**การศึกษาวิธีการแบ่งกลุ่มลูกค้าด้วยเทคนิคอาร์เอฟเอ็มร่วมกับการวิเคราะห์ข้อความ**

เอกปรียา ใบสนิ1\*, ศิริสรรพ เหล่าหะเกียรติ2

**บทคัดย่อ**

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาวิธีการแบ่งกลุ่มลูกค้า โดยใช้เทคนิคการเรียนรู้ของเครื่อง เพื่อจัดกลุ่มของข้อมูลร่วมกับการวิเคราะห์ข้อความ โดยประกอบไปด้วย 2 วิธีหลักดังนี้ 1. การแบ่งกลุ่มลูกค้าด้วยเทคนิคอาร์เอฟเอ็ม 2. การแบ่งกลุ่มด้วยเทคนิคอาร์เอฟเอ็มและการวิเคราะห์ข้อความด้วยเครื่องมือประมวลภาษาธรรมชาติ โมดูลเอ็นแอลทีเค (Natural Language Toolkit) โดยใช้ชุดข้อมูลการซื้อขายสินค้าของร้านค้าปลีกในประเทศอังกฤษที่เกิดขึ้นในช่วงวันที่ 1 ธ.ค. 2553 ถึง 9 ธ.ค. 2554 จำนวนข้อมูลทั้งหมด 541,909 แถว จากฐานข้อมูล UCI Machine Learning Repository ผู้วิจัยสนใจที่จะเปรียบเทียบผลของการจัดกลุ่มโดยใช้ฟีเจอร์ที่ได้จากเทคนิคอาร์เอฟเอ็มและเทคนิควิเคราะห์ข้อความมารวมกัน เปรียบเทียบกับการจัดกลุ่มโดยใช้ฟีเจอร์ที่ได้จากเทคนิคอาร์เอฟเอ็ม ในการเปรียบเทียบค่า Adjusted Rand Index และค่า Normalized Mutual Information ระหว่างวิธีที่ 1 และวิธีที่ 2 ซึ่งให้ค่า ARI เท่ากับ 0.5116 และ NMI เท่ากับ 0.3646 ค่า Confusion Matrix ที่ได้จากการเปรียบเทียบทั้ง 2 วิธี ผลที่ได้จากการแบ่งกลุ่มไม่สอดคล้องกันมากนัก จากการเปรียบเทียบการแบ่งกลุ่มด้วยเทคนิคอาร์เอฟเอ็มร่วมกับการวิเคราะห์ข้อความ มีประสิทธิภาพดีกว่าวิธีที่ใช้ฟีเจอร์ที่ได้จากเทคนิคอาร์เอฟเอ็มเพียงเทคนิคเดียว ในแง่ที่ผลการจัดกลุ่มที่ได้ จะให้ข้อมูลเชิงลึกของการสั่งซื้อร่วมกับข้อมูลในส่วนของศักยภาพของลูกค้าจากข้อมูลอาร์เอฟเอ็ม

**คำสำคัญ :** การแบ่งกลุ่มลูกค้า,การวิเคราะห์ข้อความ, การเรียนรู้ของเครื่อง, การจัดกลุ่มของข้อมูล,เทคนิคอาร์เอฟเอ็ม

1 หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

2 คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

\* Corresponding author:Tel.: 080-5488355 E-mail address: aekpreya.baisani@g.swu.ac.th

**The Study of Customer Segmentation by using RFM Model and Text Analytics**

Aekpreya Baisani1\*, Sirisup Laohakiat2

**Abstract**

This study aims to perform customer segmentation by utilizing clustering in machine learning based on RFM model and by using Text Analytics that consists of two main methods, as follows: (1) customer segmentation based on the RFM analysis Model Two. The process of customer segmentation was based on the RFM analysis model and Text Analytics operations with the NLTK (Natural Language Toolkit). This research used transactional dataset which contained all of the transactions occurring between 01/12/2010 and 09/12/2011, for a UK-based and registered non-store online retail with sample of 541,909 rows from UCI Machine Learning Repository. The results of the two methods were compared using clustering performance indices (adjusted rand index = 0.5116 and normalized mutual information = 0.3646) and the confusion matrix showed that the two methods were inconsistent. The results showed that the clusters obtained from customer segmentation were based on the RFM analysis model and Text Analytics method, which revealed insight on both the potential of customers based on RFM features, as well as the characteristics of the order in the clusters which was not present when using only RFM features alone.

