

การจัดกลุ่มผู้บริโภคโดยใช้พฤติกรรมการใช้โทรศัพท์มือถือเพื่อการส่งเสริมการขาย
Customer Clustering Using Mobile Phone Usage Behavior For Promotion

นิลวัจน์ บุรณปรีชา และ นิเวศ จิระวิชิตชัย

หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม

Submitted 20/7/19 ; Revised 20/8/19 ; Accepted 29/8/19

บทคัดย่อ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อการจัดกลุ่มผู้บริโภคโดยศึกษาจากพฤติกรรมการใช้โทรศัพท์มือถือเพื่อการส่งเสริมการขาย งานวิจัยเชิงทดลองนี้ใช้เทคนิค เคมีน คลัสเตอร์ในการจัดกลุ่มโดยศึกษาจากกลุ่มตัวอย่างจำนวน 1,000 ราย ผลการศึกษาพบว่าสามารถจัดกลุ่มลูกค้าได้เป็น 3 กลุ่ม ดังนี้ คือ กลุ่มที่ 1 เป็นกลุ่มลูกค้าที่ใช้บริการเครือข่ายที่มากที่สุดในจำนวนกลุ่มลูกค้าทั้งหมด และมีค่าใช้จ่ายบริการมากที่สุด กลุ่มที่ 2 เป็นกลุ่มลูกค้าที่ใช้บริการเครือข่ายระดับปานกลาง และชื่นชอบใช้บริการทางโทรศัพท์ที่หลากหลาย และกลุ่มสุดท้ายคือกลุ่มที่ 3 เป็นกลุ่มที่ไม่ค่อยได้ใช้บริการเครือข่ายหรือเลือกใช้บริการใด ๆ ปัจจัยสำคัญดังกล่าวสามารถนำมาใช้ในการกำหนดแนวทางและกลยุทธ์ในการทำรายการส่งเสริมการขายของผู้ให้บริการเครือข่าย

คำสำคัญ : การจัดกลุ่ม, เคมีน คลัสเตอร์, ส่งเสริมการขาย

Abstract

The purpose of this research is customer clustering using mobile phone usage behavior for promotion. This experimental research uses k-mean clustering techniques to segmentation by studying a sample of 1,000 people. The results of the study showed that customers can be classified into 3 groups as follows: group 1 is the customer group that uses the most network services in the total number of customers and has the highest service cost. Group 2 is a customer group that uses moderate network services and likes to use a variety of telephone services. Finally, the last of group 3 is a group that does not use the network service or choose to use any service. Such important factors can be used to determine the guidelines and strategies for the promotion of network service providers.

Keyword : Clustering, K-means Cluster, Promotion

***ผู้ประสานงานหลัก (Corresponding Author)**

E-mail: neelawat.wat@gmail.com



1. บทนำ

การจัดกลุ่มแบ่งประเภท (Clustering) เพื่อจัดการกับข้อมูลให้อยู่เป็นหมวดหมู่เพื่อการนำข้อมูลไปทำการพิจารณาหรือวิเคราะห์เพื่อประโยชน์อื่น โดยปกติแล้วถ้าข้อมูลที่มีจำนวนน้อยก็สามารถทำให้เกิดการจัดแบ่งประเภทได้โดยอาศัยประสาทสัมผัสของผู้จัดแบ่งหรือคุณสมบัติของข้อมูล ซึ่งอาจจะจัดแบ่งตามความเหมือน หรือความแตกต่างของข้อมูลนั้น [1] ก็จะสามารถบอกได้ว่าข้อมูลนั้น มีการจัดแบ่งได้กี่ประเภทและมีอะไรบ้างอย่างถูกต้อง เกิดความผิดพลาดน้อยและใช้เวลาไม่นาน แต่ถ้าต้องมีการจัดแบ่งประเภทของข้อมูลที่มีจำนวนมาก จะทำให้เกิดปัญหาขึ้นเพราะถ้าการใช้ประสาทสัมผัสของผู้จัดแบ่งหรือแม้แต่การใช้คุณสมบัติของข้อมูลเองอาจทำให้เกิดความสับสนมีโอกาสผิดพลาดสูงและต้องใช้เวลาในการจัดแบ่งนานขึ้นอีก [2] โดยใช้วิธีการวิเคราะห์แบบ K-Means Cluster Analysis

ในปัจจุบัน มีปริมาณการใช้บริการข้อมูลโทรคมนาคมของประชากรที่มากขึ้น เช่น เน้นการโทรเน้นการใช้อินเทอร์เน็ต เน้นการใช้ YouTube และเน้นการใช้ Line เป็นต้น จึงทำให้ผู้ให้บริการเครือข่ายทำการส่งเสริมการขายเพื่อดึงดูดให้กลุ่มลูกค้าให้สนใจมาเลือกใช้บริการให้เหมาะสมกับปริมาณการใช้และความต้องการของกลุ่มบุคคลต่าง ๆ

