการการแบ่งกลุ่มที่ทันท่วงทีและข่าวกรองธุรกิจแบบไดนามิกดังตารางที่ 4-9

ตารางที่ 4-8 ประสิทธิภาพด้านเวลาของอัลกอริทึมการจัดกลุ่ม ทุกกลุ่มตัวอย่าง

เทคนิค	K-Means	GMM	DBSCAN	QML
Execution Time (second)	0.09	3.60	0.48	2,904

4.6.2 การวิเคราะห์ประสิทธิภาพด้านเวลาของอัลกอริทึมการจัดกลุ่ม ตามขนาดกลุ่มตัวอย่าง เพื่อประเมินประสิทธิภาพด้านเวลาของอัลกอริทึมการจัดกลุ่มในข้อมูลขนาดต่างๆ ผู้วิจัยได้ทดสอบกับข้อมูล 8 ขนาด ได้แก่ 300, 500, 1,000, 2,000, 3,000, 4,000, 5,000 และ 6,025 รายผลการทดสอบแสดงดังตารางที่ 4-10

ตารางที่ 4-9 ประสิทธิภาพด้านเวลาของอัลกอริทึมการจัดกลุ่ม ตามขนาดกลุ่มตัวอย่าง

Sample Size	K-Means	GMM	DBSCAN	QML
300	0.06	**0.51**	**0.01**	**150.00**
500	**0.04**	0.78	0.01	254.40
1,000	0.05	2.63	0.02	470.40
2,000	0.09	4.00	0.04	925.20
3,000	**0.26**	11.84	0.10	1,578.60
4,000	0.08	7.26	0.20	2,038.20
5,000	0.08	7.15	0.38	2,270.40
6,025	**0.09**	**3.60**	**0.48**	**2,904.00**

จากตารางที่ 4-10 พบว่า K-Means และ DBSCAN แสดงประสิทธิภาพด้านความเร็วที่ยอดเยี่ยมอย่างสม่ำเสมอในทุกขนาดของกลุ่มตัวอย่าง โดย K-Means ใช้เวลาในการประมวลผลระหว่าง 0.04-0.26 วินาที และเพิ่มขึ้นเป็น 0.09 วินาทีเมื่อข้อมูลเพิ่มเป็น 6,025 รายในทำนองเดียวกัน DBSCAN ใช้เวลาระหว่าง 0.01-0.48 วินาที แสดงให้เห็นถึงความเหมาะสมสูงสำหรับการประยุกต์ใช้แบบเรียลไทม์และชุดข้อมูลขนาดใหญ่ โดยมีภาระการคำนวณที่น้อยที่สุด

GMM แสดงเวลาในการประมวลผลในระดับปานกลาง โดยเริ่มต้นที่ 0.51 วินาทีสำหรับข้อมูล 300 ราย และเพิ่มขึ้นเป็น 3.60 วินาทีเมื่อข้อมูลเพิ่มเป็น 6,025 ราย แม้จะใช้เวลานานกว่าอัลกอริทึมแบบดั้งเดิมหลายสิบเท่า แต่ยังคงอยู่ในระดับที่ยอมรับได้สำหรับการใช้งานส่วนใหญ่ ต้นทุนการคำนวณที่เพิ่มขึ้นนี้สะท้อนให้เห็นถึงความซับซ้อนของการสร้างแบบจำลองแบบความน่าจะเป็น ซึ่งต้องคำนวณความน่าจะเป็นของการเป็นสมาชิกข้ามการแจกแจงแบบเกาส์เซียนหลายตัวผ่านกระบวนการ expectation-maximization แบบวนซ้ำ

QML แสดงความซับซ้อนในการคำนวณอย่างมาก โดยเริ่มต้นที่ 150.00 วินาที สำหรับข้อมูล 300 ราย และเพิ่มขึ้นอย่างมหาศาลเป็น 2,904.00 วินาที (ประมาณ 48 นาที) เมื่อข้อมูล

เพิ่มเป็น 6,025 ราย ซึ่งช้ากว่าอัลกอริทึมอื่นๆ หลายพันเท่า ภาระที่มากนี้เกิดจากการจำลองวงจรควอนตัม การเตรียมสถานะ การแมปคุณลักษณะ และความซับซ้อนโดยธรรมชาติของการจำลองอัลกอริทึมควอนตัมบนฮาร์ดแวร์แบบคลาสสิก ซึ่งนำเสนอข้อจำกัดทางปฏิบัติที่สำคัญสำหรับการนำไปใช้ในโลกแห่งความเป็นจริงที่ต้องการการแบ่งกลุ่มที่ทันท่วงทีและข่าวกรองธุรกิจแบบไดนามิก

การวิเคราะห์แนวโน้มพบว่าเวลาในการประมวลผลของทุกเทคนิคเพิ่มขึ้นเมื่อขนาดข้อมูลเพิ่มขึ้น แต่อัตราการเพิ่มขึ้นแตกต่างกันอย่างมีนัยสำคัญ K-Means และ DBSCAN แสดงการเพิ่มขึ้นแบบเชิงเส้นที่สามารถคาดการณ์ได้ ในขณะที่ GMM แสดงการเพิ่มขึ้นแบบไม่เป็นเชิงเส้นแต่ยังอยู่ในระดับที่จัดการได้ ส่วน QML แสดงการเพิ่มขึ้นแบบทวีคูณที่ทำให้ไม่เหมาะสมสำหรับข้อมูลขนาดใหญ่

ผลการทดสอบนี้ยืนยันว่า K-Means และ DBSCAN เป็นตัวเลือกที่ดีที่สุดสำหรับงานที่ต้องการความเร็วสูง ในขณะที่ GMM ให้ความสมดุลที่ดีระหว่างความเร็วและคุณภาพการจัดกลุ่ม ส่วน QML ยังไม่เหมาะสำหรับการใช้งานจริงเนื่องจากเวลาในการประมวลผลที่นานเกินไป โดยเฉพาะในข้อมูลขนาดใหญ่ที่อาจใช้เวลาหลายชั่วโมงในการประมวลผล

4.7 การเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี

การวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มลูกค้าทั้ง 4 วิธี ได้แก่ K-Means, DBSCAN, Gaussian Mixture Model (GMM) และ Quantum Machine Learning Clustering (QML) โดยใช้เกณฑ์การประเมิน 4 ด้าน คือ Silhouette Score, Davies-Bouldin Index (DBI), Cluster Cohesion และ Execution Time จากข้อมูลลูกค้าร้านอาหารกลางคืน จำนวน 6,025 ราย

4.7.1 สรุปผลการเปรียบเทียบประสิทธิภาพโดยรวม จากการวิเคราะห์ประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี สามารถสรุปผลการเปรียบเทียบในแต่ละเมตริกได้ดังตารางที่ 4-11

ตารางที่ 4-10 ผลการเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี

เทคนิค	Silhouette Score	Davies – Bouldin Index	Cluster Cohesion	Execution Time (second)	จำนวนกลุ่ม
K-Means	0.4543	**0.9008**	2.3393	**0.09**	10
GMM	**0.4704**	0.9014	**2.2171**	3.60	10
DBSCAN	1.0000	0.0000	0.0000	0.48	54
QML	0.4090	1.1491	2.5421	2904.00	10

จากตารางที่ 4 - 11 พบว่าแต่ละเทคนิคมีข้อได้เปรียบและข้อจำกัดที่แตกต่างกันโดย DBSCAN ให้คะแนนคุณภาพการจัดกลุ่มที่สูงที่สุดแต่สร้างกลุ่มจำนวนมากถึง 54 กลุ่ม ซึ่งมากเกินไปสำหรับการนำไปใช้งานจริง ในขณะที่ K-Means GMM และ QML ให้จำนวนกลุ่ม ที่เหมาะสม (10 กลุ่ม) พร้อมคุณภาพการจัดกลุ่มที่ดี

4.7.2 การวิเคราะห์เปรียบเทียบด้านคุณภาพการจัดกลุ่ม

4.7.2.1 การวิเคราะห์ความสอดคล้องภายในกลุ่มและการแยกระหว่างกลุ่ม จากค่า Silhouette Score ที่วัดความสอดคล้องภายในกลุ่มและการแยกระหว่างกลุ่ม พบว่า DBSCAN ได้คะแนนสูงสุดที่ 1.0 ซึ่งเป็นคะแนนที่สมบูรณ์แบบ แสดงว่ากลุ่มที่ได้มีการแยกออกจากกันอย่างสมบูรณ์โดยไม่มีความซ้อนทับ อย่างไรก็ตาม DBSCAN สร้างกลุ่มจำนวนมากถึง 54 กลุ่ม ซึ่งมากเกินไปสำหรับการนำไปประยุกต์ใช้ในการพัฒนากลยุทธ์การตลาดเฉพาะกลุ่ม การที่มีกลุ่มจำนวนมากเกินไปทำให้ยากต่อการทำความเข้าใจลักษณะของแต่ละกลุ่มและพัฒนากลยุทธ์ที่เหมาะสมสำหรับแต่ละกลุ่ม นอกจากนี้ยังทำให้การจัดสรรทรัพยากรทางการตลาดซับซ้อนและไม่มีประสิทธิภาพ GMM ให้คะแนนสูงสุดในกลุ่มเทคนิคที่มีจำนวนกลุ่มเหมาะสมที่ 0.4704 รองลงมาคือ K-Means ที่ 0.4543 โดยทั้งสองเทคนิคสร้างกลุ่มจำนวน 10 กลุ่ม ซึ่งเป็นจำนวนที่เหมาะสมสำหรับการนำไปใช้งานจริง ความแตกต่างระหว่าง GMM และ K-Means มีเพียง 0.0161 หรือประมาณ 3.5% ซึ่งถือว่าน้อยมาก แสดงว่าทั้งสองเทคนิคให้ผลลัพธ์ที่ใกล้เคียงกันในแง่ของความสอดคล้องภายในกลุ่ม คะแนนที่อยู่ในระดับ 0.45-0.47 นี้ถือว่าอยู่ในระดับปานกลางถึงดี แสดงว่ากลุ่มที่ได้มีความชัดเจนพอสมควร แม้จะยังมีจุดข้อมูลบางส่วนที่อยู่บนขอบเขตระหว่างกลุ่ม แต่ก็ยังสามารถแยกแยะความแตกต่างระหว่างกลุ่มได้ QML ให้คะแนนต่ำที่สุดที่ 0.4090 แม้ว่าจะสร้างกลุ่มจำนวน 10 กลุ่มเช่นเดียวกับ K-Means และ GMM แต่คุณภาพ

การจัดกลุ่มก็ต่ำกว่า โดยต่ำกว่า K-Means ถึง 10% และต่ำกว่า GMM ถึง 13% แสดงว่ากลุ่มที่ได้จาก QML มีความซ้อนทับมากกว่าและมีจุดข้อมูลจำนวนมากที่อาจถูกจัดกลุ่มไม่ถูกต้องหรืออยู่ใกล้กับกลุ่มอื่นมากกว่ากลุ่มที่ตัวเองถูกจัดไว้ การวิเคราะห์เชิงลึกพบว่าแม้ DBSCAN จะให้คะแนน Silhouette ที่สมบูรณ์แบบ แต่การสร้างกลุ่มจำนวน 54 กลุ่มนั้นไม่สามารถนำไปใช้งานจริงได้อย่างมีประสิทธิภาพ การมีกลุ่มมากเกินไปอาจเกิดจากการที่ DBSCAN แบ่งกลุ่มตามความหนาแน่นของข้อมูล อย่างเคร่งครัด ทำให้กลุ่มย่อยเล็กๆ ที่มีลักษณะใกล้เคียงกันถูกแยกออกเป็นกลุ่มต่างหาก ในขณะที่ GMM และ K-Means ที่กำหนดจำนวนกลุ่มไว้ล่วงหน้าที่ 10 กลุ่มสามารถควบคุมจำนวนกลุ่มให้อยู่ในระดับที่เหมาะสมสำหรับการวิเคราะห์และนำไปใช้งาน

4.7.2.2 การวิเคราะห์ความคล้ายคลึงระหว่างกลุ่ม จากค่า Davies-Bouldin Index ที่วัดความคล้ายคลึงระหว่างกลุ่ม โดยค่าที่ต่ำแสดงถึงการแยกกลุ่มที่ดีกว่า พบว่า DBSCAN ได้ค่า DBI เท่ากับ 0 ซึ่งเป็นค่าที่ดีที่สุดที่เป็นไปได้ แสดงว่ากลุ่มที่ได้จาก DBSCAN แยกจากกันอย่างสมบูรณ์โดยไม่มีความคล้ายคลึงระหว่างกลุ่มเลย อย่างไรก็ตาม เมื่อพิจารณาร่วมกับจำนวนกลุ่มที่มากถึง 54 กลุ่ม จะเห็นได้ว่าการแยกกลุ่มที่สมบูรณ์แบบนี้มาพร้อมกับการแบ่งแยกข้อมูลออกเป็นกลุ่มเล็กๆ มากเกินไป ซึ่งอาจไม่เป็นประโยชน์ในการนำไปใช้งานจริง K-Means และ GMM ให้ค่า DBI ที่ใกล้เคียงกันมากคือ 0.9008 และ 0.9014 ตามลำดับ โดยความแตกต่างมีเพียง 0.0006 หรือประมาณ 0.07% ซึ่งถือว่าไม่มีนัยสำคัญ แสดงว่าทั้งสองเทคนิคให้ระดับการแยกกลุ่มที่ใกล้เคียงกันมาก ค่า DBI ที่น้อยกว่า 1 แสดงว่ากลุ่มที่ได้ยังคงแยกจากกันได้ดีพอสมควร แม้ว่าจะไม่สมบูรณ์แบบเหมือน DBSCAN แต่ระดับการแยกกลุ่มนี้เพียงพอสำหรับการสร้างกลุ่มลูกค้าที่มีลักษณะแตกต่างกันอย่างชัดเจน พร้อมทั้งมีจำนวนกลุ่มที่เหมาะสมสำหรับการจัดการและพัฒนากลยุทธ์การตลาด QML ให้ค่า DBI สูงที่สุดที่ 1.1491 ซึ่งสูงกว่า 1 แสดงว่ามีการกระจายภายในกลุ่มที่มากกว่าระยะห่างระหว่างกลุ่ม นั่นคือกลุ่มที่ได้มีความซ้อนทับสูงและขอบเขตระหว่างกลุ่มไม่ชัดเจน ค่า DBI ของ QML สูงกว่า K-Means และ GMM ประมาณ 27-28% ซึ่งถือว่าแตกต่างอย่างมีนัยสำคัญ แม้ว่า QML จะสร้างกลุ่มจำนวนเท่ากับ K-Means และ GMM ที่ 10 กลุ่ม แต่คุณภาพของการแยกกลุ่มก็ยังต่ำกว่ามาก การวิเคราะห์เพิ่มเติมพบว่าปัญหาหลักของ DBSCAN คือการสร้างกลุ่มจำนวนมากเกินไปแม้จะมีคุณภาพการแยกกลุ่มที่ดีเยี่ยม ในทางปฏิบัติธุรกิจร้านอาหารไม่สามารถจัดการกับกลุ่มลูกค้า 54 กลุ่มได้อย่างมีประสิทธิภาพ การพัฒนากลยุทธ์การตลาดเฉพาะกลุ่มสำหรับกลุ่มจำนวนมากนี้ต้องใช้ทรัพยากรมหาศาล และอาจทำให้เกิดความซับซ้อนในการดำเนินงาน ในทางตรงกันข้าม K-Means และ GMM ที่สร้างกลุ่ม 10 กลุ่มให้จำนวนที่เหมาะสมสำหรับการจัดการและการพัฒนากลยุทธ์ที่แตกต่างกันสำหรับแต่ละกลุ่ม