**Keywords** : Customer Segmentation, Text Analytics, Machine Learning, Clustering, RFM Technique

1 Data Science, Faculty of Science, Srinakharinwirot University, Bangkok, 10110, Thailand

2 Faculty of Science, Srinakharinwirot University, Bangkok, 10110, Thailand

\* Corresponding author:Tel.: 080-5488355 E-mail address: aekpreya.baisani@g.swu.ac.th

**บทนำ**

 ปัจจุบันการตลาดสมัยใหม่หมุนเวียนเปลี่ยนไปอย่างรวดเร็วและแตกต่างไปจากการทำตลาดในอดีตอย่างสิ้นเชิง การแบ่งกลุ่มลูกค้า เพื่อให้สามารถสื่อสารกับลูกค้าแต่ละกลุ่มจึงมีรูปแบบการสื่อสารที่มีเป้าหมายแตกต่างกันออกไป และมีความสำคัญมากในการพัฒนาโมเดลธุรกิจ เพื่อเพิ่มประสิทธิภาพในการดำเนินกิจกรรมทางการตลาด การจัดกลุ่มลูกค้า (Customer segmentation) เป็นวิธีการจัดกลุ่มโดยที่ลูกค้าที่มีพฤติกรรมใกล้เคียงกันหรือเหมือนกันจะถูกจัดไว้ในกลุ่มเดียวกัน ส่วนลูกค้าที่มีพฤติกรรมแตกต่างหรือไม่เหมือนกันก็จะถูกจัดไว้คนละกลุ่ม การจัดกลุ่มลูกค้าทำให้ธุรกิจสามารถทราบถึงลักษณะหรือพฤติกรรมของลูกค้าของตนเองในแต่ละกลุ่มได้ เมื่อธุรกิจเข้าใจพฤติกรรมหรือลักษณะของลูกค้าในแต่ละกลุ่มก็จะสามารถเสนอสินค้าหรือบริการที่ตรงต่อพฤติกรรมของลูกค้าได้ หากนำเทคนิคคลัสเตอริงหรือการเรียนรู้ของเครื่องประเภทการจัดกลุ่มของข้อมูล ซึ่งเป็นโมเดลที่เอาไว้ใช้การจัดกลุ่มของข้อมูล จะทำให้สามารถแบ่งกลุ่มของผลิตภัณฑ์ตามความประสงค์ของลูกค้าได้ ก็จะยิ่งส่งเสริมให้งานนั้นๆเกิดขึ้นได้อย่างมีประสิทธิภาพ ดังนั้นงานวิจัยนี้จึงศึกษาวิธีการแบ่งกลุ่มลูกค้าจากข้อมูลการซื้อขายสินค้าโดยใช้วิธีการวิเคราะห์ข้อความและเทคนิคการเรียนรู้ของเครื่อง ประเภทการจัดกลุ่มของข้อมูล เพื่อนำผลลัพธ์ที่ได้มาวิเคราะห์พฤติกรรมลูกค้าและกำหนดเป็นกลยุทธ์ทางการตลาดต่อไป