ด้วยเหตุนี้ผู้วิจัยจึงสนใจที่จะศึกษากรณีตัวอย่างพฤติกรรมการใช้โทรศัพท์มือถือของผู้บริโภค การศึกษาเรื่องนี้เป็นการศึกษาหาแนวทางในการจัดกลุ่มพฤติกรรมการใช้โทรศัพท์มือถือของผู้บริโภค โดยใช้วิธีการวิเคราะห์แบบ K-Means Cluster Analysis เหมาะสำหรับข้อมูลที่มีจำนวนมาก และคำนวณได้รวดเร็วเพื่อใช้เป็นแนวทางให้กับผู้ให้บริการเครือข่ายโทรศัพท์มือถือหรือองค์กรอื่น ๆ ที่เกี่ยวข้อง ได้รับทราบ รวมทั้งใช้เป็นแนวทางในการวางกลยุทธ์ต่าง ๆ ในการทำการส่งเสริมการขายเพื่อดึงดูดกลุ่มผู้บริโภคและกำหนดแนวทางในการบริหารจัดการกลุ่มพฤติกรรมข้อมูลผู้บริโภค

2. วัตถุประสงค์

2.1. เพื่อจัดกลุ่มผู้บริโภคและหาลักษณะปัจจัยสำคัญประจำกลุ่มของผู้บริโภคในแต่ละกลุ่ม โดยใช้เทคนิคเคมีน คลัสเตอร์

2.2. เพื่อทดสอบประสิทธิภาพการจัดกลุ่มผู้บริโภคโดยใช้พฤติกรรมการใช้โทรศัพท์มือถือเพื่อการส่งเสริมการขายโดยใช้เทคนิคเคมีน คลัสเตอร์

3. สมมุติฐาน

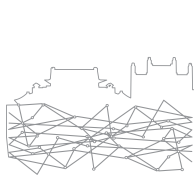
จัดกลุ่มผู้บริโภคที่มีพฤติกรรมการใช้โทรศัพท์มือถือที่อยู่ภายใน Cluster เดียวกันจะได้ปัจจัยสำคัญของกลุ่มผู้บริโภค แล้วนำมาข้อมูลที่ได้ใช้เป็นข้อมูลในการสนับสนุนในการกำหนดแนวทางและกลยุทธ์ในการทำการส่งเสริมการขายของผู้ให้บริการเครือข่ายได้

4. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ประกอบไปด้วยองค์ความรู้ในการวิเคราะห์กลุ่ม (Clustering analysis) K-Means Clustering และงานวิจัยที่เกี่ยวข้อง ดังนั้นผู้วิจัยได้มีการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการวิจัยมีรายละเอียดดังนี้

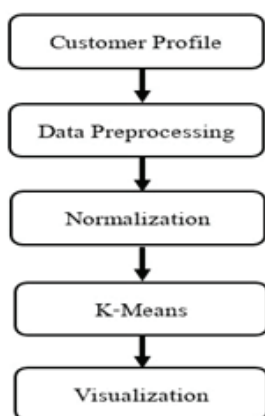
การวิเคราะห์กลุ่ม (Clustering analysis) คือ เทคนิคในการเรียนรู้แบบไม่มีผู้สอน ซึ่งมีเป้าหมายเพื่อการจำแนกกลุ่มข้อมูลที่มีคุณลักษณะคล้ายกันอยู่ในกลุ่มเดียวกัน โดยข้อมูลแต่ละกลุ่มจะถูกเรียกว่า คลัสเตอร์ (Cluster) การวิเคราะห์หรือจำแนกกลุ่มข้อมูลนั้นสามารถแบ่งออกได้เป็น 2 ประเภท ได้แก่ วิธีการแบบลำดับขั้น (Hierarchical algorithms) และวิธีการแบบไม่เป็นลำดับขั้น (Non-hierarchical - algorithms) [3]

K-Means Clustering คือ วิธีการจำแนกกลุ่มข้อมูลด้วยวิธีการแบ่งข้อมูลอัตโนมัติตามค่า k ที่กำหนด โดยกระบวนการทำงานเลือกค่า k เริ่มต้นสำหรับเป็นค่ากลางในการจัดกลุ่ม และปรับค่าตามกระบวนการดังนี้ 1) เลือกข้อมูล d_i สำหรับวัดระยะห่างกับค่า K เริ่มต้นทุกค่า 2) กำหนดชุดข้อมูล d_i ให้กับ K ที่ใกล้ที่สุด และปรับค่า K ใหม่ให้เป็นค่ากลางของกลุ่มข้อมูล และหยุดเมื่อค่า K ไม่เปลี่ยนแปลง [4] นอกจากนี้ผู้วิจัยได้ศึกษางานวิจัยที่เกี่ยวข้องใช้เป็นแนวทางของการใช้เทคนิค K-Means Cluster การศึกษาพฤติกรรมสื่อสารในบริบทออนไลน์ของกลุ่มดิจิทัลอิมมิгранท์ เพื่อจัดกลุ่มและหาคุณลักษณะของกลุ่มดิจิทัลอิมมิгранท์ และเพื่อศึกษาปัญหาทางความสัมพันธ์กับสังคมในการใช้สื่อใหม่ของกลุ่มดิจิทัลอิมมิгранท์ [5] การสร้างโมเดล การจัดการระบบนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูลเพื่อศึกษาและวิเคราะห์หาปัจจัยที่มีผลต่อความพึงพอใจของนักศึกษาในการสร้างระบบบริหารความสัมพันธ์นักศึกษา [6] ดังนั้นการใช้ K-means clustering จึงสามารถนำมาใช้งานเมื่อทราบจำนวนกลุ่มที่ต้องการจำแนกที่แน่นอน ทั้งนี้การวัดระยะห่างระหว่างข้อมูลสามารถใช้วิธีการคำนวณระยะห่างได้หลากหลาย