4.7.2.3 การวิเคราะห์ความกระชับของกลุ่ม จากค่า Cluster Cohesion ที่วัดความกระชับของกลุ่มผ่านผลรวมของระยะทางกำลังสองระหว่างจุดข้อมูลกับจุดศูนย์กลาง

โดยค่าที่ต่ำแสดงถึงความกระชับที่ดีกว่า พบว่า DBSCAN ได้ค่าเท่ากับ 0 เนื่องจากเป็นอัลกอริทึมที่ไม่ใช้แนวคิดของจุดศูนย์กลาง จึงไม่สามารถคำนวณค่านี้ได้ในรูปแบบเดียวกับเทคนิคอื่น DBSCAN ใช้แนวคิดของความหนาแน่นในการจัดกลุ่มแทน โดยจุดข้อมูลที่อยู่ในบริเวณที่มีความหนาแน่นสูงจะถูกจัดเข้ากลุ่มเดียวกัน GMM ให้ค่า Cluster Cohesion ที่ดีที่สุดในกลุ่มเทคนิคที่ใช้จุดศูนย์กลางที่ 2.2171 แสดงว่ากลุ่มที่ได้มีความกระชับดี โดยจุดข้อมูลส่วนใหญ่อยู่ใกล้กับจุดศูนย์กลางของกลุ่ม ซึ่งสอดคล้องกับลักษณะของ GMM ที่ใช้การแจกแจงแบบ Gaussian และสามารถปรับรูปร่างของกลุ่มได้ผ่าน covariance matrix การที่ GMM ให้ค่า Cohesion ที่ดีที่สุดนี้แสดงให้เห็นว่าอัลกอริทึมสามารถหาจุดศูนย์กลางที่เหมาะสมและจัดจุดข้อมูลให้อยู่ใกล้กับจุดศูนย์กลางได้อย่างมีประสิทธิภาพ พร้อมทั้งสร้างกลุ่มจำนวน 10 กลุ่มที่เหมาะสมสำหรับการนำไปใช้งาน K-Means ให้ค่า Cluster Cohesion ที่ 2.3393 ซึ่งสูงกว่า GMM เพียง 5.5% แสดงว่ากลุ่มที่ได้มีความกระชับใกล้เคียงกับ GMM แต่มีการกระจายตัวของจุดข้อมูลรอบจุดศูนย์กลางมากกว่าเล็กน้อย ซึ่งอาจเกิดจากข้อจำกัดของ K-Means ที่สมมติว่ากลุ่มมีรูปร่างเป็นทรงกลม อย่างไรก็ตาม ความแตกต่างที่เพียง 5.5% นี้ถือว่าน้อยมาก แสดงว่าในกรณีของข้อมูลนี้ สมมติฐานของ K-Means ยังคงใช้ได้ดีพอสมควร และเช่นเดียวกับ GMM ก็สร้างกลุ่มจำนวน 10 กลุ่มที่เหมาะสมสำหรับการประยุกต์ใช้ QML ให้ค่า Cluster Cohesion สูงที่สุดที่ 2.5421 ซึ่งสูงกว่า GMM ประมาณ 14.7% และสูงกว่า K-Means ประมาณ 8.7% แสดงว่ากลุ่มที่ได้จาก QML มีความกระชับน้อยที่สุด มีจุดข้อมูลจำนวนมากที่กระจายตัวห่างจากจุดศูนย์กลางของกลุ่ม ส่งผลให้กลุ่มมีความเป็นเนื้อเดียวกันน้อยกว่าเทคนิคอื่น แม้ว่า QML จะสร้างกลุ่มจำนวนเท่ากับ K-Means และ GMM แต่คุณภาพของกลุ่มก็ยังต่ำกว่าอย่างชัดเจน การวิเคราะห์เพิ่มเติมพบว่าค่า Cluster Cohesion ที่สูงของ QML นี้สอดคล้องกับค่า Silhouette Score และ Davies-Bouldin Index ที่แย่กว่าเทคนิคอื่น แสดงให้เห็นถึงปัญหาที่สอดคล้องกันในการแบ่งกลุ่มของ QML กลุ่มที่ได้จาก QML ไม่เพียงแต่มีความซ้อนทับกับกลุ่มอื่น แต่ยังมีความกระจายตัวภายในกลุ่มที่มากด้วย ทำให้กลุ่มไม่มีความชัดเจนทั้งในแง่ของการแยกจากกลุ่มอื่นและความเป็นเนื้อเดียวกันภายในกลุ่ม

4.7.3 การวิเคราะห์เปรียบเทียบด้านประสิทธิภาพการคำนวณ

4.7.3.1 การเปรียบเทียบเวลาในการประมวลผล จากการวัดเวลาในการประมวลผล พบว่ามีความแตกต่างอย่างมีนัยสำคัญระหว่างเทคนิคต่างๆ โดย K-Means แสดงประสิทธิภาพที่ดีที่สุดด้วยเวลาประมวลผลเพียง 0.09 วินาที ซึ่งรวดเร็วที่สุดในบรรดาเทคนิคทั้งหมด ตามมาด้วย DBSCAN ที่ใช้เวลา 0.48 วินาที ซึ่งแม้จะช้ากว่า K-Means ประมาณ 5 เท่า แต่ก็ยังถือว่ารวดเร็วมากและเหมาะสมสำหรับการใช้งานแบบเรียลไทม์ ทั้งสองเทคนิคนี้สามารถประมวลผลข้อมูลลูกค้า 6,025 รายได้ภายในเวลาไม่ถึง 1 วินาที ซึ่งเหมาะสมอย่างยิ่งสำหรับการนำไปใช้งานจริงที่ต้องการผลลัพธ์ทันที GMM ใช้เวลาในการประมวลผล 3.6 วินาที ซึ่งนานกว่า K-Means ประมาณ 40 เท่า

และนานกว่า DBSCAN ประมาณ 7.5 เท่า ความช้านี้เกิดจากความซับซ้อนของการคำนวณที่ต้องใช้อัลกอริทึม Expectation-Maximization (EM) ที่ต้องวนซ้ำหลายรอบเพื่อหาค่าความน่าจะเป็น ของการเป็นสมาชิกของแต่ละจุดข้อมูลในแต่ละกลุ่ม ในแต่ละรอบ GMM ต้องคำนวณพารามิเตอร์ ของการแจกแจง Gaussian หลายตัว รวมถึงค่าเฉลี่ย covariance matrix และน้ำหนักของแต่ละกลุ่ม อย่างไรก็ตาม เวลา 3.6 วินาทียังอยู่ในระดับที่ยอมรับได้สำหรับการใช้งานส่วนใหญ่ โดยเฉพาะงานที่ไม่ต้องการผลลัพธ์แบบเรียลไทม์ QML แสดงข้อจำกัดที่สำคัญที่สุดด้านเวลาในการประมวลผล โดยใช้เวลาถึง 2,904 วินาที หรือประมาณ 48 นาที ซึ่งช้ากว่า K-Means ประมาณ 32,267 เท่า ช้ากว่า DBSCAN ประมาณ 6,050 เท่า และช้ากว่า GMM ประมาณ 807 เท่า ความช้าอย่างมากนี้เป็นอุปสรรคสำคัญต่อการนำไปใช้งานจริง ความช้านี้เกิดจากความจำเป็นในการจำลองระบบควอนตัมบนคอมพิวเตอร์แบบคลาสสิก ซึ่งต้องใช้ทรัพยากรในการคำนวณสูงมาก ขั้นตอนการทำงานของ QML ประกอบด้วยหลายส่วนที่ใช้เวลามาก ได้แก่ การเตรียมสถานะควอนตัม การดำเนินการวงจรควอนตัม และการวัดผลและแปลงกลับเป็นข้อมูลแบบคลาสสิก

4.7.3.2 การวิเคราะห์ความสัมพันธ์ระหว่างเวลาและคุณภาพ เมื่อพิจารณาความสัมพันธ์ระหว่างเวลาในการประมวลผลกับคุณภาพของการจัดกลุ่ม พบว่า GMM ให้ผลลัพธ์ที่ดีที่สุดในแง่ของอัตราส่วนระหว่างคุณภาพต่อเวลาเมื่อพิจารณาจากเทคนิคที่มีจำนวนกลุ่มเหมาะสมสำหรับการนำไปใช้งาน โดยให้คุณภาพการจัดกลุ่มสูงสุดด้วย Silhouette Score ที่ 0.4704 และ Cluster Cohesion ที่ 2.2171 แม้ว่าจะใช้เวลาในการประมวลผล 3.6 วินาที ซึ่งนานกว่าเทคนิคอื่น แต่เมื่อเทียบกับคุณภาพที่ได้รับ เวลานี้ยังอยู่ในระดับที่ยอมรับได้และคุ้มค่ากับผลลัพธ์ที่ดีกว่า K-Means แสดงให้เห็นถึงความสมดุลที่ดีระหว่างความเร็วและคุณภาพ โดยใช้เวลาในการประมวลผลเพียง 0.09 วินาที ซึ่งรวดเร็วที่สุด และให้คุณภาพการจัดกลุ่มที่ดีด้วย Silhouette Score ที่ 0.4543 ซึ่งต่ำกว่า GMM เพียง 3.5% เท่านั้น สำหรับการใช้งานที่ต้องการความเร็วเป็นหลักและยอมรับคุณภาพที่ลดลงเล็กน้อย K-Means เป็นทางเลือกที่ยอดเยี่ยม นอกจากนี้ K-Means ยังมีข้อได้เปรียบด้านความเรียบง่ายในการใช้งานและการตีความผลลัพธ์ ทำให้เหมาะสมสำหรับผู้ใช้ที่ไม่มีความเชี่ยวชาญเฉพาะทางด้านเทคนิค DBSCAN แม้จะให้คุณภาพการจัดกลุ่มที่ดีเยี่ยมด้วยคะแนน Silhouette ที่สมบูรณ์แบบที่ 1.0 และใช้เวลาในการประมวลผลเพียง 0.48 วินาที แต่ปัญหาหลักคือการสร้างกลุ่มจำนวน 54 กลุ่ม ซึ่งมากเกินไปสำหรับการนำไปใช้งานจริงในบริบทของธุรกิจร้านอาหาร การมีกลุ่มจำนวนมากเกินไปทำให้ยากต่อการจัดการและพัฒนากลยุทธ์การตลาดเฉพาะกลุ่ม ดังนั้นแม้ว่า DBSCAN จะมีอัตราส่วนคุณภาพต่อเวลาที่ดีในเชิงเทคนิค แต่จำนวนกลุ่มที่มากเกินไป ทำให้ไม่เหมาะสมสำหรับการประยุกต์ใช้งานจริง QML แสดงให้เห็นถึงอัตราส่วนคุณภาพต่อเวลาที่แย่ที่สุด โดยใช้เวลาในการประมวลผลมากที่สุดที่ 2,904 วินาที แต่ให้คุณภาพการจัดกลุ่มที่ต่ำที่สุด ด้วย Silhouette Score ที่ 0.4090 และ Davies-Bouldin Index

ที่ 1.1491 การแลกเปลี่ยนนี้ ไม่คุ้มค่าสำหรับการใช้งานจริง เนื่องจากทั้งเวลาที่ใช้และคุณภาพที่ได้ไม่เป็นไปตามที่คาดหวัง ความล้มเหลวในด้านนี้เป็นข้อจำกัดหลักที่ทำให้ QML ยังไม่เหมาะสมสำหรับการนำไปใช้งานในปัจจุบัน

4.7.4 การวิเคราะห์จำนวนกลุ่มและความเหมาะสมในการนำไปใช้งาน

4.7.4.1 ปัญหาของ DBSCAN กับจำนวนกลุ่มที่มากเกินไป แม้ว่า DBSCAN จะให้คุณภาพการจัดกลุ่มที่ดีเยี่ยมในเชิงเทคนิค แต่การสร้างกลุ่มจำนวน 54 กลุ่มเป็นข้อจำกัดสำคัญที่ทำให้ไม่สามารถนำไปใช้งานจริงได้อย่างมีประสิทธิภาพ ในทางปฏิบัติ ธุรกิจร้านอาหารไม่สามารถพัฒนาและดำเนินกลยุทธ์การตลาดเฉพาะกลุ่มสำหรับกลุ่มลูกค้า 54 กลุ่มได้ การมีกลุ่มจำนวนมากเกินไปนี้ก่อให้เกิดปัญหาหลายประการ ประการแรกคือความซับซ้อนในการทำความเข้าใจการวิเคราะห์ลักษณะเฉพาะของลูกค้า 54 กลุ่มต้องใช้เวลาและความพยายามอย่างมากและอาจทำให้เกิดความสับสนในการทำความเข้าใจความแตกต่างระหว่างกลุ่มต่างๆ โดยเฉพาะกลุ่ม ที่มีลักษณะใกล้เคียงกัน ประการที่สองคือการพัฒนากลยุทธ์การตลาด การสร้างกลยุทธ์ที่แตกต่างกันสำหรับ 54 กลุ่มต้องใช้ทรัพยากรมหาศาล ทั้งในแง่ของเวลา งบประมาณ และบุคลากรซึ่งอาจไม่คุ้มค่ากับผลตอบแทนที่ได้รับ ประการที่สามคือความยากในการดำเนินการ การนำกลยุทธ์การตลาด 54 แบบไปปฏิบัติใช้งานจริงเป็นเรื่องที่ซับซ้อนและยากในการจัดการ อาจทำให้เกิดความสับสนในทีมงานและลดประสิทธิภาพในการดำเนินงาน ประการที่สี่คือการจัดสรรทรัพยากร การแบ่งทรัพยากรทางการตลาดให้กับกลุ่มลูกค้า 54 กลุ่มทำให้แต่ละกลุ่มได้รับทรัพยากรน้อยลงอาจไม่เพียงพอที่จะสร้างผลกระทบที่มีนัยสำคัญ