 แม้ว่าจะมีงานวิจัยจำนวนไม่น้อยที่ศึกษาเกี่ยวกับการจัดกลุ่มลูกค้า [1, 2, 3, 4, 5] การศึกษาเหล่านั้นอาจแบ่งเป็น 2 แนวทางได้แก่ แนวทางที่ 1 ผู้ศึกษาจะอาศัยข้อมูลจากใบสั่งซื้อ เพื่อจัดกลุ่มข้อมูลโดยอาศัยความคล้ายคลึงกันของสินค้าที่สั่งซื้อของลูกค้าแต่ละราย ส่วนในแนวทางที่ 2 ผู้ศึกษาจะจัดกลุ่มโดยอาศัยแบบจำลองอาร์เอฟเอ็ม ซึ่งเป็นการจัดกลุ่มลูกค้า โดยจำแนกตามศักยภาพของลูกค้าผ่านแบบจำลองอาร์เอฟเอ็ม โดยวิเคราะห์จากใบสั่งซื้อ การแบ่งกลุ่มเช่นนี้ อาจทำให้เราไม่สามารถมองเห็นภาพรวมของลูกค้าได้ชัดเจนในทุกๆแง่มุมได้ดีพอ ในการศึกษาที่ผ่านมาผู้วิจัยจึงทำการจัดกลุ่มลูกค้าโดยอาศัยข้อมูลจากทั้งสองส่วน อันได้แก่ ข้อมูลการสั่งซื้อสินค้าในใบสั่งซื้อ ร่วมกับการใช้แบบจำลองอาร์เอฟเอ็ม เพื่อวัดศักยภาพของลูกค้า มาประกอบร่วมกัน เพื่อหาแนวทางในการสร้างแบบจำลองการจัดกลุ่มลูกค้า เพื่อให้มีประสิทธิภาพในการจัดกลุ่มสูงสุด นอกจากนี้ ในการวิเคราะห์ข้อมูลการสั่งซื้อสินค้าในใบสั่งซื้อ ผู้วิจัยยังใช้เทคนิคการวิเคราะห์ข้อความร่วมด้วย เพื่อช่วยในการจัดกลุ่มสินค้า เพื่อให้แบบจำลองสามารถจัดกลุ่มได้อย่างมีความแม่นยำมากขึ้น

**วิธีดำเนินการ**

**ขั้นตอนที่ 1** **:** แนะนำชุดข้อมูลที่ใช้ในการศึกษานี้

ผู้วิจัยนำข้อมูลเกี่ยวกับการซื้อขายสินค้ามาวิเคราะห์ในการทำวิจัยครั้งนี้ โดยใช้ชุดข้อมูลการซื้อขายสินค้าของร้านค้าปลีกในประเทศอังกฤษที่เกิดขึ้นในช่วงวันที่ 1 ธ.ค. 2553 ถึง 9 ธ.ค. 2554 จำนวนข้อมูลทั้งหมด 541,909 แถว จากฐานข้อมูล UCI Machine Learning Repository [6] ประกอบด้วยตัวแปรอิสระ 8 ตัวแปร

**ขั้นตอนที่ 2** **:** การนำเข้าข้อมูล ตรวจสอบข้อมูล และพิจารณาข้อมูล

ใช้ภาษาไพทอนในการวิเคราะห์ข้อมูลและการเรียนรู้ของเครื่อง เริ่มต้นด้วยการนำเข้าโมดูลสำคัญสำหรับการสร้างแบบจำลอง ต่อมานำเข้าไฟล์ข้อมูลและข้อมูลที่ใช้สำหรับสร้างแบบจำลอง เริ่มกระบวนการตรวจสอบและสำรวจข้อมูลเบื้องต้น เพื่อหาข้อมูลเชิงลึกจากข้อมูลธุรกรรมโดยใช้ไลบรารี่ Pandas, Numpy และ Dataprep.eda สำรวจข้อมูลกับข้อมูลคอลัมน์ชื่อสินค้า เพื่อพยายามแบ่งกลุ่มสินค้าออกเป็นกลุ่ม แล้วตั้งชื่อหมวดหมู่ให้แต่ละกลุ่ม จากนั้นหาคำซ้ำที่ปรากฏในแต่ละข้อความโดยใช้ไลบรารี่ Pycaret