3. วิจัยดำเนินการวิจัย

งานวิจัยนี้เป็นการศึกษาพฤติกรรมการใช้โทรศัพท์มือถือของผู้บริโภค และจัดกลุ่ม Clustering โดยใช้เทคนิค K means พบว่ามีประชากรผู้ใช้งานบริการโทรศัพท์เครือข่ายมือถือเป็นจำนวนมาก ผู้วิจัยจึงได้ทำการศึกษาและรวบรวมข้อมูลที่เกี่ยวข้องกับการใช้โทรศัพท์มือถือตามกรอบแนวคิดในงานวิจัย ดังแสดงในภาพที่ 1



ภาพที่ 1 กรอบแนวคิดในงานวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

ข้อมูลได้จากฐานข้อมูลกลางของบริษัท IBM Knowledge Center เป็นฐานข้อมูลตัวอย่างของบริษัทโทรคมนาคมของต่างประเทศเป็นฐานข้อมูลในการวิเคราะห์ซึ่งเป็นข้อมูลลูกค้า 1,000 คน ที่ใช้บริการโทรคมนาคม (กรณีศึกษา) ตัวแปรที่ใช้ในงานวิจัย 14 ตัวแปร ดังนี้

การโทรทางไกล (long distance) หมายถึง การใช้บริการโทรทางไกลผ่านรหัส

โทรศัพท์ฟรี (toll free) หมายถึง ไม่เสียค่าธรรมเนียม

อุปกรณ์เสริม (equipment) หมายถึง อุปกรณ์เสริมสำหรับใช้งานควบคู่กันไปเพื่อเพิ่มความหลากหลายในการใช้งาน

บัตรโทรศัพท์ (calling card) หมายถึง บัตรที่เอาไว้ใช้ในการติดต่อสื่อสาร

อินเทอร์เน็ตไร้สาย (wireless) หมายถึง การเชื่อมต่ออินเทอร์เน็ตแบบไร้สายของผู้ให้บริการเครือข่ายต่าง ๆ

การจัดกลุ่มผู้บริโภคโดยใช้พฤติกรรมการใช้โทรศัพท์มือถือเพื่อการส่งเสริมการขาย

ไลน์ (lines) หมายถึง โปรแกรมที่ไว้ใช้สำหรับการติดต่อสื่อสาร

ฝากข้อความเสียง (voice mail) หมายถึง ระบบที่ช่วยเก็บเสียงพูดของผู้โทรเข้าในกรณีที่ไม่สามารถติดต่อได้

เพจจิ้ง (paging service) หมายถึง บริการส่งข้อความติดตามตัวทางเดียว

อินเทอร์เน็ต (internet) หมายถึง การเชื่อมต่ออินเทอร์เน็ตผ่านเครือข่ายของผู้ให้บริการมือถือ

ระบุผู้โทร (caller ID) หมายถึง การแสดงหมายเลขโทรศัพท์หรือชื่อผู้โทร

บริการเรียกซ้อน (call waiting) หมายถึง การที่คนโทรเข้ามาหาเราในขณะที่เรากำลังสนทนาโทรศัพท์กับอีกคนหนึ่งอยู่ โดยจะมีเสียงเตือนเข้ามาทางหูฟัง

บริการโอนสายอัตโนมัติ (call forwarding) หมายถึง โอนสายไปที่เบอร์ปลายทางที่เรากำหนดในกรณีที่ไม่สามารถติดต่อได้

การประชุมสาย (3 - way calling) หมายถึง การติดต่อสื่อสารตั้งแต่ 3 คนขึ้นไป

ใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic - billing) หมายถึง บริการแสดงใบแจ้งยอดอิเล็กทรอนิกส์

3.2 เครื่องมือที่ใช้ในการวิจัย

การวิเคราะห์ข้อมูลในครั้งนี้ใช้เครื่องมือในการวิจัย คือ โปรแกรมการวิเคราะห์ทางสถิติ คือ IBM SPSS Modeler โดยใช้เทคนิคเคมีน คลัสเตอร์ เพื่อใช้ในการจัดกลุ่มพฤติกรรมข้อมูลผู้บริโภค