4.7.4.2 ความเหมาะสมของจำนวนกลุ่ม 10 กลุ่ม K-Means และ GMM ที่สร้างกลุ่มจำนวน 10 กลุ่มให้จำนวนที่เหมาะสมสำหรับการนำไปใช้งานจริงในบริบทของธุรกิจร้านอาหารกลางคืน จำนวนกลุ่ม 10 กลุ่มนี้เป็นจำนวนที่สามารถจัดการได้ในทางปฏิบัติ โดยธุรกิจสามารถ ทำความเข้าใจลักษณะเฉพาะของแต่ละกลุ่มได้อย่างลึกซึ้ง และพัฒนากลยุทธ์การตลาดที่เฉพาะเจาะจงและมีประสิทธิภาพสำหรับแต่ละกลุ่ม จำนวนกลุ่ม 10 กลุ่มช่วยให้สามารถแบ่งแยกลูกค้าออกเป็นกลุ่มที่มีความแตกต่างกันอย่างชัดเจน โดยไม่มากเกินไปจนยากต่อการจัดการ นอกจากนี้ยังช่วยให้สามารถจัดสรรทรัพยากรทางการตลาดได้อย่างมีประสิทธิภาพ โดยแต่ละกลุ่มจะได้รับทรัพยากรเพียงพอที่จะสร้างผลกระทบที่มีนัยสำคัญ การมีกลุ่มจำนวน 10 กลุ่มยังช่วยให้ทีมงานสามารถจำและทำความเข้าใจลักษณะของแต่ละกลุ่มได้ง่าย ทำให้การสื่อสารภายในองค์กรและการดำเนินกลยุทธ์เป็นไปอย่างราบรื่น

4.7.4.3 การเปรียบเทียบความเหมาะสมในการใช้งานจริง เมื่อพิจารณาจากทั้งคุณภาพการจัดกลุ่ม เวลาในการประมวลผล และจำนวนกลุ่มที่เหมาะสม พบว่า GMM เป็นเทคนิคที่เหมาะสมที่สุดสำหรับการนำไปใช้งานจริงในการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน โดยให้คุณภาพการจัดกลุ่ม

สูงสุดในบรรดาเทคนิคที่มีจำนวนกลุ่มเหมาะสม ด้วย Silhouette Score ที่ 0.4704 และ Cluster Cohesion ที่ 2.2171 พร้อมทั้งสร้างกลุ่มจำนวน 10 กลุ่มที่สามารถจัดการได้ในทางปฏิบัติ แม้ว่าจะใช้เวลาในการประมวลผล 3.6 วินาที ซึ่งนานกว่าเทคนิคอื่น แต่ก็ยังอยู่ในระดับที่ยอมรับได้สำหรับการใช้งานทั่วไป K-Means เป็นทางเลือกที่ดีรองลงมา โดยเฉพาะสำหรับการใช้งานที่ต้องการความเร็วและความเรียบง่าย ให้คุณภาพการจัดกลุ่มที่ดีเพียงพอด้วย Silhouette Score ที่ 0.4543 ซึ่งต่ำกว่า GMM เพียง 3.5% แต่มีข้อได้เปรียบด้านเวลาในการประมวลผลที่รวดเร็วมากเพียง 0.09 วินาที และความเรียบง่ายในการใช้งานและการตีความผลลัพธ์ K-Means เหมาะสำหรับผู้ใช้ที่ต้องการเริ่มต้นใช้งานได้รวดเร็วหรือต้องการประมวลผลข้อมูลบ่อยครั้ง DBSCAN แม้จะให้คุณภาพการจัดกลุ่มที่ดีเยี่ยมในเชิงเทคนิค แต่จำนวนกลุ่ม 54 กลุ่มทำให้ไม่เหมาะสมสำหรับการนำไปใช้งานจริงในบริบทของธุรกิจ การพัฒนาและดำเนินกลยุทธ์การตลาดสำหรับกลุ่มจำนวนมากเกินไปนี้ต้องใช้ทรัพยากรมหาศาลและมีความซับซ้อนสูง ซึ่งอาจไม่คุ้มค่ากับผลตอบแทนที่ได้รับ ดังนั้น DBSCAN จึงไม่เหมาะสมสำหรับงานวิจัยนี้ แม้จะมีคะแนนคุณภาพที่สูงที่สุด QML ยังไม่เหมาะสมสำหรับการใช้งานจริงในปัจจุบัน เนื่องจากทั้งคุณภาพการจัดกลุ่มที่ต่ำกว่าเทคนิคอื่นอย่างชัดเจนและเวลาในการประมวลผลที่นานเกินไป แม้ว่าจะสร้างกลุ่มจำนวน 10 กลุ่มที่เหมาะสมเช่นเดียวกับ K-Means และ GMM แต่ประสิทธิภาพโดยรวมยังไม่สามารถแข่งขันกับเทคนิคแบบคลาสสิกได้

4.7.5 การวิเคราะห์ข้อได้เปรียบและข้อจำกัดของแต่ละเทคนิค จากผลการทดลองพบว่า K-Means ให้ผลลัพธ์ที่สมดุลดีระหว่างคุณภาพและประสิทธิภาพ ด้วย Silhouette Score ที่ 0.4543 ซึ่งอยู่ในระดับดี และเวลาในการประมวลผลที่รวดเร็วที่สุดเพียง 0.09 วินาที K-Means สร้างกลุ่มจำนวน 10 กลุ่มซึ่งเป็นจำนวนที่เหมาะสมสำหรับการนำไปใช้งานจริง ความเรียบง่ายในการใช้งานและการตีความผลลัพธ์เป็นจุดเด่นสำคัญของเทคนิคนี้ ทำให้เหมาะสมสำหรับผู้ใช้ที่ไม่มีความเชี่ยวชาญเฉพาะทางลึกซึ้ง K-Means เป็นเทคนิคที่ได้รับการยอมรับอย่างกว้างขวางและมีการใช้งานมากที่สุดในอุตสาหกรรม ทำให้มีเครื่องมือและไลบรารีที่รองรับอย่างครบถ้วน การดำเนินงานที่รวดเร็วทำให้เหมาะสมสำหรับการวิเคราะห์ข้อมูลที่ต้องทำบ่อยครั้งหรือต้องการผลลัพธ์ทันที อย่างไรก็ตาม ข้อจำกัดของ K-Means คือต้องกำหนดจำนวนกลุ่ม k ล่วงหน้าและสมมติว่ากลุ่มมีรูปร่างเป็นทรงกลม ทำให้ไม่สามารถจับกลุ่มที่มีรูปร่างซับซ้อนได้ดี นอกจากนี้ K-Means ยังมีความอ่อนไหวต่อจุดแปลกปลอม สำหรับข้อมูลในงานวิจัยนี้ K-Means สามารถให้ผลลัพธ์ที่ดีพอสมควร แสดงให้เห็นว่าข้อมูลลูกค้าร้านอาหารกลางคืนมีลักษณะที่เข้ากับสมมติฐานของ K-Means ได้ในระดับหนึ่ง และจำนวนกลุ่ม 10 กลุ่มที่ได้เป็นจำนวนที่เหมาะสมสำหรับการพัฒนากลยุทธ์การตลาดเฉพาะกลุ่ม K-Means เป็นทางเลือกที่ดีสำหรับธุรกิจที่ต้องการเริ่มต้นการแบ่งกลุ่มลูกค้าอย่างรวดเร็วและมีประสิทธิภาพ GMM แสดงผลการทำงานที่ดี

ที่สุดในบรรดาเทคนิคที่มีจำนวนกลุ่มเหมาะสมสำหรับการนำไปใช้งาน โดยให้ Silhouette Score สูงสุดที่ 0.4704 และ Cluster Cohesion ดีที่สุดที่ 2.2171 GMM สร้างกลุ่มจำนวน 10 กลุ่ม เช่นเดียวกับ K-Means ซึ่งเป็นจำนวนที่เหมาะสมสำหรับการจัดการและพัฒนากลยุทธ์การตลาด ข้อได้เปรียบหลักของ GMM คือความสามารถในการให้ความน่าจะเป็นของการเป็นสมาชิกของแต่ละกลุ่ม ทำให้สามารถระบุจุดข้อมูลที่อยู่บนขอบเขตระหว่างกลุ่มได้ ซึ่งมีประโยชน์สำหรับการวิเคราะห์เชิงลึกและการพัฒนากลยุทธ์การตลาดที่ยืดหยุ่น ความยืดหยุ่นสูงในการปรับรูปร่างและขนาดของกลุ่มผ่าน covariance matrix ทำให้ GMM สามารถจัดการกับกลุ่มที่มีรูปร่างเป็นวงรีหรือมีการกระจายตัวที่ไม่เท่ากันในแต่ละมิติได้ดี การที่ GMM ให้คุณภาพการจัดกลุ่มที่ดีกว่า K-Means แม้ว่าจะใช้เวลาในการประมวลผลนานกว่า แสดงให้เห็นว่าความซับซ้อนเพิ่มเติมนั้นคุ้มค่ากับผลลัพธ์ที่ดีกว่า อย่างไรก็ตาม ข้อจำกัดหลักของ GMM คือเวลาในการประมวลผลที่นานกว่าที่ 3.6 วินาที ซึ่งอาจเป็นปัญหาสำหรับงานที่ต้องการผลลัพธ์แบบเรียลไทม์

นอกจากนี้ GMM ยังต้องกำหนดจำนวนกลุ่มล่วงหน้า และอัลกอริทึม EM ที่ใช้อาจลู่เข้าสู่ local optimum ได้ สำหรับข้อมูลในงานวิจัยนี้ GMM สามารถสร้างกลุ่มที่มีคุณภาพดีที่สุด ได้ แสดงให้เห็นว่าสมมติฐานของการแจกแจงแบบ Gaussian เหมาะสมกับข้อมูลลูกค้าร้านอาหารกลางคืน และจำนวนกลุ่ม 10 กลุ่มที่ได้เป็นจำนวนที่เหมาะสมสำหรับการประยุกต์ใช้งานจริง GMM เป็นทางเลือกที่ดีที่สุดสำหรับธุรกิจที่ต้องการคุณภาพการแบ่งกลุ่มที่สูงและยอมรับเวลาในการประมวลผลที่นานขึ้นเล็กน้อย DBSCAN แสดงผลการทำงานที่ขัดแย้งกัน ในแง่ของคุณภาพการจัดกลุ่มตามเมตริกต่างๆ DBSCAN ให้ผลลัพธ์ที่ดีเยี่ยมด้วย Silhouette Score ที่สมบูรณ์แบบที่ 1.0 และ Davies-Bouldin Index ที่ 0 แสดงว่ากลุ่มที่ได้มีการแยกจากกันอย่างสมบูรณ์และมีความสอดคล้องภายในกลุ่มสูง นอกจากนี้ DBSCAN ยังใช้เวลาในการประมวลผลเพียง 0.48 วินาที ซึ่งถือว่ารวดเร็วมาก อย่างไรก็ตาม ปัญหาหลักที่ทำให้ DBSCAN ไม่เหมาะสมสำหรับการนำไปใช้งานจริงในงานวิจัยนี้คือการสร้างกลุ่มจำนวน 54 กลุ่ม ซึ่งมากเกินไปสำหรับการจัดการในทางปฏิบัติ จำนวนกลุ่มที่มากเกินไปนี้เกิดจากลักษณะของอัลกอริทึม DBSCAN ที่แบ่งกลุ่มตามความหนาแน่นของข้อมูลอย่างเคร่งครัด ทำให้กลุ่มย่อยเล็กๆ ที่มีความหนาแน่นแตกต่างกันถูกแยกออกเป็นกลุ่มต่างหาก แม้ว่ากลุ่มเหล่านี้อาจมีลักษณะที่ใกล้เคียงกัน ในทางปฏิบัติธุรกิจร้านอาหารไม่สามารถพัฒนาและดำเนินกลยุทธ์การตลาดสำหรับกลุ่มลูกค้า 54 กลุ่มได้อย่างมีประสิทธิภาพ การมีกลุ่มจำนวนมากเกินไปทำให้ยากต่อการทำความเข้าใจความแตกต่างระหว่างกลุ่ม ต้องใช้ทรัพยากรมหาศาลในการพัฒนากลยุทธ์ และมีความซับซ้อนสูงในการดำเนินการ แม้ว่า DBSCAN จะมีข้อได้เปรียบในการค้นหากลุ่มที่มีรูปร่างใดก็ได้และการระบุจุดแปลกปลอมได้อย่างมีประสิทธิภาพ แต่สำหรับงานวิจัยนี้ ข้อจำกัดด้านจำนวนกลุ่มที่มากเกินไปทำให้ DBSCAN ไม่เหมาะสมสำหรับการนำไปใช้งานจริง หากต้องการใช้ DBSCAN อาจต้องมีการปรับแต่งพารามิเตอร์

เพื่อลดจำนวนกลุ่มให้อยู่ในระดับที่จัดการได้ หรือพิจารณาใช้เทคนิคอื่นที่สามารถควบคุมจำนวนกลุ่มได้ดีกว่า QML แสดงผลการทำงานที่ต่ำกว่าเทคนิคอื่นอย่างชัดเจนในทุกด้าน โดยให้ Silhouette Score ที่ต่ำที่สุดที่ 0.4090 Davies-Bouldin Index สูงที่สุดที่ 1.1491 และ Cluster Cohesion สูงที่สุดที่ 2.5421 แสดงให้เห็นว่ากลุ่มที่ได้จาก QML มีความซ้อนทับสูง มีการกระจายตัวภายในกลุ่มมาก และแยกจากกลุ่มอื่นได้ไม่ชัดเจน นอกจากนี้ เวลาในการประมวลผลที่ 2,904 วินาที หรือประมาณ 48 นาที เป็นข้อจำกัดสำคัญที่ทำให้ QML ไม่เหมาะสมสำหรับการใช้งานจริงในปัจจุบัน ความล้มเหลวของ QML ในงานวิจัยนี้น่าจะเกิดจากหลายปัจจัย ประการแรกคือการจำลองระบบควอนตัมบนคอมพิวเตอร์แบบคลาสสิกซึ่งอาจทำให้เกิดข้อผิดพลาดหรือความไม่แม่นยำในการคำนวณ ประการที่สองคือความซับซ้อนของอัลกอริทึมที่อาจต้องการการปรับแต่งพารามิเตอร์อย่างละเอียด ประการที่สามคืออาจมีข้อจำกัดในการแปลงข้อมูลแบบคลาสสิกให้เป็นสถานะควอนตัมผ่าน quantum feature mapping ซึ่งอาจทำให้สูญเสียข้อมูลบางส่วนหรือสร้างความผิดเพี้ยนในข้อมูล แม้ว่า QML จะสร้างกลุ่มจำนวน 10 กลุ่มซึ่งเป็นจำนวนที่เหมาะสมเช่นเดียวกับ K-Means และ GMM แต่คุณภาพของกลุ่มและเวลาในการประมวลผลทำให้ QML ไม่สามารถแข่งขันกับเทคนิคแบบคลาสสิกได้ อย่างไรก็ตาม QML ยังคงเป็นเทคโนโลยีที่มีศักยภาพสูงในอนาคต โดยเฉพาะเมื่อมีฮาร์ดแวร์ควอนตัมที่แท้จริงมาใช้งานได้ ซึ่งจะสามารถแก้ปัญหาด้านเวลาในการประมวลผลและอาจให้คุณภาพที่ดีขึ้นด้วย