**ขั้นตอนที่ 3** **:** การวิเคราะห์ข้อความเพื่อสร้างกลุ่มสินค้า

 การจัดเตรียมข้อมูล ผู้วิจัยได้ตัดข้อมูลบางแถวออกไม่นำข้อมูลต่อไปนี้มาร่วมวิเคราะห์ ได้แก่ ใบคำสั่งซื้อมีการยกเลิกหรือแก้ไข, ข้อมูลที่ไม่มีรหัสลูกค้า และทำการกำหนดช่วงวันที่ของข้อมูลที่ใช้ในการวิเคราะห์ข้อมูลเท่ากับ 1 ปี ต่อมาวิเคราะห์ข้อความจากชื่อสินค้า ใช้ข้อมูลคอลัมน์ชื่อสินค้า เลือกเฉพาะสินค้าที่ไม่ซ้ำโดยมีรหัสสินค้าเป็นตัวบ่งบอก เนื่องจากชื่อสินค้าอยู่ในรูปแบบภาษาปกติของมนุษย์จึงจำเป็นต้องแปลงชื่อสินให้อยู่ในรูปแบบที่คอมพิวเตอร์สามารถเข้าใจ โดยผ่านกระบวนการวิเคราะห์ข้อความ เริ่มจากนำข้อมูลจากคอลัมน์สินค้ามาตัดคำชื่อสินค้าภาษาอังกฤษและกำจัดคำหยุด โดยอาศัยพจนานุกรมคำศัพท์มาช่วยในกระบวนการตัดคำ สำหรับงานวิจัยนี้ผู้วิจัยใช้เครื่องมือประมวลภาษาธรรมชาติ โมดูล NLTK (Natural Language Toolkit) จากนั้นหาคำหลักที่พบบ่อยที่สุด สำหรับนำไปใช้กำหนดประเภทให้กับชื่อสินค้า ผู้วิจัยใช้เทคนิค K-Means Clustering ในการแบ่งกลุ่ม เพื่อกำหนดประเภทให้กับชื่อสินค้า ใช้ Silhouette Score เป็นเทคนิคที่ใช้วัดตัวอย่างนั้นๆมีความเหมือนกับกลุ่มที่อยู่มากเพียงใด เมื่อเทียบกับกลุ่มอื่นๆ และทำการรวมข้อมูลประเภทสินค้าโดยจัดกลุ่มตามแอตทริบิวต์รหัสลูกค้า

**ขั้นตอนที่ 4** **:** การสร้างคุณลักษณะข้อมูลสำหรับแบบจำลองอาร์เอฟเอ็ม

 ผู้วิจัยทำการรวมข้อมูลการซื้อขายสินค้าโดยจัดกลุ่มตามแอตทริบิวต์รหัสลูกค้า จากนั้นสร้างตัวแปรใหม่ขึ้น โดยมีรายละเอียดการคำนวณดังนี้

1. จำนวนวันที่ลูกค้าซื้อล่าสุด (Recency) ดังสมการที่ 1

(1)

 Ri = (Dn - Dl)-1

 Ri หมายถึง ค่ารีเซนซีของลูกค้าที่ i

 Dn หมายถึง วันที่กำหนดให้เป็นวันที่ปัจจุบัน

 Dl หมายถึง วันที่ลูกค้าซื้อสินค้าล่าสุด

 เพื่อให้ค่า Recency มีทิศทางเดียวกับค่า ความถี่และค่าจำนวนเงิน โดยค่ามากหมายถึงการที่ลูกค้ามีศักยภาพมาก เราจึงใช้ส่วนกลับของเวลาครั้งล่าสุด แทนที่จะเป็นค่าเวลาล่าสุดตามปกติ

1. ความบ่อยในการซื้อ (Frequency) ดังสมการที่ 2

(2)

Fi =

$$\sum\_{i=m}^{n}Ii$$

 Fi หมายถึง ค่าฟรีเควนซีของลูกค้าที่ i

 I หมายถึง จำนวนใบสั่งซื้อ

1. คำนวณจำนวนเงินที่ลูกค้าใช้จ่าย (Monetary) ดังสมการที่ 3

(3)

Mi =

$$\sum\_{i=m}^{n}Ti$$

 Mi หมายถึง โมนทารี ผลรวมยอดซื้อสินค้าทุกครั้งเข้าด้วยกัน

 T หมายถึง จำนวนเงินที่ลูกค้าจ่าย

 ดังนั้นลูกค้าแต่ละรายจะมีค่า RFM เป็นของตนเอง ซึ่งคุณลักษณะตัวนี้ จะเป็นไปตามลูกค้าแต่ละราย มิใช่คุณลักษณะของแต่ละรายการสั่งซื้อ