3.3 วิเคราะห์การจัดกลุ่ม (Clustering Analysis)

เป็นกระบวนการของการรวมชุดของข้อมูลออกเป็นกลุ่มๆซึ่งมีความคล้ายคลึงกันในแต่ละกลุ่ม เรียกว่าการจัดกลุ่ม (Clustering) โดยภายในกลุ่มอาจจะมีคุณลักษณะที่มีความคล้ายคลึงกันตั้งแต่หนึ่งคุณลักษณะจนถึงหลายคุณลักษณะ ในขณะที่เดียวกันก็มีคุณลักษณะที่แตกต่างกันระหว่างกลุ่มเช่นกัน [3]

3.4 ขั้นตอนการวิเคราะห์และจัดกลุ่มของวิธี K-Means

1. นำข้อมูลและตัวแปรที่ได้จากฐานข้อมูลของบริษัท IBM Knowledge Center (กรณีศึกษา) กำหนดสูตรเพื่อปรับค่าของข้อมูลตัวแปรบางตัวให้อยู่ในระดับ Ratio Scale และ Interval Scale จะได้ผลลัพธ์ที่ได้แล้วนำเข้าไปทำการวิเคราะห์ด้วยโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS Modeler โดยกำหนดค่าต่าง ๆ และปรับค่าตัวแปรต่าง ๆ เพื่อกำจัดความแตกต่างของข้อมูล [7]



2. ทำการวิเคราะห์ข้อมูลโดยใช้การวิเคราะห์การจัดกลุ่ม (Cluster Analysis) โดยการจัดกลุ่มด้วยวิธี K-Means เป็นการจำแนก Cases ออกเป็นกลุ่มย่อย โดยผู้วิจัยทำการวิเคราะห์ด้วยวิธี K-Means หลาย ๆ ครั้ง แต่ละครั้งกำหนดจำนวนกลุ่มแตกต่างกันไป เช่น 3, 4 หรือ 5 กลุ่ม แล้วพิจารณาหาจำนวนกลุ่มที่เหมาะสมจนได้จำนวนกลุ่ม หรือจำนวน Cluster ออกเป็น 3 กลุ่ม

3. จัดกลุ่มตัวแปรออกเป็นกลุ่มย่อย ๆ เป็นกลุ่มข้อมูลที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่เหมือนกันหรือคล้ายคลึงกัน โดยมีวิธีดังนี้

3.1 สุ่มค่าเริ่มต้นจำนวน k เรียกว่า Cluster Centroid

3.2 หาค่าระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางมากที่สุด หากข้อมูลตัวใดใกล้กับศูนย์กลางกลุ่มใดมากที่สุดจะอยู่กลุ่มนั้น

3.3 หาค่าเฉลี่ย จากนั้นนำมาจัดตามกลุ่มแต่ละกลุ่มให้เป็นค่าจุดศูนย์กลางใหม่ [8]

3.4 กลับไปยังข้อ 3.2 และเริ่มทำใหม่ จนกระทั่งค่าเฉลี่ยในแต่ละกลุ่มไม่เปลี่ยนแปลงโดยจะมีการทำงานหลาย ๆ รอบ (Iteration) โดยในแต่ละรอบจะมีการรวบรวม Cases ให้ไปอยู่ในกลุ่มใดกลุ่มหนึ่ง โดยเลือกกลุ่มที่ Case นั้นมีระยะห่างจากค่ากลางของกลุ่มน้อยที่สุด แล้วคำนวณค่ากลางของกลุ่มใหม่จะทำงานนี้จนกระทั่งค่ากลางของกลุ่มไม่เปลี่ยนแปลงหรือครบจำนวนรอบที่กำหนดไว้

4. ผลการทดลอง

ตารางที่ 1 Initial Cluster Centers

Initial Cluster Centers			
	Cluster		
	1	2	3
โทรทางไกล (long distance)			
โทรศัพท์ฟรี (toll free)	2.48	-1.70	.12
อุปกรณ์เสริม (equipment)	2.34	-.20	-.39
บัตรโทรศัพท์ (calling card)	1.34	-.65	.59
อินเทอร์เน็ตไร้สาย (wireless)	2.49	-.86	-1.28
ไลน์ (lines)	1.14	-1.75	1.42
ฝากข้อความเสียง (voice mail)	1.05	-.95	1.05
เพจจิ้ง (paging service)	1.51	1.51	1.51
อินเทอร์เน็ต (internet)	1.68	1.68	1.68
ระบุผู้โทร (caller ID)	1.31	-.76	1.31
บริการเรียกซ้อน (call waiting)	1.04	1.04	-.96
บริการโอนสายอัตโนมัติ (call forwarding)	1.03	-.97	1.03
การประชุมสาย (3 - way calling)	1.01	1.01	-.99
ใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing)	1.00	1.00	-1.00
	.77	-.77	1.30