4.7.6 การจัดอันดับและสรุปผลการเปรียบเทียบ

4.7.6.1 การจัดอันดับตามเกณฑ์การประเมิน จากการวิเคราะห์ประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธี ผู้วิจัยได้ทำการจัดอันดับแต่ละเทคนิคตามเกณฑ์การประเมินทั้ง 4 ด้าน โดยให้คะแนนอันดับจาก 1 ซึ่งเป็นอันดับที่ดีที่สุด ถึง 4 ซึ่งเป็นอันดับที่แย่ที่สุด โดยพิจารณาเฉพาะเทคนิคที่มีจำนวนกลุ่มเหมาะสมสำหรับการนำไปใช้งานจริง ดังตารางที่ 4-12

ตารางที่ 4-12 การจัดอันดับเทคนิคการแบ่งกลุ่มตามเกณฑ์การประเมินแต่ละด้าน

เทคนิค	Silhouette Score	Davies – Bouldin Index	Cluster Cohesion	Execution Time	คะแนนรวม	อันดับรวม
GMM	1	1	1	2	5	1
K-Means	2	2	2	1	7	2
DBSCAN	-	-	-	-	-	-
QML	3	3	3	3	12	3

หมายเหตุ: *คะแนนรวมที่ต่ำกว่าแสดงถึงประสิทธิภาพที่ดีกว่า (พิจารณาเฉพาะเทคนิคที่มีจำนวนกลุ่มเหมาะสม) DBSCAN ไม่ได้รับการพิจารณาในการจัดอันดับรวมเนื่องจากจำนวนกลุ่ม 54 กลุ่ม มากเกินไปสำหรับการนำไปใช้งานจริง

จากตารางที่ 4-12 เมื่อพิจารณาเฉพาะเทคนิคที่มีจำนวนกลุ่มเหมาะสมสำหรับการนำไปใช้งาน พบว่า GMM ได้อันดับ 1 ด้วยคะแนนรวม 5 คะแนน โดยได้อันดับ 1 ในทุกด้าน ของคุณภาพการจัดกลุ่มและได้อันดับ 2 ในด้านเวลาการประมวลผล รองลงมาคือ K-Means ที่ได้คะแนนรวม 7 คะแนน โดยได้อันดับ 2 ในด้านคุณภาพการจัดกลุ่มแต่ได้อันดับ 1 ในด้านเวลา การประมวลผล สำหรับ QML ได้คะแนนรวม 12 คะแนน เนื่องจากได้อันดับ 3 ในทุกด้าน

4.7.6.2 สรุปผลการเปรียบเทียบโดยรวม จากการเปรียบเทียบประสิทธิภาพของเทคนิคการแบ่งกลุ่มทั้ง 4 วิธีอย่างครอบคลุม สามารถสรุปได้ว่า GMM เป็นเทคนิคที่เหมาะสมที่สุดสำหรับการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืนในงานวิจัยนี้ โดยให้คุณภาพการจัดกลุ่มสูงสุดด้วย Silhouette Score ที่ 0.4704 Davies-Bouldin Index ที่ 0.9014 และ Cluster Cohesion ที่ 2.2171 พร้อมทั้งสร้างกลุ่มจำนวน 10 กลุ่มที่เหมาะสมสำหรับการจัดการและพัฒนากลยุทธ์การตลาด แม้ว่าจะใช้เวลาในการประมวลผล 3.6 วินาที ซึ่งนานกว่าเทคนิคอื่น แต่ก็ยังอยู่ในระดับที่ยอมรับได้และคุ้มค่ากับคุณภาพที่ได้รับ K-Means เป็นทางเลือกที่ดีรองลงมา โดยเฉพาะสำหรับการใช้งานที่ต้องการความเร็วและความเรียบง่าย ให้คุณภาพการจัดกลุ่มที่ดีเพียงพอและใกล้เคียงกับ GMM แต่มีข้อได้เปรียบด้านเวลาในการประมวลผลที่รวดเร็วมากเพียง 0.09 วินาที K-Means เหมาะสำหรับธุรกิจที่ต้องการเริ่มต้นการแบ่งกลุ่มลูกค้าอย่างรวดเร็วหรือต้องการประมวลผลข้อมูลบ่อยครั้ง DBSCAN แม้จะให้คุณภาพการจัดกลุ่มที่ดีเยี่ยมในเชิงเทคนิคด้วยคะแนน Silhouette

ที่สมบูรณ์แบบและใช้เวลาในการประมวลผลเพียง 0.48 วินาที แต่จำนวนกลุ่ม 54 กลุ่ม ทำให้ไม่เหมาะสมสำหรับการนำไปใช้งานจริงในบริบทของธุรกิจร้านอาหาร การมีกลุ่มจำนวนมากเกินไปทำให้ยากต่อการจัดการและพัฒนากลยุทธ์การตลาดเฉพาะกลุ่มอย่างมีประสิทธิภาพ ดังนั้น DBSCAN จึงไม่เหมาะสมสำหรับงานวิจัยนี้ QML ยังไม่เหมาะสมสำหรับการใช้งานจริงในปัจจุบัน เนื่องจากทั้งคุณภาพการจัดกลุ่มที่ต่ำกว่าเทคนิคอื่นอย่างชัดเจนและเวลาในการประมวลผลที่นานเกินไป แม้ว่าจะสร้างกลุ่มจำนวน 10 กลุ่มที่เหมาะสมเช่นเดียวกับ K-Means และ GMM แต่ประสิทธิภาพโดยรวมยังไม่สามารถแข่งขันกับเทคนิคแบบคลาสสิกได้ อย่างไรก็ตาม ผลการวิจัยนี้เป็นข้อมูลที่มีคุณค่าสำหรับการพัฒนา QML ในอนาคต ผลการวิจัยนี้แสดงให้เห็นว่าเทคนิคแบบคลาสสิก โดยเฉพาะ GMM และ K-Means ยังคงมีประสิทธิภาพดีกว่า QML สำหรับการแบ่งกลุ่มลูกค้าในบริบทของร้านอาหารกลางคืน และการพิจารณาจำนวนกลุ่มที่เหมาะสมเป็นปัจจัยสำคัญในการเลือกเทคนิคที่เหมาะสม นอกเหนือจากคุณภาพการจัดกลุ่มตามเมตริกต่างๆ เพียงอย่างเดียว

พระจอมเกล้าพระนครเหนือ มหาวิทยาลัยเทคโนโลยี

บทที่ 5
สรุปผล อภิปรายผล และข้อเสนอแนะ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมการจัดกลุ่มแบบควอนตัมและแบบคลาสสิกสำหรับการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน โดยใช้ข้อมูลจากลูกค้าจริง 6,025 รายที่เก็บรวบรวมตลอดระยะเวลา 3 เดือน ผลการวิจัยที่ได้แสดงให้เห็นถึงความแตกต่างอย่างชัดเจนระหว่างประสิทธิภาพของอัลกอริทึมแต่ละประเภท ซึ่งนำไปสู่ข้อค้นพบที่สำคัญเกี่ยวกับการเลือกใช้เทคนิคการจัดกลุ่มที่เหมาะสมสำหรับธุรกิจบริการ โดยอธิบายรายละเอียดได้ดังนี้

5.1 สรุปผลการวิจัย

5.2 การอภิปรายผลการวิจัย

5.3 ข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมการจัดกลุ่มแบบควอนตัมและแบบคลาสสิกสำหรับการแบ่งกลุ่มลูกค้าร้านอาหารกลางคืน โดยวิเคราะห์ข้อมูลจากลูกค้าจำนวน 6,025 รายที่เก็บรวบรวมตลอดระยะเวลา 3 เดือนจากร้านอาหารกลางคืนในกรุงเทพมหานคร โดยใช้เกณฑ์การประเมิน 4 ด้าน ได้แก่ Silhouette Score, Davies-Bouldin Index, Cluster Cohesion และ Execution Time

ผลการวิจัยพบว่า Gaussian Mixture Model (GMM) เป็นเทคนิคที่มีประสิทธิภาพสูงสุด โดยให้ค่า Silhouette Score สูงสุด (0.4704), Davies-Bouldin Index ต่ำสุด (0.9014) และ Cluster Cohesion ที่ดีที่สุด (2.2171) แสดงให้เห็นความสมดุลที่ดีระหว่างการแยกกลุ่มและความกระชับภายในกลุ่ม ซึ่งสอดคล้องกับงานวิจัยก่อนหน้านี้ที่แสดงให้เห็นว่า GMM มีประสิทธิภาพในการแบ่งกลุ่มลูกค้าผ่านการสร้างแบบจำลองเชิงความน่าจะเป็น ลักษณะ เชิงความน่าจะเป็นของ GMM ช่วยให้สามารถกำหนดขอบเขตกลุ่มได้อย่างยืดหยุ่น ซึ่งมีประโยชน์อย่างมากในการจับลักษณะเฉพาะของลูกค้าที่มีความทับซ้อนกันในสถานการณ์ร้านอาหารกลางคืน นอกจากนี้ GMM ยังสร้างกลุ่มจำนวน 10 กลุ่มซึ่งเป็นจำนวนที่เหมาะสมสำหรับการจัดการและพัฒนากลยุทธ์การตลาดอย่างมีประสิทธิภาพ แม้ว่าจะใช้เวลาในการประมวลผล 3.6 วินาที ซึ่งนานกว่าเทคนิคอื่น แต่ก็ยังอยู่ในระดับที่ยอมรับได้และคุ้มค่ากับคุณภาพที่ได้รับ

K-Means แสดงประสิทธิภาพในระดับปานกลาง (Silhouette Score = 0.4543, Davies-Bouldin Index = 0.9008, Cluster Cohesion = 2.3393) สอดคล้องกับเอกสารวิชาการ

ที่บันทึกการใช้งานอย่างแพร่หลายเนื่องจากความเรียบง่ายและประสิทธิภาพในการคำนวณที่รวดเร็วที่สุดเพียง 0.09 วินาที K-Means เป็นทางเลือกที่ดีรองลงมา โดยเฉพาะสำหรับงานที่ต้องการความเร็วและความเรียบง่าย ให้คุณภาพการจัดกลุ่มที่ต่ำกว่า GMM เพียง 3.5% แต่มีข้อได้เปรียบด้านเวลาในการประมวลผลที่รวดเร็วกว่ามาก เหมาะสำหรับธุรกิจที่ต้องการเริ่มต้นการแบ่งกลุ่มลูกค้าอย่างรวดเร็วหรือต้องการประมวลผลข้อมูลบ่อยครั้ง อย่างไรก็ตาม ข้อจำกัดของอัลกอริทึมในการตรวจจับเฉพาะกลุ่มทรงกลมและความอ่อนไหวต่อจุดศูนย์กลางเริ่มต้น ทำให้มีข้อจำกัดในการจัดการกับรูปแบบพฤติกรรมลูกค้าที่ซับซ้อนและไม่เป็นเชิงเส้น

DBSCAN ได้ค่าเมตริกที่สมบูรณ์แบบในทางทฤษฎี (Silhouette Score = 1.0, Davies-Bouldin Index = 0, Cluster Cohesion = 0) แสดงว่ากลุ่มที่ได้มีการแยกจากกันอย่างสมบูรณ์และมีความสอดคล้องภายในกลุ่มสูง พร้อมทั้งใช้เวลาในการประมวลผลเพียง 0.48 วินาที อย่างไรก็ตาม DBSCAN สร้างกลุ่มถึง 54 กลุ่มซึ่งมีการแยกส่วนมากเกินไป ทำให้ไม่เหมาะสมกับการนำไปใช้ทางธุรกิจ งานวิจัยนี้เผยให้เห็นข้อจำกัดสำคัญว่าค่าเมตริกทางทฤษฎีที่สมบูรณ์แบบไม่ได้รับประกันความเหมาะสมในการนำไปใช้งานจริงทางธุรกิจเมื่อจำนวนกลุ่มที่ได้มากเกินกว่าความต้องการในการแบ่งกลุ่มที่สามารถนำไปปฏิบัติได้ ผลการค้นพบนี้ท้าทายข้อสมมติที่ว่าวิธีการที่อ้างอิงความหนาแน่นจะดีกว่าในทุกกรณี โดยเน้นย้ำความสำคัญของการสร้างสมดุลระหว่างการเพิ่มประสิทธิภาพทางสถิติกับความเป็นไปได้ในทางปฏิบัติ

Quantum Machine Learning (QML) แสดงความไม่มีประสิทธิภาพในการคำนวณอย่างรุนแรงและคุณภาพการจัดกลุ่มที่ต่ำ โดยใช้เวลาในการประมวลผล 2,904 วินาที หรือประมาณ 48 นาที ซึ่งช้ากว่า K-Means ประมาณ 32,267 เท่า ช้ากว่า DBSCAN ประมาณ 6,050 เท่า และช้ากว่า GMM ประมาณ 807 เท่า ผลลัพธ์เหล่านี้แตกต่างอย่างชัดเจนจากข้อได้เปรียบทางทฤษฎีของความขนานแบบควอนตัมและการค้นพบความสัมพันธ์ที่ซับซ้อนที่ได้รับการกล่าวถึงในเอกสารวิชาการด้านคอมพิวเตอร์ควอนตัม ยืนยันข้อจำกัดของคอมพิวเตอร์ควอนตัมในปัจจุบัน