**ขั้นตอนที่ 5** **:** การสร้างแบบจำลองสำหรับจัดกลุ่มข้อมูล

ในการศึกษานี้ ผู้วิจัยเปรียบเทียบแบบจำลอง 2 แบบ โดยแบบจำลองรูปแบบที่ 1 จะเป็นการจัดกลุ่มด้วยเทคนิคอาร์เอฟเอ็ม ส่วนในแบบจำลองที่ 2 จะเป็นการจัดกลุ่มข้อมูลจากใบสั่งซื้อร่วมกับการจัดกลุ่มศักยภาพของลูกค้าด้วยเทคนิคอาร์เอฟเอ็ม โดยแบบจำลองแต่ละแบบมีรายละเอียดดังนี้

* แบบจำลองที่ 1 การจัดกลุ่มด้วยการใช้ข้อมูลอาร์เอฟเอ็ม ผู้วิจัยนำฟีเจอร์ที่ได้จากขั้นตอนการวิเคราะห์ข้อมูลลูกค้าด้วยเทคนิคอาร์เอฟเอ็ม จากนั้นทำการจัดกลุ่มของข้อมูลด้วยอัลกอริทึมเคมีน
* แบบจำลองที่ 2 การจัดกลุ่มด้วยการใช้ข้อมูลอาร์เอฟเอ็มและการวิเคราะห์ข้อความ ผู้วิจัยนำฟีเจอร์ที่ได้จากขั้นตอนการวิเคราะห์ข้อมูลลูกค้าด้วยเทคนิคอาร์เอฟเอ็มร่วมกับการแบ่งกลุ่มสินค้าด้วยเทคนิควิเคราะห์ข้อความ
* แบบจำลองที่ 3 การสร้างคะแนนอาร์เอฟเอ็มร่วมกับการจัดกลุ่มข้อมูลตามใบสั่งซื้อ

**ผลการวิจัยและอภิปรายผลการวิจัย**

 การทดลองนี้ใช้เว็บแอพพลิเคชั่นจูปิเตอร์โน๊ตบุ๊ค สำหรับสร้างตัวแบบแบ่งกลุ่มลูกค้า มีจำนวนทั้งสิ้น 3,155 เรคคอร์ด ประกอบด้วย 2 วิธีดังนี้

 **วิธีที่ 1** การจัดกลุ่มด้วยเทคนิคอาร์เอฟเอ็ม ประกอบด้วยคุณลักษณะดังนี้

 1) จำนวนวันที่ลูกค้าซื้อล่าสุด (Recency)

 2) ความบ่อยในการซื้อ (Frequency)

 3) จำนวนเงินที่ลูกค้าใช้จ่าย (Monetary)

 จากการจัดกลุ่มของข้อมูลด้วยเทคนิคอาร์เอฟเอ็มจะเป็นการแบ่งกลุ่มของลูกค้าออกเป็น 5 คลัสเตอร์ โดยได้ผลลัพธ์ดังนี้ คลัสเตอร์ที่ 1 มีจำนวน 616 เรคคอร์ด คลัสเตอร์ที่ 2 มีจำนวน 408 เรคคอร์ด คลัสเตอร์ที่ 3 มีจำนวน 389 เรคคอร์ด คลัสเตอร์ที่ 4 มีจำนวน 1,283 เรคคอร์ด และคลัสเตอร์ที่ 5 มีจำนวน 459 เรคคอร์ด

 **วิธีที่ 2** การจัดกลุ่มด้วยเทคนิคอาร์เอฟเอ็มและการวิเคราะห์ข้อความ

 ประกอบด้วยคุณลักษณะ ดังนี้

 1) จำนวนวันที่ลูกค้าซื้อล่าสุด (Recency)

 2) ความบ่อยในการซื้อ (Frequency)

 3) จำนวนเงินที่ลูกค้าใช้จ่าย (Monetary)

 4) กลุ่มสินค้าที่ 1 (categ\_0)

 5) กลุ่มสินค้าที่ 2 (categ\_1)

 6) กลุ่มสินค้าที่ 3 (categ\_2)

 7) กลุ่มสินค้าที่ 4 (categ\_3)