จากตารางที่ 1 เป็นค่าตัวแปรของค่าสังเกตมีระยะห่างในแต่ละกลุ่ม แสดงค่าเฉลี่ยของตัวแปรแต่ละตัวที่ Standardized ใน Cluster ต่าง ๆ หรือเป็นค่ากลางของ Cluster ในตอนเริ่มต้นได้กำหนดไว้ 3 กลุ่มหรือ 3 Cluster ผลการจัดกลุ่มและลักษณะปัจจัยประจำกลุ่ม ตามลำดับจากมากไปน้อย สรุปได้ดังนี้

กลุ่มที่ 1 ปัจจัยที่สำคัญ คือ การโทรทางไกล (long distance) โทรศัพท์ฟรี (toll free) อุปกรณ์เสริม (equipment) บัตรโทรศัพท์ (calling card) ไลน์ (lines) ฝากข้อความเสียง (voice mail) เพจจิ้ง (paging service) อินเทอร์เน็ต (internet) ระบุผู้โทร (caller ID) บริการเรียกซ้อน (call waiting) บริการโอนสายอัตโนมัติ (call forwarding) การประชุมสาย (3 - way calling)

กลุ่มที่ 2 ปัจจัยที่สำคัญ คือ ฝากข้อความเสียง (voice mail) เพจจิ้ง (paging service) ระบุผู้โทร (caller ID) บริการโอนสายอัตโนมัติ (call forwarding) การประชุมสาย (3 - way calling)

กลุ่มที่ 3 มีปัจจัยที่สำคัญ คือ อินเทอร์เน็ตไร้สาย (wireless) ไลน์ (lines) ฝากข้อความเสียง (voice mail) เพจจิ้ง (paging - service) อินเทอร์เน็ต (internet) บริการเรียกซ้อน (call waiting) และใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing)

ตารางที่ 2 Iteration History

Iteration	Change in Cluster Centers		
	1	2	3
1	3.298	3.590	3.491
2	1.016	.427	.931
3	.577	.320	.420
4	.240	.180	.195
5	.119	.125	.108
6	.093	.083	.027
7	.069	.094	.032
8	.059	.051	.018
9	.035	.085	.063
10	.025	.359	.333
11	.068	.439	.287
12	.079	.368	.177
13	.125	.139	.078
14	.077	.096	.020
15	.041	.047	.015
16	.014	.027	.000
17	.019	.038	.000
18	.000	.000	.000

จากตารางที่ 2 เป็นการแสดงค่าเฉลี่ยหรือค่ากลางของแต่ละ Cluster ที่เปลี่ยนไปในขั้นตอนของการหมุนในการวิเคราะห์ข้อมูลนี้ ได้กำหนดจำนวนรอบสูงสุดไว้ที่ 20 รอบ จากตารางนี้ ในการหมุนช่วงตอนต้นจะมีความต่างกันมาก (รอบที่ 1-12) ต่อมาช่วงรอบที่ 14 เริ่มได้ตำแหน่งสุดท้ายเพราะค่าหลังจากนั้นมีการปรับตัวน้อยมาก ในการหมุนครั้งนี้จึงหยุดอยู่รอบที่ 18 ถือว่าดี เพราะว่าการหมุนหรือ Iteration ต้องน้อยกว่าที่เรากำหนดไว้ แต่ถ้าหยุดอยู่รอบที่ 20 ที่เรากำหนด เราต้องไปตั้งค่า Iteration ที่เพิ่มขึ้น

ตารางที่ 3 Final Cluster Centers

Final Cluster Centers	Cluster		
	1	2	3
โทรทางไกล (long distance)	.05	.22	-.16
โทรศัพท์ฟรี (toll free)	.24	.12	-1.05
อุปกรณ์เสริม (equipment)	.81	-.19	-.69
บัตรโทรศัพท์ (calling card)	.17	.02	-.17
อินเทอร์เน็ตไร้สาย (wireless)	.42	-.75	-1.00
ไลน์ (lines)	.48	-.29	-.05
ฝากข้อความเสียง (voice mail)	1.26	-.24	-.44
เพจจิ้ง (paging service)	1.43	-.38	-.44
อินเทอร์เน็ต (internet)	.81	-.59	-.02
ระบุผู้โทร (caller ID)	.82	.71	-.81
บริการเรียกซ้อน (call waiting)	.76	.72	-.80
บริการโอนสายอัตโนมัติ (call forwarding)	.78	.69	-.79
การประชุมสาย (3 - way calling)	.74	.67	-.75
ใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing)	.70	.63	.05