งานวิจัยนี้ดำเนินการภายใต้กรอบแนวคิด CQ (Classical Data-Quantum Computing) คือการประมวลผลข้อมูลลูกค้าแบบคลาสสิกโดยใช้อัลกอริทึมควอนตัม ซึ่งเป็นหมวดหมู่ที่มีการวิจัยมากที่สุดและมีการอ้างว่ามีการประยุกต์ใช้ทางธุรกิจในทางปฏิบัติ อย่างไรก็ตาม ผลการวิจัยแสดงให้เห็นว่าเทคโนโลยีคอมพิวเตอร์ควอนตัมในปัจจุบันยังคงอยู่ในระยะเริ่มต้นของการพัฒนายังไม่เพียงพอสำหรับการนำไปใช้จริงในการวิเคราะห์เชิงพาณิชย์ที่ต้องการข้อมูลเชิงลึกที่ทันเวลาและประสิทธิภาพในการดำเนินงาน แม้ว่า QML จะสร้างกลุ่มจำนวน 10 กลุ่มซึ่งเป็นจำนวนที่เหมาะสมเช่นเดียวกับ K-Means และ GMM แต่ประสิทธิภาพโดยรวมยังไม่สามารถเทียบเท่ากับเทคนิคแบบคลาสสิกได้

งานวิจัยนี้ได้ตอบคำถามพื้นฐานเกี่ยวกับประสิทธิภาพของการจัดกลุ่มแบบควอนตัมเมื่อเทียบกับแบบคลาสสิกผ่านการวิเคราะห์ และโดยสรุปได้ว่า Gaussian Mixture Model ให้ความสมดุลที่เหมาะสมที่สุดระหว่างคุณภาพการจัดกลุ่ม ประสิทธิภาพการคำนวณ และประโยชน์ใช้สอยทางธุรกิจ ในขณะที่แนวทางควอนตัมในปัจจุบันยังไม่เหมาะสมในทางการค้า แนวทางของ GMM ที่ใช้ความน่าจะเป็นพร้อม 10 กลุ่มที่สามารถนำไปปฏิบัติได้ ให้ผลลัพธ์ที่สามารถนำไปใช้งานได้จริงสำหรับการพัฒนากลยุทธ์การตลาดเฉพาะกลุ่มและการปรับปรุงการดำเนินงานในธุรกิจร้านอาหารกลางคืน

งานวิจัยนี้แสดงให้เห็นว่าแนวทางดั้งเดิมที่ให้ความสำคัญกับค่าวัดประสิทธิภาพทางสถิติอย่างเดียว ไม่เพียงพอสำหรับการทำธุรกิจ การที่ GMM กลายเป็นวิธีการที่ดีที่สุดแสดงให้เห็นว่าประสิทธิภาพในทางปฏิบัติต้องการการพิจารณาหลายมิติพร้อมกัน ได้แก่ คุณภาพการจัดกลุ่ม ประสิทธิภาพการคำนวณ และความเป็นไปได้ในการดำเนินงาน ซึ่งเป็นปัญหาการหาค่าที่เหมาะสมแบบหลายมิติ ไม่ใช่การหาค่าสูงสุดของเมตริกเดียว การค้นพบนี้เปลี่ยนแปลงวิธีที่องค์กรควรประเมินและเลือกเทคโนโลยีการจัดกลุ่ม โดยเปลี่ยนจุดสนใจจากความสมบูรณ์แบบทางทฤษฎีไปสู่ความเหมาะสมในการประยุกต์ใช้งานจริง

ความไม่เหมาะสมของคอมพิวเตอร์ควอนตัมในปัจจุบันสำหรับการแบ่งกลุ่มลูกค้านอกเหนือจากการประยุกต์ใช้ ช่องว่างด้านประสิทธิภาพที่ใหญ่มากระหว่างข้อได้เปรียบทางทฤษฎีของควอนตัมกับการนำไปใช้จริง โดยเวลาในการประมวลผลช้ากว่าหลายพันเท่าพร้อมกับคุณภาพการจัดกลุ่มที่ต่ำกว่า แสดงให้เห็นถึงอุปสรรคด้านความพร้อมทางเทคโนโลยีพื้นฐานมากกว่าข้อจำกัดเฉพาะการนำไปใช้งาน หลักฐานนี้ท้าทายสมมติฐานเกี่ยวกับระยะเวลาในการนำเทคโนโลยีมาใช้ที่แพร่หลายในเอกสารวิชาการด้านคอมพิวเตอร์ควอนตัม ชี้ให้เห็นว่าการประยุกต์ใช้ทางธุรกิจในทางปฏิบัติต้องการความก้าวหน้าอย่างมีนัยสำคัญในความสามารถของฮาร์ดแวร์ควอนตัม ไม่ใช่เพียงแค่การปรับปรุงอัลกอริทึม

กรณีของ DBSCAN เผยให้เห็นข้อมูลเชิงลึกที่สำคัญซึ่งสามารถนำไปใช้ในงานวิจัยด้านการเรียนรู้ของเครื่องต่างๆ คือ เมตริกการประเมินทางสถิติที่แยกจาก ธุรกิจอาจนำไปสู่การเลือกอัลกอริทึมที่ไม่เหมาะสม ประสิทธิภาพทางทฤษฎีที่สมบูรณ์แบบซึ่งสร้างกลุ่ม 54 กลุ่มที่ไม่สามารถ ใช้งานได้ แสดงให้เห็นว่ากรอบการประเมินควรรวมข้อจำกัดด้านประโยชน์ใช้สอยควบคู่ไปกับการประเมินทางคณิตศาสตร์ การค้นพบนี้ชี้ให้เห็นว่างานวิจัยด้านการแบ่งกลุ่มลูกค้าควรพัฒนาเมตริกการประเมินแบบผสมผสานที่ให้น้ำหนักกับความสามารถในการตีความทางธุรกิจและความเป็นไปได้ในการนำไปปฏิบัติอย่างชัดเจนควบคู่ไปกับมาตรวัดคุณภาพทางสถิติ

สำหรับผู้ปฏิบัติงานทางธุรกิจ งานวิจัยนี้ให้คำแนะนำที่ชัดเจน GMM เป็นวิธีการที่แนะนำสำหรับการแบ่งกลุ่มลูกค้าที่ต้องการทั้งความซับซ้อนทางการวิเคราะห์และความเป็นไปได้

ในการดำเนินงาน องค์กรควรใช้ GMM เมื่อกลยุทธ์การแบ่งกลุ่มต้องการความเข้าใจที่ละเอียดเกี่ยวกับลักษณะลูกค้าที่ทับซ้อนกันและรูปร่างกลุ่มที่ยืดหยุ่น K-Means ยังคงมีคุณค่าสำหรับการประยุกต์ใช้แบบเรียลไทม์ที่ประสิทธิภาพการคำนวณมีความสำคัญมากกว่าความซับซ้อนของการจัดกลุ่ม

สรุปได้ว่า งานวิจัยนี้ชี้ให้เห็นว่าเทคนิคแบบคลาสสิก โดยเฉพาะ GMM และ K-Means ยังคงมีประสิทธิภาพดีกว่า QML สำหรับการแบ่งกลุ่มลูกค้าในบริบทของร้านอาหารกลางคืน การพิจารณาจำนวนกลุ่มที่เหมาะสมเป็นปัจจัยสำคัญในการเลือกเทคนิคที่เหมาะสม นอกเหนือจากคุณภาพการจัดกลุ่มตามเมตริกต่างๆ เพียงอย่างเดียว ผลการวิจัยนี้มีส่วนช่วยให้เข้าใจการประยุกต์ใช้เทคโนโลยีควอนตัมในการวิเคราะห์เชิงพาณิชย์มากยิ่งขึ้น และให้แนวทางที่ชัดเจนสำหรับการเลือกใช้เทคนิคการจัดกลุ่มที่เหมาะสมในการพัฒนากลยุทธ์การตลาดสำหรับธุรกิจร้านอาหารกลางคืน

5.2 อภิปรายผลการวิจัย

การประเมินเปรียบเทียบอัลกอริทึมการจัดกลุ่มแบบควอนตัมและแบบคลาสสิกสำหรับการแบ่งกลุ่มลูกค้าร้านอาหารในช่วงเวลากลางคืนของการศึกษานี้ เผยให้เห็นข้อมูลเชิงลึกที่สำคัญเกี่ยวกับการประยุกต์ใช้งานจริงของเทคโนโลยีควอนตัมที่กำลังพัฒนาเมื่อเทียบกับวิธีการแบบคลาสสิกที่มีอยู่แล้วในการวิเคราะห์ธุรกิจในโลกแห่งความเป็นจริง

5.2.1 Gaussian Mixture Model (GMM)

Gaussian Mixture Model (GMM) กลายเป็นวิธีการที่เหมาะสมที่สุดด้วยค่า Silhouette Score สูงสุด (0.4704) และ Davies-Bouldin Index ที่ดีกว่า (0.9014) แสดงให้เห็นสมดุลที่มีประสิทธิภาพระหว่างการแยกกลุ่มและความกระชับของกลุ่ม ผลการค้นพบนี้สอดคล้อง กับงานวิจัยก่อนหน้าที่แสดงให้เห็นประสิทธิผลของ GMM ในการแบ่งกลุ่มลูกค้าผ่านการสร้างแบบจำลองแบบความน่าจะเป็น [10] ลักษณะความน่าจะเป็นของ GMM ทำให้สามารถมีขอบเขตกลุ่มที่ยืดหยุ่นและการกำหนดแบบอ่อน ซึ่งมีคุณค่าอย่างยิ่งในการจับลักษณะเฉพาะของลูกค้าที่ทับซ้อนกันในสถานการณ์ร้านอาหารกลางคืน ที่ซึ่งรูปแบบพฤติกรรมอาจไม่แตกต่างกันอย่างชัดเจน

5.2.2 K-Means

K-Means แสดงประสิทธิภาพในระดับปานกลาง (Silhouette Score = 0.4543, Davies-Bouldin Index = 0.9008) สอดคล้องกับเอกสารวิชาการที่บันทึกการประยุกต์ใช้อย่างแพร่หลายเนื่องจากความเรียบง่ายและประสิทธิภาพในการคำนวณ [3, 4] อย่างไรก็ตาม ข้อจำกัดของอัลกอริทึมในการตรวจจับเฉพาะกลุ่มทรงกลมและความอ่อนไหวต่อจุดศูนย์กลางเริ่มต้น ซึ่งได้รับการแก้ไขบางส่วนผ่านการกำหนดค่าเริ่มต้นแบบ k-means++ ทำให้ประสิทธิผลจำกัดลง

เมื่อต้องรับมือกับรูปแบบพฤติกรรมลูกค้าที่ซับซ้อนและไม่เป็นเชิงเส้นที่เป็นลักษณะเฉพาะของสถานประกอบการรับประทานอาหารกลางคืน

5.2.3 DBSCAN

DBSCAN ได้ค่าเมตริกที่สมบูรณ์แบบในทางทฤษฎี (Silhouette Score = 1.0, Davies-Bouldin Index = 0, Cluster Cohesion = 0) แต่สร้างกลุ่มถึง 54 กลุ่มพร้อมการแยกส่วนที่มากเกินไป ทำให้ไม่เหมาะสมกับการประยุกต์ใช้ทางธุรกิจ แม้งานวิจัยก่อนหน้าจะบันทึกความเป็นเลิศของ DBSCAN ในการรับมือกับรูปร่างผิดปกติและการตรวจจับค่าผิดปกติ [13] [14] แต่งานวิจัยนี้เผยให้เห็นข้อจำกัดที่สำคัญ: ค่าเมตริกทางทฤษฎีที่สมบูรณ์แบบไม่ได้รับประกันประโยชน์ใช้สอยทางธุรกิจในทางปฏิบัติเมื่อการแพร่กระจายของกลุ่มเกินกว่าความต้องการในการแบ่งกลุ่มที่สามารถนำไปปฏิบัติได้ ผลการค้นพบนี้ท้าทายข้อสันนิษฐานที่ว่าวิธีการตามความหนาแน่นจะเหนือกว่าในทุกกรณี โดยเน้นย้ำถึงความสำคัญของการสร้างสมดุลระหว่างการเพิ่มประสิทธิภาพทางสถิติกับความเป็นไปได้ในทางปฏิบัติ

5.2.4 Quantum Machine Learning Clustering

Quantum Machine Learning (QML) แสดงความไม่มีประสิทธิภาพในการคำนวณอย่างรุนแรงและคุณภาพการจัดกลุ่มที่ไม่ดี โดยใช้เวลาในการประมวลผลช้ากว่าวิธีการแบบคลาสสิกหลายพันเท่า ผลลัพธ์เหล่านี้แตกต่างอย่างสิ้นเชิงจากศักยภาพทางทฤษฎีของควอนตัมคอมพิวเตอร์ที่เอกสารวิชาการกล่าวถึง [7] [20] ผลลัพธ์นี้สอดคล้องกับการศึกษาเปรียบเทียบล่าสุดที่บันทึกผลลัพธ์ที่หลากหลายในการประยุกต์ใช้การเรียนรู้ของเครื่องแบบควอนตัม แม้งานวิจัยบางชิ้นจะรายงานประสิทธิภาพที่เหนือกว่าของอัลกอริทึมควอนตัมในชุดข้อมูลเฉพาะ เช่น QSVM ที่ได้ความแม่นยำ 100% ในชุดข้อมูล Seeds เทียบกับ 94.97% ของวิธีการแบบคลาสสิก [28]

งานวิจัยนี้ดำเนินการภายใต้กรอบแนวคิด CQ (Classical Data-Quantum Computing) คือการประมวลผลข้อมูลลูกค้าแบบคลาสสิกโดยใช้อัลกอริทึมควอนตัม ซึ่งเป็นหมวดหมู่ที่มีการวิจัย อย่างแพร่หลายที่สุดพร้อมการประยุกต์ใช้ทางธุรกิจในทางปฏิบัติที่ถูกอ้างว่ามี [6] อย่างไรก็ตาม ผลลัพธ์แสดงให้เห็นว่าเทคโนโลยีคอมพิวเตอร์ควอนตัมในปัจจุบันยังคงอยู่ในระยะเริ่มต้นของการพัฒนา ไม่เพียงพอสำหรับการนำไปใช้จริงในการวิเคราะห์เชิงพาณิชย์ที่ต้องการข้อมูลเชิงลึก ที่ทันเวลาและประสิทธิภาพในการดำเนินงาน

5.3 การบูรณาการทฤษฎีกับกรอบการวิจัย

การวิจัยนี้บูรณาการทฤษฎีและงานวิจัยที่เกี่ยวข้องทั้ง 6 หัวข้ออย่างเป็นระบบ โดยนำแบบจำลอง RFM (Recency, Frequency, Monetary) จากงานวิจัยของ Maniyara และคณะ [10] มาออกแบบตัวแปรในแบบสอบถาม ได้แก่ ความถี่ในการมา ช่วงเวลาที่มา