 เช่นเดียวกับวิธีที่ 1 เราเลือกจัดกลุ่มข้อมูลเป็น 5 คลัสเตอร์ ได้ผลลัพธ์ดังนี้ คลัสเตอร์ที่ 1 มีจำนวน 813 เรคคอร์ด คลัสเตอร์ที่ 2 มีจำนวน 1,298 เรคคอร์ด คลัสเตอร์ที่ 3 มีจำนวน 627 เรคคอร์ด คลัสเตอร์ที่ 4 มีจำนวน 219 เรคคอร์ด และคลัสเตอร์ที่ 5 มีจำนวน 198 เรคคอร์ด

**การเปรียบเทียบผลของการจัดกลุ่มทั้งสองแบบจำลอง**

ในการเปรียบเทียบค่า Adjusted Rand Index และค่า Normalized Mutual Information ระหว่างแบบจำลองที่ 1 กับ 2 เราได้ค่าดังตารางที่ 1

**ตารางที่ 1** ค่า Adjusted Rand index และค่า Normalized Mutual Information ที่ได้จากการเปรียบเทียบกลุ่มข้อมูลจากทั้งสองแบบจำลอง

|  |  |
| --- | --- |
| **ตัววัด** | **ค่า** |
| Adjusted Rand index | 0.5116 |
| Normalized Mutual Information | 0.3646 |

โดยมีค่า Confusion Matrix ดังนี้

**ตารางที่ 2** ค่า Confusion Matrix ได้จากการเปรียบเทียบแบบจำลองทั้ง 2 โดย M1 M2 หมายถึงแบบจำลองที่ 1 และ 2 ตามลำดับ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  **M1** **M2** | **0** | **1** | **2** | **3** | **4** |
| 0 | 159 | 217 | 210 | 3 | 224 |
| 1 | 100 | 0 | 0 | 1198 | 0 |
| 2 | 230 | 114 | 114 | 14 | 155 |
| 3 | 71 | 40 | 29 | 44 | 35 |
| 4 | 56 | 37 | 36 | 24 | 45 |

 จากตารางที่ 1 และ 2 จะเห็นได้ว่า ผลที่ได้จากการแบ่งกลุ่ม ไม่สอดคล้องกันมากนัก ยกเว้น ในกลุ่ม (1, 3) ซึ่งสมาชิกในกลุ่มทั้งสองนี้ จะแทบสอดคล้องกันทั้งสิ้น ในส่วนกลุ่มอื่นๆ จะค่อนข้างกระจัดกระจายไปในคลัสเตอร์ต่างๆ เมื่อเราไปพิจารณาดูค่า จุดศูนย์กลางของแต่ละคลัสเตอร์ในภาพที่ 1 และ 2 ดังนี้



**ภาพที่ 1** จุดศูนย์กลางของแต่ละคลัสเตอร์ในแบบจำลองที่ 1



**ภาพที่ 2** จุดศูนย์กลางของแต่ละคลัสเตอร์ในแบบจำลองที่ 2

 จากผลของจุดศูนย์กลางที่ได้ จะเห็นว่า ค่า Recency ไม่มีบทบาทในการจัดกลุ่มมากนัก เนื่องจากค่า Recency ของจุดศูนย์กลางของทุกกลุ่มจะมีค่าใกล้เคียงกัน ในขณะที่ ฟีเจอร์ที่มีค่าแตกต่างกันอย่างเด่นชัดจะเป็นค่า Monetary ที่ในแต่ละกลุ่มของทั้งสองแบบจำลองมีค่าแยกกันอย่างชัดเจน โดยจะเห็นว่า จุดศูนย์กลางกลุ่ม 3 ในแบบจำลองที่ 1 จะมีค่าใกล้เคียงกับ จุดศูนย์กลางกลุ่ม 1 ในแบบจำลองที่ 2 ซึ่งสอดคล้องกับผลที่ได้จากการหา Confusion Matrix ในตารางที่ 2