จากตารางที่ 3 เป็นการแสดงค่าเฉลี่ยสำหรับตัวแปรกลุ่มสุดท้ายที่สะท้อนคุณลักษณะของตัวอย่างของแต่ละกลุ่ม ค่าเฉลี่ยในตารางคือ ค่ากลางของแต่ละ Cluster จะพบว่า ค่าเฉลี่ยของการส่งข้อความ (paging) ใน Cluster ที่ 1 จะแตกต่างกับ Cluster ที่ 2 และ 3 มากเมื่อเทียบกับลักษณะปัจจัยอื่น ๆ โดยลูกค้ากลุ่มที่ 1 มีแนวโน้มเป็นกลุ่มที่ใช้บริการมากที่สุด และจ่ายเงินมากที่สุด ลูกค้ากลุ่มที่ 2 ใช้บริการระดับปานกลางแต่เป็นกลุ่มที่ชอบใช้บริการทางโทรศัพท์ เช่น ฝากข้อความโทรกลับ บริการเรียกซ้อน ลูกค้ากลุ่มที่ 3 ไม่ค่อยได้ใช้หรือซื้อบริการที่หลากหลาย

ตารางที่ 4 Distances between Final Cluster Centers

Distances between Final Cluster Centers			
Cluster	1	2	3
1		3.500	4.863
2	3.500		3.396
3	4.863	3.396	

จากตารางที่ 4 จะเห็นได้ว่า Cluster กลุ่มที่ 1 กับ Cluster กลุ่มที่ 3 มีความแตกต่างกันมาก แต่ Cluster กลุ่มที่ 2 จะมีความคลึงกับ Cluster กลุ่มที่ 1 และกลุ่มที่ 3 เป็นการแสดงระยะห่างระหว่างค่ากลางของทั้ง 3 cluster พบว่า Cluster ที่ 1 มีระยะห่างจาก Cluster ที่ 2 มากที่สุด คือ 3.500 และใกล้ Cluster ที่ 3 มากที่สุดคือ 4.863 และ Cluster ที่ 3 ใกล้กับ Cluster ที่ 2 มากที่สุดเช่นกัน คือ 3.396

ตารางที่ 5 Number of Cases in each Cluster

Number of Cases in each Cluster		
Cluster	1	226.000
	2	292.000
	3	482.000
Valid		1000.000
Missing		.000

จากตารางที่ 5 เป็นการแสดงจำนวนข้อมูลผู้ใช้โทรคมนาคมในแต่ละ Cluster และพบว่าลูกค้ากลุ่มที่ 3 ส่วนใหญ่ มีจำนวนที่เยอะ จึงถูกระบุให้อยู่ในกลุ่มที่ไม่สามารถทำกำไรให้องค์กรธุรกิจได้มากเท่าที่ควร



	ANOVA					
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
โทรทางไกล (long distance)	13.063	2	.975	997	13.387	.000
โทรศัพท์ฟรี (toll free)	43.418	2	.820	472	52.932	.000
อุปกรณ์เสริม (equipment)	99.056	2	.488	383	202.999	.000
บัตรโทรศัพท์ (calling card)	6.301	2	.984	675	6.402	.002
อินเทอร์เน็ตไร้สาย (wireless)	52.879	2	.646	293	81.873	.000
ไลน์ (lines)	38.032	2	.926	997	41.084	.000
ฝากข้อความเสียง (voice mail)	236.301	2	.528	997	447.554	.000
เพจจิ้ง (paging service)	298.992	2	.402	997	743.348	.000
อินเทอร์เน็ต (internet)	123.447	2	.754	997	153.542	.000
ระบุผู้โทร (caller ID)	308.104	2	.384	997	802.474	.000
บริการเรียกซ้อน (call waiting)	294.674	2	.411	997	717.172	.000
บริการโอนสายอัตโนมัติ (call forwarding)	288.343	2	.424	997	680.718	.000
การประชุมสาย (3 - way calling)	262.397	2	.476	997	551.678	.000
ใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing)	112.782	2	.776	997	145.381	.000

จากตารางที่ 6 เป็นการวิเคราะห์ความแปรปรวน ใช้เป็นตัววัดค่าตัวแปรใดที่มีผลมากที่สุดต่อการกำหนดกลุ่มที่ได้ระบุเอาไว้ ตัวแปรใดที่มีค่า F สูงสุด แสดงว่าเป็นตัวแปรที่แบ่งแยกกลุ่มได้ดีที่สุด แสดงค่า Mean Square ระหว่าง Mean Square Cluster และ Mean Square Error และให้ค่าสถิติ F โดยที่จะไม่ใช่ค่าสถิติ F และค่า Significance ใน Column สุดท้ายของตาราง ในการทดสอบค่าความแตกต่างระหว่างค่าเฉลี่ยของแต่ละตัวแปรเมื่ออยู่ต่าง Cluster กัน จะพบว่าค่าเฉลี่ยของปัจจัยการระบุการโทร (Caller id) เมื่อต่างกลุ่มกันจะมีความต่างกันมากที่สุด เนื่องจากค่าสถิติ F สูงสุด คือ 802.474 และของปัจจัยการส่งข้อความ (paging) รองลงมาคือ F = 743.348 ในตารางนี้ใช้ค่า Mean Square มาเปรียบเทียบกัน ส่วนปัจจัยการใช้บัตรโทรศัพท์ (calling card) มีค่าเฉลี่ยแตกต่างกันน้อยที่สุดเมื่ออยู่ต่าง Cluster กัน (F = 6.402)