และค่าใช้จ่ายเฉลี่ย พร้อมขยายขอบเขตด้วยข้อมูลทั่วไปและพฤติกรรมการใช้บริการตามแนวทางของ Yang [12] สำหรับวิธีการจัดกลุ่มแบบดั้งเดิม ผู้วิจัยเลือกใช้สามเทคนิคจากการทบทวนวรรณกรรม คือ K-Means จาก Julian และ Haridasan [3] เนื่องจากความเรียบง่ายและประสิทธิภาพสูง DBSCAN จาก Sakina และคณะ [13] ที่ให้ Davies-Bouldin Index ต่ำที่สุด และ GMM จาก Jerin และคณะ [15] ที่จัดการข้อมูลซับซ้อนได้ดี ผู้วิจัยพัฒนาอัลกอริทึม QML ภายใต้กรอบแนวคิด CQ (Classical Data-Quantum Computing) ตาม Jiang [6] โดยใช้หลักการ Superposition และ Entanglement จาก Pushpak และ Jain [17] ร่วมกับ Quantum Genetic Algorithm จาก Shao และคณะ [23] และ quantum feature mapping จาก Maouaki และคณะ [25]

การบูรณาการนี้นำไปสู่การออกแบบกรอบการวิจัยที่ใช้ 4 เทคนิค ประเมินด้วย 4 เมตริก ทดสอบกับข้อมูล 8 ระดับคือ 300-6,025 หน่วยตัวอย่าง และออกแบบตัวแปรที่ครอบคลุม ทำให้งานวิจัยสามารถเปรียบเทียบผลกับงานวิจัยก่อนหน้าได้

5.4 ข้อเสนอแนะการวิจัย

จากผลการวิจัยที่ได้ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมการจัดกลุ่มแบบควอนตัมและแบบคลาสสิก ผู้วิจัยมีข้อเสนอแนะดังนี้

5.4.1 แนวทางการแบ่งกลุ่มลูกค้าและพัฒนากลยุทธ์การตลาดสำหรับร้านอาหารกลางคืน

ร้านอาหารกลางคืนควรนำ Gaussian Mixture Model มาใช้ในการแบ่งกลุ่มลูกค้าเพื่อพัฒนากลยุทธ์การตลาดที่เหมาะสมกับแต่ละกลุ่ม โดยสามารถเริ่มต้นด้วยการจัดทำฐานข้อมูลลูกค้า ที่มีการบันทึกข้อมูลอย่างสม่ำเสมอ แล้วนำผลการวิเคราะห์ไปใช้ในการออกแบบโปรโมชั่นและแพ็คเกจอาหารที่ตรงกับความต้องการของลูกค้าแต่ละกลุ่ม สำหรับธุรกิจขนาดเล็กที่ต้องการความรวดเร็วในการประมวลผล K-Means ก็เป็นทางเลือกที่ดีเนื่องจากใช้เวลาน้อยและให้ผลลัพธ์ที่ใกล้เคียงกับ GMM

5.4.2 แนวทางการขยายผลการวิจัย

นักวิจัยที่สนใจศึกษาต่อยอดสามารถขยายการวิจัยไปยังธุรกิจบริการประเภทอื่น เช่น คาเฟ่ บาร์ หรือร้านอาหารที่เปิดในช่วงเวลาต่างกัน เพื่อดูว่าผลการวิจัยสามารถนำไปใช้ได้กับบริบทอื่นหรือไม่ การเพิ่มตัวแปรที่สะท้อนพฤติกรรมยุคดิจิทัล เช่น การรีวิวออนไลน์หรือการแชร์ภาพอาหารบนโซเชียลมีเดีย อาจทำให้ได้ผลการแบ่งกลุ่มที่สะท้อนพฤติกรรมลูกค้าจริงมากขึ้น นอกจากนี้ การศึกษาการเปลี่ยนแปลงของกลุ่มลูกค้าตามช่วงเวลาหรือฤดูกาลก็เป็นหัวข้อที่น่าสนใจ

5.4.3 โอกาสและข้อจำกัดของเทคโนโลยีควอนตัม

สำหรับเทคโนโลยีควอนตัม แม้ว่าผลการวิจัยจะแสดงให้เห็นว่ายังไม่เหมาะสมกับการใช้งานจริงในปัจจุบัน แต่ควรติดตามความก้าวหน้าของเทคโนโลยีนี้ต่อไป โดยเฉพาะเมื่อมีการพัฒนาฮาร์ดแวร์ควอนตัมที่มีประสิทธิภาพดีขึ้น การทดสอบ QML อีกครั้งในอนาคตอาจให้ผลลัพธ์ที่แตกต่างออกไป การพัฒนาอัลกอริทึมแบบผสมผสานที่ใช้จุดแข็งของทั้งแนวทางคลาสสิกและควอนตัมร่วมกันอาจเป็นทางเลือกที่น่าสนใจ

การวิจัยครั้งนี้แสดงให้เห็นว่าการเลือกอัลกอริทึมที่เหมาะสมไม่ควรพิจารณาเพียงค่าเมตริก ทางสถิติเท่านั้น แต่ต้องคำนึงถึงความเป็นไปได้ในการนำไปใช้จริงด้วย ดังนั้น งานวิจัยในอนาคต ควรพัฒนากรอบการประเมินที่รวมทั้งตัวชี้วัดทางสถิติและตัวชี้วัดทางธุรกิจเข้าด้วยกัน เพื่อให้ได้ผลการประเมินที่สมบูรณ์และตรงกับความต้องการของผู้ใช้งานจริง

พระจอมเกล้าพระนครเหนือ มหาวิทยาลัยเทคโนโลยี

บรรณานุกรม

[1] B. Balakrishnan, S. Gawali, S. Tiwari, S. S. Kalsi, and V. Naregal, "Data Analytics of Restaurants using AI," in 2023 9th International Conference on Smart Computing and Communications (ICSCC), 17-19 Aug. 2023 2023, pp. 506-511, doi: 10.1109/ICSCC59169.2023.10335092.

[2] K. Kumarasiri, J. Premarathne, N. Mendis, M. Bandhuratne, A. Karunasena, and L. Weerasinghe, "SmartDine: A Comprehensive System for Enhancing Customer Satisfaction and Optimizing Business Operations in the Restaurant Context," in 2024 6th International Conference on Advancements in Computing (ICAC), 12-13 Dec. 2024 2024, pp. 378-383, doi: 10.1109/ICAC64487.2024.10851011.

[3] A. Julian and R. H. S, "Optimizing Customer Segmentation through Machine Learning," in 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), 9-10 Feb. 2024 2024, vol. 5, pp. 413-416, doi: 10.1109/IC2PCT60090.2024.10486699.

[4] D. Parmar, P. Tanna, and N. Raza, "Clustering and Optimization Techniques for Enhanced Segmentation in Retail," in 2024 IEEE 2nd International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP), 6-8 Dec. 2024 2024, pp. 1-7, doi: 10.1109/IHCSP63227.2024.10960248.

[5] S. Manimozhi, D. Ruby, and K. Biruntha, "An Elegant Evaluation: Triangulating Clustering Methods for Customer Segmentation," in 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 22-23 Feb. 2024 2024, pp. 1-6, doi: 10.1109/ic-ETITE58242.2024.10493354.

[6] J. R. Jiang, "A Quick Overview of Quantum Machine Learning," in 2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE),

27-29 Oct. 2023 2023, pp. 301-304, doi: 10.1109/ECICE59523.2023.10383149.

[7] S. Misra and P. Rani, "Quantum Machine Learning: A Comprehensive Overview and Analysis," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 24-28 June 2024 2024, pp. 1-5, doi: 10.1109/ICCCNT61001.2024.10724803.

[8] A. S. Selvi, S. Narendrakumar, G. V. M. R. Samanvay, C. T. A. Vinayak, S. A. Vembu, and S. S. Pandi, "From Classical to Quantum: Evaluating Machine Learning Enhancements," in 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), 28-29 Nov. 2024 2024, vol. 1, pp. 778-784, doi: 10.1109/ICAICCIT64383.2024.10912309.

[9] G. S. Walia, S. P. S, K. Karthikeya, D. Suresh, P. Sudheer, and V. Syambabu, "Improving Data Encoding for Enhanced AI Performance in Complex Datasets via Quantum Feature Space Mapping: Harnessing Quantum Algorithms," in 2025 International Conference on Pervasive Computational Technologies (ICPCT), 8-9 Feb. 2025 2025, pp. 985-990, doi: 10.1109/ICPCT64145.2025.10940638.

[10] K. Maniyara, K. Shah, V. H. Gaidhane, and R. Wanjari, "An Effective Customer Segmentation using RFM Ranking Techniques," in 2024 International Conference on Modeling, Simulation & Intelligent Computing (MoSICom), 9-11 Dec. 2024 2024, pp. 592-597, doi: 10.1109/MoSICom63082.2024.10882067.

[11] S. Chen, "E-Commerce Customer Segmentation and Precision Marketing Strategy Based on Cluster Analysis," in 2024 International Conference on Internet of Things, Robotics and Distributed Computing (ICIRDC), 29-31 Dec. 2024 2024, pp. 10-15, doi: 10.1109/ICIRDC65564.2024.00008.

[12] S. Yang, "Market Segmentation and Profiling Based on Customer Lifetime Value using Data Mining Techniques: First Bank USA One Case Study," in 2024 20th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 27-29 July 2024 2024, pp. 1-8, doi: 10.1109/ICNC-FSKD64080.2024.10702210.

[13] N. Sakina, A. P. Arun, R. P, V. P. H, and P. K. Gupta, "Optimizing Customer Segmentation: A Comparative Analysis of Clustering Algorithms Using Evaluation Metrics," in 2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS), 7-9 Nov. 2024 2024, pp. 1-6, doi: 10.1109/CSITSS64042.2024.10816952.

[14] V. P. Salve and M. P. Ghatule, "Comprehensive Analysis of Clustering Methods: Focusing on Fuzzy Clustering," in 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI), 20-22 Jan. 2025 2025, pp. 1109-1117, doi: 10.1109/ICMSCI62561.2025.10894658.

[15] J. Q. Jerin, N. K. Khan, S. Biswas, and N. Sharmin, "Comparative Study of Clustering Algorithms: Scenario Based on Boston Crime Dataset," in 2023 26th International Conference on Computer and Information Technology (ICCIT), 13-15 Dec. 2023 2023, pp. 1-6, doi: 10.1109/ICCIT60459.2023.10441641.

[16] T. Babu, P. M. E, and R. R. Nair, "Data-Driven Insights: Mall Customer Segmentation through Fuzzy C-Means Clustering," in 2024 International BIT Conference (BITCON), 7-8 Dec. 2024 2024, pp. 1-5, doi: 10.1109/BITCON63716.2024.10984994.

[17] S. N. Pushpak and S. Jain, "An Introduction To Quantum Machine Learning Techniques," in 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 3-4 Sept. 2021 2021, pp. 1-6, doi: 10.1109/ICRITO51393.2021.9596240.

[18] P. Kuppusamy, N. Y. Kumar, J. Dontireddy, and C. Iwendi, "Quantum Computing and Quantum Machine Learning Classification – A Survey," in 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), 8-9 Oct. 2022 2022, pp. 200-204, doi: 10.1109/ICCCMLA56841.2022.9989137.

[19] I. Manan, F. Rehman, H. Sharif, N. Riaz, M. Atif, and M. Aqeel, "Quantum Computing and Machine Learning Algorithms - A Review," in 2022 3rd International Conference on Innovations in Computer Science & Software Engineering (ICONICS), 14-15 Dec. 2022 2022, pp. 1-6, doi: 10.1109/ICONICS56716.2022.10100452.

[20] Arshdeep, K. Rani, K. Singh, S. Chauhan, H. Kumar, and V. Mishra, "Quantum Machine Learning: Redefining Computational Boundaries," in 2025 Seventh International Conference on Computational Intelligence andCommunication Technologies (CCICT), 11-12 April 2025 2025, pp. 48-54, doi: 10.1109/CCICT65753.2025.00018.

[21] D. Bandi et al., "Quantum Machine Learning for Security Data Analysis," in 2023 IEEE World AI IoT Congress (AIIoT), 7-10 June 2023 2023, pp. 0460-0465, doi: 10.1109/AIIoT58121.2023.10174430.

[22] N. M. T, A. Hiremath, M. N, S. L. Peng, R. S. M, and K. P. S, "A Survey on Machine Learning Techniques Using Quantum Computing," in 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 26-27 Dec. 2022 2022, pp. 1-6, doi: 10.1109/ICERECT56837.2022.10059764.

[23] Z. Shao, S. Zhang, and S. Kumar, "Quantum K-Means Model based on Optimization Framework," in 2023 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 14-16 April 2023 2023, pp. 372-376, doi: 10.1109/IPEC57296.2023.00071.

[24] A. U. Haque, J. Cho, and H. Shin, "Continuous-Variable Quantum K-Means Clustering with Photonic Gaussian States," in 2024 15th International Conference on Information and Communication Technology Convergence (ICTC), 16-18 Oct. 2024 2024, pp. 500-505, doi: 10.1109/ICTC62082.2024.10826857.

[25] W. E. Maouaki, N. Innan, A. Marchisio, T. Said, M. Bennai, and M. Shafique, "Quantum Clustering for Cybersecurity," in 2024 IEEE International Conference on Quantum Computing and Engineering (QCE), 15-20 Sept. 2024 2024, vol. 02, pp. 5-10, doi: 10.1109/QCE60285.2024.10243.

[26] O. Ouedrhiri, U. Faghihi, F. Toure, and O. Banouar, "Quantum Fidelity Based Fuzzy C-Means Clustering Algorithm," in 2024 IEEE International Conference on Quantum Computing and Engineering (QCE), 15-20 Sept. 2024 2024, vol. 02, pp. 138-143, doi: 10.1109/QCE60285.2024.10267.

[27] M. Nabatian, J. Tanha, A. R. Ebrahimzadeh, N. Samadi, and N. Razzaghi–Asl, "An improvement on quantum clustering," in 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS), 29-30 Dec. 2021 2021, pp. 01-06, doi: 10.1109/ICSPIS54653.2021.9729349.

[28] P. Bhaskaran and S. Prasanna, "An Accuracy Analysis Of Classical And Quantum Machine Learning Algorithms Using Different Datasets," in 2024 2nd International Conference on Networking and Communications (ICNWC), 2-4 April 2024 2024, pp. 1-8, doi: 10.1109/ICNWC60771.2024.10537433.

[29] M. Nadim, M. Hassan, A. K. Mandal, and C. K. Roy, "Quantum Vs. Classical Machine Learning Algorithms for Software Defect Prediction: Challenges and Opportunities," in 2025 IEEE/ACM International Workshop on Quantum Software Engineering (Q-SE), 3-3 May 2025 2025, pp. 35-42, doi: 10.1109/Q-SE66736.2025.00012.