 ทั้งนี้ จะเห็นได้ว่า ในการจัดกลุ่มด้วยการใช้แบบจำลองที่ 2 นอกจากจะสะท้อนศักยภาพของลูกค้าแล้ว ยังสามารถจำแนกลักษณะการสั่งซื้อสินค้าหลัก ในแต่ละกลุ่มได้อีกด้วย เช่น ในกลุ่มที่ 0 จะเป็นพวกที่มีลักษณะ คือ สั่งสินค้าทั้ง 5 กลุ่มในปริมาณที่เท่าๆกัน แต่จะมียอดสั่งซื้อต่ำ และมียอดการใช้จ่ายต่ำ ในขณะที่ในกลุ่มที่มียอดการสั่งซื้อสูง ในกลุ่มที่ 4 จะนิยมสั่งซื้อสินค้า ใน category\_0 มากที่สุด ซึ่งจากข้อมูล Insight ในส่วนนี้ ทำให้เราสามารถวางแผนการสั่งสินค้าให้สอดคล้องกับความต้องการของลูกค้าที่มียอดการสั่งซื้อสูงได้

**สรุปผลการวิจัย**

 งานวิจัยนี้เป็นการศึกษาวิธีการแบ่งกลุ่มลูกค้าด้วยเทคนิคอาร์เอฟเอ็มร่วมกับเทคนิควิเคราะห์ข้อความ โดยศึกษา 2 วิธีดังนี้ 1) การจัดกลุ่มด้วยเทคนิคอาร์เอฟเอ็ม โดยใช้ฟีเจอร์ที่ได้จากขั้นตอนการวิเคราะห์ข้อมูลลูกค้าด้วยเทคนิคอาร์เอฟเอ็ม 2) การจัดกลุ่มด้วยเทคนิคอาร์เอฟเอ็มและการวิเคราะห์ข้อความ ใช้ฟีเจอร์ที่ได้จากขั้นตอนการวิเคราะห์ข้อมูลลูกค้าด้วยเทคนิคอาร์เอฟเอ็มร่วมกับการแบ่งกลุ่มสินค้าด้วยเทคนิควิเคราะห์ข้อความ ผลการทดลองพบว่าวิธีการจัดกลุ่มด้วยเทคนิคอาร์เอฟเอ็มและการวิเคราะห์ข้อความมีประสิทธิภาพที่ดีกว่า เพราะสามารถสะท้อนให้เห็นกลุ่มลูกค้าจากการซื้อขายสินค้า อีกทั้งยังสะท้อนให้เห็นกลุ่มสินค้าที่ใกล้เคียงกันอีกด้วย งานวิจัยในอนาคต ผู้วิจัยจะศึกษาเทคนิควิธีการเรียนรู้ของเครื่อง ประเภทการจัดกลุ่มของข้อมูลแบบอื่นๆที่สามารถนำมาสร้างตัวแบบเพื่อใช้ในการแบ่งกลุ่มลูกค้าให้มีประสิทธิภาพมากยิ่งขึ้น

**กิตติกรรมประกาศ**

การจัดทำวิจัยได้รับการสนับสนุนจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ในการนำเสนอผลงานวิจัย ผู้วิจัยจึงขอขอบคุณมา ณ ที่นี้

**เอกสารอ้างอิง**

[1] Maryani I, Riana D, Astuti RD, Ishaq A, Pratama EA, editors. Customer Segmentation based on RFM model and Clustering Techniques With K-Means Algorithm. 2018Third International Conference on Informatics and Computing (ICIC); 2018: IEEE.

[2] Chen D, Sain SL, Guo K. Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. Journal of Database Marketing & Customer Strategy Management. 2012;19(3):197-208.

[3] Kansal T, Bahuguna S, Singh V, Choudhury T, editors. Customer segmentation using k-means clustering. 2018International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS); 2018: IEEE.

[4] Hu X, Zhang H, Wu X, Chen J, Xiao Y, Xue Y, et al., editors. A Novel Approach for Customer Segmentation Based on Biclustering. International Conference on Web Information Systems Engineering; 2013: Springer.

[5] Wang B, Miao Y, Zhao H, Jin J, Chen Y. A biclustering-based method for market segmentation using customer pain points. Engineering Applications of Artificial Intelligence. 2016;47:101-9.

[6] Chen, D. (2015). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Public Analytics group,London South Bank University, School of Engineering