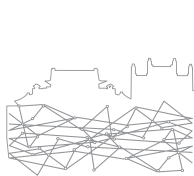
5. อภิปรายผลการวิจัย

ผู้บริโภคกลุ่มที่ 1 มีแนวโน้มเป็นกลุ่มที่ใช้บริการเครือข่ายที่มากที่สุด และใช้จ่ายมากที่สุด มีค่าตัวแปรดังนี้ การโทรทางไกล (long distance) มีค่าเท่ากับ .05 บริการโทรศัพท์ฟรี (toll free) มีค่าเท่ากับ .24 อุปกรณ์เสริม (equipment) มีค่าเท่ากับ .81 บัตรโทรศัพท์ (calling card) มีค่าเท่ากับ .17 อินเทอร์เน็ตไร้สาย (wireless) มีค่าเท่ากับ .42 ไลน์ (lines)

มีค่าเท่ากับ .48 ฝากข้อความเสียง (voice mail) มีค่าเท่ากับ 1.26 เพจจิ้ง (paging service) มีค่าเท่ากับ 1.43 อินเทอร์เน็ต (internet) มีค่าเท่ากับ .81 ระบุผู้โทร (caller ID) มีค่าเท่ากับ .82 บริการเรียกซ้อน (call waiting) มีค่าเท่ากับ .86 บริการ โอนสายอัตโนมัติ (call forwarding) มีค่าเท่ากับ .78 การประชุมสาย (3 - way calling) มีค่าเท่ากับ .74 และใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing) มีค่าเท่ากับ .70

ผู้บริโภคกลุ่มที่ 2 ใช้บริการระดับปานกลาง แต่เป็นกลุ่มที่ชอบใช้บริการทางโทรศัพท์ มีค่าตัวแปรดังนี้ การโทรทางไกล (long distance) มีค่าเท่ากับ .22 บริการโทรศัพท์ฟรี (toll free) มีค่าเท่ากับ .12 อุปกรณ์เสริม (equipment) มีค่าเท่ากับ .19 บัตรโทรศัพท์ (calling card) มีค่าเท่ากับ .02 อินเทอร์เน็ตไร้สาย (wireless) มีค่าเท่ากับ .75 ไลน์ (lines) มีค่าเท่ากับ .29 ฝากข้อความเสียง (voice mail) มีค่าเท่ากับ .24 เพจจิ้ง (paging service) มีค่าเท่ากับ .38 อินเทอร์เน็ต (internet) มีค่าเท่ากับ .59 ระบุผู้โทร (caller ID) มีค่าเท่ากับ .71 บริการเรียกซ้อน (call waiting) มีค่าเท่ากับ .72 บริการโอนสายอัตโนมัติ (call forwarding) มีค่าเท่ากับ .69 การประชุมสาย (3 - way calling) มีค่าเท่ากับ .67 และใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing) มีค่าเท่ากับ .63

การจัดกลุ่มผู้บริโภคโดยใช้พฤติกรรมการใช้โทรศัพท์มือถือเพื่อการส่งเสริมการขาย



ผู้บริโภคกลุ่มที่ 3 เป็นกลุ่มลูกค้าที่ไม่ค่อยได้ใช้บริการหรือซื้อบริการที่มีความหลากหลาย จึงมีแต่ค่าที่ติดลบไม่สามารถที่จะทำอะไรหรือผลประโยชน์ให้กับองค์กรได้มีค่าตัวแปรดังนี้ ดังนี้ การโทรทางไกล (long distance) มีค่าเท่ากับ -.16 บริการโทรศัพท์ฟรี (toll free) มีค่าเท่ากับ -.05 อุปกรณ์เสริม (equipment) มีค่าเท่ากับ -.69 บัตรโทรศัพท์ (calling card) มีค่าเท่ากับ -.17 อินเทอร์เน็ตไร้สาย (wireless) มีค่าเท่ากับ -.00 ไลน์ (lines) มีค่าเท่ากับ -.05 ฝากข้อความเสียง (voice mail) มีค่าเท่ากับ -.44 เพจจิง (paging service) มีค่าเท่ากับ -.44 อินเทอร์เน็ต (internet) มีค่าเท่ากับ -.02 ระบุผู้โทร (caller ID) มีค่าเท่ากับ -.81 บริการเรียกซ้อน (call waiting) มีค่าเท่ากับ -.80 บริการโอนสายอัตโนมัติ (call forwarding) มีค่าเท่ากับ -.79 การประชุมสาย (3 - way calling) มีค่าเท่ากับ -.75 และใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic billing) มีค่าเท่ากับ -.05