[30] P. Liu, H. Zhu, Y. Hua, H. Wu, Y. Yu, and R. Wang, "Clustering Analysis of Electric Vehicle Charging Behavior Based on the Improved Quantum K-Means

Algorithm," in 2025 IEEE International Conference on Power and Integrated Energy Systems (ICPIES), 7-9 April 2025 2025, pp. 242-246, doi: 10.1109/ICPIES65420.2025.11070203.

[31] S. Hemajothi, K. Stella, P. Mahesh, B. K, R. J, and S. H. N, "Enhancing Customer Retention in Restaurants using Machine Learning Algorithm," in 2025 5th International Conference on Soft Computing for Security Applications (ICSCSA), 4-6 Aug. 2025 2025, pp. 1497-1503, doi: 10.1109/ICSCSA66339.2025.11170715.

[32] N. Kaur, N. Mahajan, V. Singh, and A. Gupta, "Artificial Intelligence Revolutionizing The Restaurant Industry - Analyzing Customer Experience Through Data Mining and Thematic Content Analysis," in 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM), 22-24 Feb. 2023 2023, pp. 1-5, doi: 10.1109/ICIPTM57143.2023.10117897.

[33] D. P. Kaur, Tanay, P. Malhotra, R. Dutta, and H. J. Kaur, "Restaurant Automation: A Framework using Data Acquisition and Visualization for Improved Customer Experiences," in 2023 4th International Conference on Data Analytics for Business and Industry (ICDABI), 25-26 Oct. 2023 2023, pp. 409-413, doi: 10.1109/ICDABI60145.2023.10629293.

ภาคผนวก ก

แบบสอบถามการวิจัย

แบบสอบถามเรื่อง “การวิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน”

ผู้วิจัย นาย ธัญวัฒน์ แก้วประดิษฐ์ นักศึกษาปริญญาโท สาขาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ข้อมูลติดต่อ Email: tanyavat_kae@outlook.com โทรศัพท์ +66 855549395

แบบสอบถามมีทั้งหมด 2 ส่วน

ส่วนที่ 1 ข้อมูลเกี่ยวกับสถานภาพและข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

1. เพศ

☐ 1.ชาย ☐ 2. หญิง ☐ 3. LGBTQ+

2. อายุ

☐ 1. 18-22 ปี ☐ 2. 23-30 ปี
☐ 3. 31-40 ปี ☐ 4. 40 ปีขึ้นไป

3. สถานะ

☐ 1. โสด ☐ 2. มีแฟน / สมรส ☐ 3. หย่า / แยกทาง

4. อาชีพ

☐ 1. นักเรียน/นักศึกษา ☐ 2. พนักงานบริษัท ☐ 3. ธุรกิจส่วนตัว
☐ 4. ข้าราชการ ☐ 5. อาชีพอิสระ ☐ 6. ว่างงาน

5. รายได้ส่วนบุคคล/เดือน

☐ 1. น้อยกว่า 15,000 บาท ☐ 2. 15,001-30,000 บาท
☐ 3. 30,001-50,000 บาท ☐ 4. มากกว่า 50,000 บาท

ส่วนที่ 2 พฤติกรรมการใช้บริการ / การรับรู้สื่อ

1. คุณมาร้านนี้บ่อยแค่ไหน

☐ 1. ครั้งแรก ☐ 2. 2-3 ครั้ง / เดือน

☐ 3. 1 ครั้ง / สัปดาห์ ☐ 4. 2-3 ครั้ง / สัปดาห์

☐ 5. เกือบทุกวัน ☐ 6. มากกว่า 1 ครั้ง/ เดือน

2. ช่วงเวลาที่คุณมาบ่อยที่สุด

☐ 1. 17.00-19.00 ☐ 2. 19.01-21.00

☐ 3. 21.01-23.00 ☐ 4. 23.01 – 02.00

3. วันที่คุณมาบ่อยที่สุด

☐ 1. วันธรรมดา ☐ 2. วันเสาร์ อาทิตย์

4. ปกติคุณมาที่ร้านกับใคร

☐ 1. คนเดียว ☐ 2. เพื่อน 2-3 คน

☐ 3. กลุ่มเพื่อนมากกว่า 3 คน ☐ 4. แฟน / คู่รัก

☐ 5. เพื่อนร่วมงาน

5. ค่าใช้จ่ายเฉลี่ย / ครั้ง

☐ 1. น้อยกว่า 500 บาท ☐ 2. 500 - 1,000 บาท

☐ 3. 1,001 - 2,000 บาท ☐ 4. 2,001- 3,000 บาท

☐ 5. มากกว่า 3,000 บาท

6. คุณเคยใช้บริการ ร้านอื่นหรือไม่

☐ 1. ไม่เคย ☐ 2. 1 ร้าน

☐ 3. 2-3 ร้าน ☐ 4. 4-5 ร้าน

☐ 5. มากกว่า 5 ร้าน

7. ช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร

☐ 1. 06.00-12.00 ☐ 2. 12.01-18.00

☐ 3. 18.01-24.00 ☐ 4. ไม่ต้องการข้อมูลข่าวสาร

ภาคผนวก ข

แบบฟอร์มที่ให้ผู้เชี่ยวชาญแสดงความคิดเห็น

เพื่อพิจารณาความสอดคล้องของข้อคำถาม และการให้ข้อเสนอแนะ

จากการคำนวณหาค่า IOC (Item Objective Congruence)

แบบฟอร์มที่ให้ผู้เชี่ยวชาญแสดงความคิดเห็น (1)
ร่างแบบสอบถามเพื่อการวิจัย (IOC)

เรื่อง “แบบสอบถามสำหรับงานวิจัยการ วิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน”

คำชี้แจง

1. แบบสอบถามฉบับนี้อยู่ในขั้นตอนการศึกษาเพื่อ วิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน

2. แบบสอบถามฉบับนี้มุ่งตรวจสอบเพื่อหาความเที่ยงตรง (Validity) โดยการวิเคราะห์ดัชนีความสอดคล้อง (Index of Item Objective Congruence: IOC) ของแบบสอบถามและข้อเสนอแนะของผู้เชี่ยวชาญ เพื่อนำไปปรับปรุงแบบสอบถามให้สมบูรณ์ยิ่งขึ้น

3. แบบสอบถามฉบับนี้มีทั้งหมด 2 ส่วน ประกอบด้วย

ส่วนที่ 1 ข้อมูลทั่วไป

ส่วนที่ 2 พฤติกรรมการใช้บริการ / การรับรู้สื่อ

4. ขอความกรุณาผู้ทรงคุณวุฒิหรือท่านผู้เชี่ยวชาญ ช่วยพิจารณาร่างแบบสอบถามว่ามีความสอดคล้องกับวัตถุประสงค์ของการวิจัยเรื่องนี้หรือไม่ ด้วยการให้คะแนนในแต่ละข้อคำถามในระบบ IOC โดยการทำเครื่องหมาย ✓ ลงในช่องว่าง

ชื่อผู้เชี่ยวชาญ ท่าน ดร. วสุพล ตรีโสภากุล

ตำแหน่ง MD at UResearch Consulting & Partnering

เกณฑ์การให้คะแนนในระบบ IOC

1) ให้ 1 คะแนน เมื่อแน่ใจว่าข้อนั้นมีเนื้อหาที่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

2) ให้ 0 คะแนน เมื่อไม่แน่ใจว่าข้อนั้นมีเนื้อหาที่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

3) ให้ -1 คะแนน เมื่อแน่ใจว่าข้อนั้นมีเนื้อหาไม่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

ผู้วิจัยขอความกรุณาท่านผู้ทรงคุณวุฒิหรือท่านผู้เชี่ยวชาญ ให้ข้อเสนอแนะหรือความคิดเห็นเพิ่มเติมในประเด็นที่ยังไม่สมบูรณ์ โดยการเขียนข้อเสนอแนะไว้ท้ายข้อความนั้นๆ ผู้วิจัยขอขอบพระคุณท่านที่กรุณาเสียสละเวลาในการตอบแบบสอบถามทุกข้ออย่างถูกต้องครบถ้วน

ส่วนที่ 1 ข้อมูลเกี่ยวกับสถานภาพ และข้อมูลทั่วปของผู้ตอบแบบสอบถาม

คำชี้แจงสำหรับผู้เชี่ยวชาญ: โปรดพิจารณาว่าข้อความที่เกี่ยวกับสถานภาพของผู้ตอบเหมาะสมหรือไม่อย่างไร

คำชี้แจง: โปรดใส่เครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
1	เพศ ☐ 1. ชาย ☐ 2. หญิง ☐ 3. LGBTQ+	✓			ไม่ควรระบุเพศสภาพปล่อยอิสระ
2	อายุ ☐ 1. 18-22 ปี ☐ 2. 23-30 ปี ☐ 3. 31-40 ปี ☐ 4. 40 ปีขึ้นไป	✓			
3	สถานะ ☐ 1. โสด ☐ 2. มีแฟน/สมรส ☐ 3. หย่า/แยกทาง	✓			
4	อาชีพ ☐ 1. นักศึกษา ☐ 2. พนักงานบริษัท ☐ 3. ธุรกิจส่วนตัว ☐ 4. ข้าราชการ ☐ 5. อาชีพอิสระ ☐ 6. ว่างงาน	✓			

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
5	รายได้ส่วนบุคคล/เดือน ☐ 1. น้อยกว่า 15,000 บาท ☐ 2. 15,000-30,000 บาท ☐ 3. 30,001-50,000 บาท ☐ 4. มากกว่า 50,000 บาท	✓			

ส่วนที่ 2 ข้อมูลเกี่ยวกับพฤติกรรมการใช้บริการ และการรับรู้สื่อของผู้ตอบแบบสอบถาม

คำชี้แจงสำหรับผู้เชี่ยวชาญ: โปรดพิจารณาว่าข้อความที่เกี่ยวกับพฤติกรรมการใช้บริการและการรับรู้สื่อของผู้ตอบเหมาะสมหรือไม่อย่างไร

คำชี้แจง: โปรดใส่เครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
1	คุณมาร้านนี้บ่อยแค่ไหน ☐ 1. ครั้งแรก ☐ 2. 2-3 ครั้ง / เดือน ☐ 3. 1 ครั้ง / สัปดาห์ ☐ 4. 2-3 ครั้ง / สัปดาห์ ☐ 5. เกือบทุกวัน ☐ 6. มากกว่า 1 ครั้ง / เดือน	✓			
2	ช่วงเวลาที่คุณมาบ่อยที่สุด ☐ 1. 17.00-19.00 ☐ 2. 19.01-21.00 ☐ 3. 21.01-23.00 ☐ 4. 23.01–02.00	✓			

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
3	วันที่คุณมาบ่อยที่สุด ☐ 1. วันธรรมดา ☐ 2. วันเสาร์ - อาทิตย์	✓			
4	ปกติคุณมาที่ร้านกับใคร ☐ 1. คนเดียว ☐ 2. เพื่อน 2-3 คน ☐ 3. กลุ่มเพื่อนมากกว่า 3 คน ☐ 4. แฟน / คู่รัก ☐ 5. เพื่อนร่วมงาน	✓			ระบุคำตอบเป็นช่วงจะทำให้ผู้ตอบตอบสะดวกกว่า
5	ค่าใช้จ่ายเฉลี่ย / ครั้ง ☐ 1. น้อยกว่า 500 บาท ☐ 2. 500 - 1,000 บาท ☐ 3. 1,001 - 2,000 บาท ☐ 4. 2,001 – 3,000 บาท ☐ 5. มากกว่า 3,000 บาท	✓			
6	คุณเคยใช้บริการ ร้านอื่นหรือไม่ ☐ 1. ไม่เคย ☐ 2. 1 ร้าน ☐ 3. 2-3 ร้าน ☐ 4. 4-5 ร้าน ☐ 5. มากกว่า 5 ร้าน	✓			
7	ช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร ☐ 1. 06.00-12.00 ☐ 2. 12.01-18.00 ☐ 3. 18.01-24.00 ☐ 4. ไม่ต้องการข้อมูลข่าวสาร	✓			

ข้อเสนอแนะของผู้เชี่ยวชาญ

แนะนำตัวเลือกคำตอบควรเป็นช่วงเพื่อให้ผู้ตอบสะดวกในการตอบ เนื่องจากผู้ตอบอยู่ในร้านควรใช้เวลาในการตอบไม่นาน

แบบฟอร์มที่ให้ผู้เชี่ยวชาญแสดงความคิดเห็น (2)
ร่างแบบสอบถามเพื่อการวิจัย (IOC)

เรื่อง “แบบสอบถามสำหรับงานวิจัยการ วิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน”

คำชี้แจง

1. แบบสอบถามฉบับนี้อยู่ในขั้นตอนการศึกษาเพื่อ วิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน

2. แบบสอบถามฉบับนี้มุ่งตรวจสอบเพื่อหาความเที่ยงตรง (Validity) โดยการวิเคราะห์ดัชนีความสอดคล้อง (Index of Item Objective Congruence: IOC) ของแบบสอบถามและข้อเสนอแนะของผู้เชี่ยวชาญ เพื่อนำไปปรับปรุงแบบสอบถามให้สมบูรณ์ยิ่งขึ้น

3. แบบสอบถามฉบับนี้มีทั้งหมด 2 ส่วน ประกอบด้วย

ส่วนที่ 1 ข้อมูลทั่วไป

ส่วนที่ 2 พฤติกรรมการใช้บริการ / การรับรู้สื่อ

4. ขอความกรุณาผู้ทรงคุณวุฒิหรือท่านผู้เชี่ยวชาญ ช่วยพิจารณาร่างแบบสอบถามว่ามีความสอดคล้องกับวัตถุประสงค์ของการวิจัยเรื่องนี้หรือไม่ ด้วยการให้คะแนนในแต่ละข้อคำถามในระบบ IOC โดยการทำเครื่องหมาย ✓ ลงในช่องว่าง

ชื่อผู้เชี่ยวชาญ ท่าน นฤมล สิมะขจรบุญ

ตำแหน่ง Former Deputy Director at True Corporation

เกณฑ์การให้คะแนนในระบบ IOC

1) ให้ 1 คะแนน เมื่อแน่ใจว่าข้อนั้นมีเนื้อหาที่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

2) ให้ 0 คะแนน เมื่อไม่แน่ใจว่าข้อนั้นมีเนื้อหาที่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

3) ให้ -1 คะแนน เมื่อแน่ใจว่าข้อนั้นมีเนื้อหาไม่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

ผู้วิจัยขอความกรุณาท่านผู้ทรงคุณวุฒิหรือท่านผู้เชี่ยวชาญ ให้ข้อเสนอแนะหรือความคิดเห็นเพิ่มเติมในประเด็นที่ยังไม่สมบูรณ์ โดยการเขียนข้อเสนอแนะไว้ท้ายข้อความนั้นๆ ผู้วิจัยขอขอบพระคุณท่านที่กรุณาเสียสละเวลาในการตอบแบบสอบถามทุกข้ออย่างถูกต้องครบถ้วน

ส่วนที่ 1 ข้อมูลเกี่ยวกับสถานภาพ และข้อมูลทั่วปของผู้ตอบแบบสอบถาม

คำชี้แจงสำหรับผู้เชี่ยวชาญ: โปรดพิจารณาว่าข้อความที่เกี่ยวกับสถานภาพของผู้ตอบเหมาะสมหรือไม่อย่างไร

คำชี้แจง: โปรดใส่เครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
1	เพศ ☐ 1. ชาย ☐ 2. หญิง ☐ 3. LGBTQ+	✓			
2	อายุ ☐ 1. 18-22 ปี ☐ 2. 23-30 ปี ☐ 3. 31-40 ปี ☐ 4. 40 ปีขึ้นไป	✓			
3	สถานะ ☐ 1. โสด ☐ 2. มีแฟน/สมรส ☐ 3. หย่า/แยกทาง	✓			
4	อาชีพ ☐ 1. นักศึกษา ☐ 2. พนักงานบริษัท ☐ 3. ธุรกิจส่วนตัว ☐ 4. ข้าราชการ ☐ 5. อาชีพอิสระ ☐ 6. ว่างงาน	✓			คำตอบเดียวจะชัดเจนกว่าบางคนมีหลายอาชีพ

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
5	รายได้ส่วนบุคคล/เดือน ☐ 1. น้อยกว่า 15,000 บาท ☐ 2. 15,000-30,000 บาท ☐ 3. 30,001-50,000 บาท ☐ 4. มากกว่า 50,000 บาท	✓			

ส่วนที่ 2 ข้อมูลเกี่ยวกับพฤติกรรมการใช้บริการ และการรับรู้สื่อของผู้ตอบแบบสอบถาม

คำชี้แจงสำหรับผู้เชี่ยวชาญ: โปรดพิจารณาว่าข้อความที่เกี่ยวกับพฤติกรรมการใช้บริการและการรับรู้สื่อของผู้ตอบเหมาะสมหรือไม่อย่างไร

คำชี้แจง: โปรดใส่เครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
1	คุณมาร้านนี้บ่อยแค่ไหน ☐ 1. ครั้งแรก ☐ 2. 2-3 ครั้ง / เดือน ☐ 3. 1 ครั้ง / สัปดาห์ ☐ 4. 2-3 ครั้ง / สัปดาห์ ☐ 5. เกือบทุกวัน ☐ 6. มากกว่า 1 ครั้ง / เดือน	✓			
2	ช่วงเวลาที่คุณมาบ่อยที่สุด ☐ 1. 17.00-19.00 ☐ 2. 19.01-21.00 ☐ 3. 21.01-23.00 ☐ 4. 23.01–02.00	✓			

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
3	วันที่คุณมาบ่อยที่สุด ☐ 1. วันธรรมดา ☐ 2. วันเสาร์ - อาทิตย์	✓			
4	ปกติคุณมาที่ร้านกับใคร ☐ 1. คนเดียว ☐ 2. เพื่อน 2-3 คน ☐ 3. กลุ่มเพื่อนมากกว่า 3 คน ☐ 4. แฟน / คู่รัก ☐ 5. เพื่อนร่วมงาน	✓			ระบุคำตอบเป็นช่วงจะทำให้ผู้ตอบตอบสะดวกกว่า
5	ค่าใช้จ่ายเฉลี่ย / ครั้ง ☐ 1. น้อยกว่า 500 บาท ☐ 2. 500 - 1,000 บาท ☐ 3. 1,001 - 2,000 บาท ☐ 4. 2,001 – 3,000 บาท ☐ 5. มากกว่า 3,000 บาท	✓			
6	คุณเคยใช้บริการ ร้านอื่นหรือไม่ ☐ 1. ไม่เคย ☐ 2. 1 ร้าน ☐ 3. 2-3 ร้าน ☐ 4. 4-5 ร้าน ☐ 5. มากกว่า 5 ร้าน	✓			
7	ช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร ☐ 1. 06.00-12.00 ☐ 2. 12.01-18.00 ☐ 3. 18.01-24.00 ☐ 4. ไม่ต้องการข้อมูลข่าวสาร	✓			

ข้อเสนอแนะของผู้เชี่ยวชาญ

บังคับให้ตอบคำตอบเดียวเรื่องอาชีพ เพราะตอนวิเคราะห์ข้อมูลจะได้แยกกลุ่มได้ชัดเจน ไม่ทับซ้อนกัน

แบบฟอร์มที่ให้ผู้เชี่ยวชาญแสดงความคิดเห็น (3)
ร่างแบบสอบถามเพื่อการวิจัย (IOC)

เรื่อง “แบบสอบถามสำหรับงานวิจัยการ วิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน”

คำชี้แจง

1. แบบสอบถามฉบับนี้อยู่ในขั้นตอนการศึกษาเพื่อ วิเคราะห์พฤติกรรมลูกค้าสำหรับธุรกิจร้านอาหารกลางคืน

2. แบบสอบถามฉบับนี้มุ่งตรวจสอบเพื่อหาความเที่ยงตรง (Validity) โดยการวิเคราะห์ดัชนีความสอดคล้อง (Index of Item Objective Congruence: IOC) ของแบบสอบถามและข้อเสนอแนะของผู้เชี่ยวชาญ เพื่อนำไปปรับปรุงแบบสอบถามให้สมบูรณ์ยิ่งขึ้น

3. แบบสอบถามฉบับนี้มีทั้งหมด 2 ส่วน ประกอบด้วย

ส่วนที่ 1 ข้อมูลทั่วไป

ส่วนที่ 2 พฤติกรรมการใช้บริการ / การรับรู้สื่อ

4. ขอความกรุณาผู้ทรงคุณวุฒิหรือท่านผู้เชี่ยวชาญ ช่วยพิจารณาร่างแบบสอบถามว่ามีความสอดคล้องกับวัตถุประสงค์ของการวิจัยเรื่องนี้หรือไม่ ด้วยการให้คะแนนในแต่ละข้อคำถามในระบบ IOC โดยการทำเครื่องหมาย ✓ ลงในช่องว่าง

ชื่อผู้เชี่ยวชาญ ท่าน นิธิรุจน์ สวัสดิ์พรปิติ

ตำแหน่ง หุ้นส่วนเจ้าของร้าน PROMpt Charan Sanit Wong

เกณฑ์การให้คะแนนในระบบ IOC

1) ให้ 1 คะแนน เมื่อแน่ใจว่าข้อนั้นมีเนื้อหาที่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

2) ให้ 0 คะแนน เมื่อไม่แน่ใจว่าข้อนั้นมีเนื้อหาที่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

3) ให้ -1 คะแนน เมื่อแน่ใจว่าข้อนั้นมีเนื้อหาไม่สอดคล้องกับวัตถุประสงค์ที่ต้องการศึกษา

ผู้วิจัยขอความกรุณาท่านผู้ทรงคุณวุฒิหรือท่านผู้เชี่ยวชาญ ให้ข้อเสนอแนะหรือความคิดเห็นเพิ่มเติมในประเด็นที่ยังไม่สมบูรณ์ โดยการเขียนข้อเสนอแนะไว้ท้ายข้อความนั้นๆ ผู้วิจัยขอขอบพระคุณท่านที่กรุณาเสียสละเวลาในการตอบแบบสอบถามทุกข้ออย่างถูกต้องครบถ้วน

ส่วนที่ 1 ข้อมูลเกี่ยวกับสถานภาพ และข้อมูลทั่วปของผู้ตอบแบบสอบถาม

คำชี้แจงสำหรับผู้เชี่ยวชาญ: โปรดพิจารณาว่าข้อความที่เกี่ยวกับสถานภาพของผู้ตอบเหมาะสมหรือไม่อย่างไร

คำชี้แจง: โปรดใส่เครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
1	เพศ ☐ 1. ชาย ☐ 2. หญิง ☐ 3. LGBTQ+	✓			
2	อายุ ☐ 1. 18-22 ปี ☐ 2. 23-30 ปี ☐ 3. 31-40 ปี ☐ 4. 40 ปีขึ้นไป	✓			
3	สถานะ ☐ 1. โสด ☐ 2. มีแฟน/สมรส ☐ 3. หย่า/แยกทาง	✓			
4	อาชีพ ☐ 1. นักศึกษา ☐ 2. พนักงานบริษัท ☐ 3. ธุรกิจส่วนตัว ☐ 4. ข้าราชการ ☐ 5. อาชีพอิสระ ☐ 6. ว่างงาน	✓			

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
5	รายได้ส่วนบุคคล/เดือน ☐ 1. น้อยกว่า 15,000 บาท ☐ 2. 15,000-30,000 บาท ☐ 3. 30,001-50,000 บาท ☐ 4. มากกว่า 50,000 บาท	✓			

ส่วนที่ 2 ข้อมูลเกี่ยวกับพฤติกรรมการใช้บริการ และการรับรู้สื่อของผู้ตอบแบบสอบถาม

คำชี้แจงสำหรับผู้เชี่ยวชาญ: โปรดพิจารณาว่าข้อความที่เกี่ยวกับพฤติกรรมการใช้บริการและการรับรู้สื่อของผู้ตอบเหมาะสมหรือไม่อย่างไร

คำชี้แจง: โปรดใส่เครื่องหมาย ✓ ลงในช่องที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
1	คุณมาร้านนี้บ่อยแค่ไหน ☐ 1. ครั้งแรก ☐ 2. 2-3 ครั้ง / เดือน ☐ 3. 1 ครั้ง / สัปดาห์ ☐ 4. 2-3 ครั้ง / สัปดาห์ ☐ 5. เกือบทุกวัน ☐ 6. มากกว่า 1 ครั้ง / เดือน	✓			
2	ช่วงเวลาที่คุณมาบ่อยที่สุด ☐ 1. 17.00-19.00 ☐ 2. 19.01-21.00 ☐ 3. 21.01-23.00 ☐ 4. 23.01–02.00	✓			

ข้อ	ข้อคำถาม	ความคิดเห็นผู้เชี่ยวชาญ			ข้อเสนอแนะ
		1	0	-1	
3	วันที่คุณมาบ่อยที่สุด ☐ 1. วันธรรมดา ☐ 2. วันเสาร์ - อาทิตย์	✓			
4	ปกติคุณมาที่ร้านกับใคร ☐ 1. คนเดียว ☐ 2. เพื่อน 2-3 คน ☐ 3. กลุ่มเพื่อนมากกว่า 3 คน ☐ 4. แฟน / คู่รัก ☐ 5. เพื่อนร่วมงาน	✓			
5	ค่าใช้จ่ายเฉลี่ย / ครั้ง ☐ 1. น้อยกว่า 500 บาท ☐ 2. 500 - 1,000 บาท ☐ 3. 1,001 - 2,000 บาท ☐ 4. 2,001 – 3,000 บาท ☐ 5. มากกว่า 3,000 บาท	✓			
6	คุณเคยใช้บริการ ร้านอื่นหรือไม่ ☐ 1. ไม่เคย ☐ 2. 1 ร้าน ☐ 3. 2-3 ร้าน ☐ 4. 4-5 ร้าน ☐ 5. มากกว่า 5 ร้าน	✓			
7	ช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร ☐ 1. 06.00-12.00 ☐ 2. 12.01-18.00 ☐ 3. 18.01-24.00 ☐ 4. ไม่ต้องการข้อมูลข่าวสาร	✓			

ข้อเสนอแนะของผู้เชี่ยวชาญ

ช่วงเวลารับข้อมูลข่าวสารอาจจะให้แคบมากกว่านี้

ตารางที่ ข-1 สรุปผลของคะแนนความสอดคล้องในการคำนวนหาค่า IOC

ข้อที่	ข้อคำถาม	ประมาณค่าความคิดเห็นของผู้ทรงคุณวุฒิ			ค่า IOC	แปลผล
		1	2	3		
1	เพศ	1	1	1	1	ใช้ได้
2	อายุ	1	1	1	1	ใช้ได้
3	สถานะ	1	1	1	1	ใช้ได้
4	อาชีพ	1	1	1	1	ใช้ได้
5	รายได้ส่วนบุคคล / เดือน	1	1	1	1	ใช้ได้
6	คุณมาร้านนี้บ่อยแค่ไหน	1	1	1	1	ใช้ได้
7	ช่วงเวลาที่คุณมาบ่อยที่สุด	1	1	1	1	ใช้ได้
8	วันที่คุณมาบ่อยที่สุด	1	1	1	1	ใช้ได้
9	ปกติคุณมาที่ร้านกับใคร	1	1	1	1	ใช้ได้
10	ค่าใช้จ่ายเฉลี่ย / ครั้ง	1	1	1	1	ใช้ได้
11	คุณเคยใช้บริการ ร้านอื่นหรือไม่	1	1	1	1	ใช้ได้
12	ช่วงเวลาที่ต้องการรับข้อมูลข่าวสาร	1	1	1	1	ใช้ได้

ผลการประเมิน: ผู้ทรงคุณวุฒิทั้ง 3 ท่านให้คะแนนทุกข้อเป็น "1" (หมายถึงสอดคล้อง) ทั้งหมด ส่งผลให้ค่า IOC ของทุกข้อเท่ากับ 1 และผลการแปลความหมายคือ "ใช้ได้" ทุกข้อ แสดงว่าแบบสอบถามหรือเครื่องมือที่นำมาประเมินมีความตรงเชิงเนื้อหา และผู้เชี่ยวชาญทั้ง 3 ท่านเห็นตรงกันว่าทุกข้อคำถามมีความเหมาะสมและสอดคล้องกับวัตถุประสงค์ของการวิจัย

ประวัติผู้เขียน

ชื่อ	นายธัญวัฒน์ แก้วประดิษฐ์
ชื่อวิทยานิพนธ์	การเปรียบเทียบประสิทธิภาพอัลกอริทึมการจัดกลุ่มด้วยควอนตัมแมชชีนเลิร์นนิ่งสำหรับการแบ่งกลุ่มลูกค้าในการวิเคราะห์ธุรกิจร้านอาหาร
สาขาวิชา	วิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
ประวัติ	ประวัติการศึกษา พ.ศ. 2547 สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์ และเทคโนโลยี มหาวิทยาลัยอัสสัมชัญ

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