จากผลการศึกษาในตารางที่ 3 สรุปได้ว่า การที่จัดกลุ่มผู้บริโภคได้เป็น 3 กลุ่ม เกิดจากผู้วิจัยได้ทำการวิเคราะห์ด้วยวิธี K-Means หลายๆ ครั้ง แล้วได้ผลการจัดกลุ่มเป็น 3 กลุ่มที่มีค่าที่ดีที่สุด โดยแต่ละกลุ่มมีลักษณะปัจจัยสำคัญประจำกลุ่ม ดังนี้ คือ 1) ผู้บริโภคกลุ่มที่ 1 เป็นกลุ่มผู้บริโภคที่ใช้บริการเครือข่ายที่มากที่สุด ในจำนวนกลุ่มผู้บริโภคทั้งหมด และมีค่าใช้จ่ายบริการที่มากที่สุด ลักษณะปัจจัยสำคัญของตัวแปรกลุ่มนี้ ได้แก่ การโทรทางไกล (long distance) บริการโทรศัพท์ฟรี (toll free) อุปกรณ์เสริม (equipment) บัตรโทรศัพท์ (calling card) อินเทอร์เน็ตไร้สาย (wireless) ไลน์ (lines) ฝากข้อความเสียง (voice mail) เพจจิง (paging service) อินเทอร์เน็ต (internet) ระบุผู้โทร (caller ID) บริการเรียกซ้อน (call waiting) การบริการโอนสายอัตโนมัติ (call - forwarding) การประชุมสาย (3 - way calling) และใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic - billing) 2) ผู้บริโภคกลุ่มที่ 2 เป็นกลุ่มใช้บริการระดับปานกลาง แต่เป็นกลุ่มที่ชอบใช้บริการทางโทรศัพท์ และลักษณะปัจจัยสำคัญของกลุ่มนี้ ได้แก่ การโทรทางไกล (long distance) บริการโทรศัพท์ฟรี (toll free) อุปกรณ์เสริม (equipment) บัตรโทรศัพท์ (calling card) อินเทอร์เน็ตไร้สาย (wireless) ไลน์ (lines) ฝากข้อความเสียง (voice mail) เพจจิง (paging service) อินเทอร์เน็ต (internet) ระบุผู้โทร (caller ID)

บริการเรียกซ้อน (call waiting) บริการโอนสายอัตโนมัติ (call forwarding) การประชุมสาย (3 - way calling) และใบแจ้งค่าบริการอิเล็กทรอนิกส์ (electronic - billing) 3) ผู้บริโภคกลุ่มที่ 3 เป็นกลุ่มที่ไม่ค่อยได้ใช้บริการหรือเลือกซื้อบริการที่มีความหลากหลาย ผลการศึกษานี้สามารถนำไปประยุกต์ใช้เป็นปัจจัยสำคัญประจำกลุ่มของกลุ่มผู้บริโภคมาเพื่อเป็นข้อมูลในการสนับสนุนในการกำหนดแนวทางในการทำกลยุทธ์ทางการตลาดหรือการทำรายการส่งเสริมการขายของผู้ให้บริการเครือข่ายต่าง ๆ ได้

6. เอกสารอ้างอิง

- [1] Tsai, C. F., Tsai, C. T., Hung, C. S., & Hwang, P. S. (2011). Data Mining Techniques for Identifying Students at Risk of Failing a Computer Proficiency Test Required for Graduation. *Australian Journal of Educational Technology*, 27(3), 481-498.
- [2] Han, J., & Kamber, M. (2006). *Data Mining Concepts and Techniques Second Edition*. USA: Morgan Kaufmann Publishers.
- [3] Wagstaff, K., Cardie, C., Rogers, S., & Schrödl, S. (2001). Constrained K-means Clustering with Background Knowledge (pp. 577-584). In *Proceedings of the Eighteenth International Conference on Machine Learning*. San Francisco, USA.
- [4] Chirawichitchai, N. (2009). An Experimental Study on Feature Reduction Techniques and Classification Algorithms of Thai Documents. *Journal of Science Ladkrabang*, 18(2), 11-24.
- [5] กุลยา วานิชย์บัญชา. (2552). *การวิเคราะห์ข้อมูลหลายตัวแปร* (พิมพ์ครั้งที่ 4). กรุงเทพมหานคร : ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.



- [6] บุหงา ชัยสุวรรณ. (2561). การจัดกลุ่มและหาคุณลักษณะของกลุ่มดิจิทัลอิมมิแกรนตตามพฤติกรรมการสื่อสารในบริบทออนไลน์และปัญหาทางความสัมพันธ์กับสังคม. *วารสารพัฒนาบริหารศาสตร์*, 58(1), 54-73.
- [7] ปราณี มณีรัตน์. (2554). *การสร้างโมเดลการจัดกระบวนนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูล*. รายงานการวิจัย คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม.
- [8] อรุณ ชัยหมื่น (2548). *การศึกษาเปรียบเทียบแบ่งกลุ่มข้อมูลลูกค้าสินค้าหัตถกรรมไทยโดยวิธีขั้นตอนของ SOM กับ K-Means Algorithm และ Hierarchical Clustering and K-Means Algorithm*. วิทยานิพนธ์วิทยาศาสตร์ มหาบัณฑิต มหาวิทยาลัยเกษตรศาสตร์.