

การวิเคราะห์แบบ RFM ในการแบ่งกลุ่มผู้ใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารใน องค์กร กรณีศึกษาสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์

RFM Analysis to Segment Users of Printers and Copiers in Organizations:

A Case Study of School of Informatics, Walailak University

กาญจนา หฤพรพงษ์^{1*} และ ปิยมาศ จิตตระ²

สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ จ.นครศรีธรรมราช 80160¹

ศูนย์เทคโนโลยีดิจิทัล มหาวิทยาลัยวลัยลักษณ์ จ.นครศรีธรรมราช 80160²

hkanjana@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อทำการค้นหาคำตอบจากการวิเคราะห์ข้อมูลเพื่อจัดกลุ่มผู้ใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารในองค์กร โดยการทำเหมืองข้อมูลด้วยวิธีการจัดกลุ่ม ข้อมูลที่ใช้ในการวิเคราะห์ประกอบด้วย 3 ส่วน คือ การพิมพ์งานด้วยเครื่องพิมพ์ การพิมพ์งานด้วยเครื่องถ่ายเอกสาร และการถ่ายเอกสาร งานวิจัยนี้นำเสนอวิธีในการแบ่งกลุ่มผู้ใช้ตามพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารด้วยแบบจำลอง Recency, Frequency and Monetary หรือ RFM ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ จำนวน 84,764 รายการ ตั้งแต่ปี 2554-2559 ผู้ใช้งานทั้งหมด 78 คน วิเคราะห์ด้วย RapidMiner Studio 7.4

ผลจากการวิจัยพบว่าสามารถแบ่งกลุ่มผู้ใช้ตามพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารออกเป็น 14 กลุ่มได้อย่างชัดเจน โดยกลุ่มที่น่าสนใจ คือ กลุ่มที่มีความถี่ในการใช้บ่อยและมีปริมาณใช้งานมาก จำนวน 15 คน จากผลการวิจัยทำให้ทราบรูปแบบการใช้งานและนำไปใช้ในการตรวจสอบพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรว่ามีการใช้งานที่มีความสอดคล้องกับงานที่ทำอยู่หรือไม่เพื่อช่วยลดค่าใช้จ่าย ผลการวิจัยดังกล่าวเมื่อนำมาตรวจสอบไม่พบความผิดปกติในการใช้งานเนื่องจากกลุ่มผู้ใช้ในกลุ่มดังกล่าวเป็นเจ้าหน้าที่ที่รับผิดชอบทำงานงานด้านเอกสาร นอกจากนี้ยังพบว่าพฤติกรรมของบุคลากรในการพิมพ์เอกสารส่วนใหญ่นิยมพิมพ์ทางเครื่องพิมพ์มากกว่าการพิมพ์ทางเครื่องถ่ายเอกสาร

คำสำคัญ: การวิเคราะห์แบบ RFM, การแบ่งกลุ่มผู้ใช้, การค้นหาคำตอบ, การทำเหมืองข้อมูล

ABSTRACT

This research aims to discover knowledge by cluster analysis data of printer and copier usage in the organization using data mining techniques in clustering. There are three types of data to analyze: printing information with the printer, printing with a copier, and photocopying usage. This research applies the Recency, Frequency, and Monetary (RFM) model to segment the users according to the usage behavior regarding the printer and copier. This paper presents a method for analyzing such data using printer and copier data comprising 84,764 transactions of 78 users in the School of Informatics, Walailak University, within the period of use from 2011 to 2016 using RapidMiner Studio version 7.4.

The results of the RFM segmentation can be clearly divided into 14 groups. The interesting findings were that there are 15 users who are frequent users and have a high usage rate. This knowledge can help organizations monitor staff behaviors regarding using printers and copiers that are consistent with their jobs. We found that no abnormal results were found because most users in this group are staff that are responsible for documentation work. In addition, most of the printing behavior was using the printer rather than the copier.

Keywords: RFM Analysis, User Segmentation, Knowledge Discovery, Data Mining

บทนำ

ปัญหาค่าใช้จ่ายในองค์กรที่สำคัญประการหนึ่งเกิดจากค่าใช้จ่ายในการจัดทำเอกสารที่มีปริมาณมาก โดยเฉพาะในสถาบันการศึกษาซึ่งยังใช้การจัดทำเอกสารในรูปแบบกระดาษในการทำงานเป็นส่วนใหญ่ พฤติกรรมในการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรในหน่วยงานเป็นข้อมูลที่ถูกหน่วยงานมีการเก็บข้อมูลไว้อยู่แล้ว ดังนั้นการนำข้อมูลการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรมาวิเคราะห์เพื่อดูพฤติกรรมการใช้งานสามารถช่วยให้หน่วยงานค้นพบข้อมูลที่มีประโยชน์ เช่น การจัดกลุ่มพฤติกรรมการใช้งานออกเป็นกลุ่มเพื่อช่วยให้การนำไปตรวจสอบหรือวางแผนในการให้บริการเครื่องพิมพ์หรือเครื่องถ่ายเอกสารให้มีความเหมาะสมต่อพฤติกรรมการใช้งานของบุคลากร หรือการตรวจสอบการใช้งานในกลุ่มที่มีการใช้งานบ่อยและมีปริมาณมากเพื่อดูว่ามีความสอดคล้องกับหน้าที่การทำงานที่ทำอยู่หรือไม่เพื่อช่วยลดค่าใช้จ่ายในองค์กรได้ ทางผู้วิจัยจึงมองเห็นว่าข้อมูลการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ ที่จัดเก็บมาตั้งแต่ปี พ.ศ. 2554 – 2559 ซึ่งเก็บข้อมูลพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารเป็นรายคน จำนวน 78 คน มีจำนวน 84,764 รายการ และจำนวนการพิมพ์เอกสารมากถึง 1,431,218 แผ่น [1] ข้อมูลดังกล่าวนี้สามารถนำมาวิเคราะห์เพื่อให้เกิดประโยชน์ต่อหน่วยงาน จึงเสนอวิธีการในการวิเคราะห์ข้อมูลเพื่อค้นหาความรู้โดยประยุกต์ใช้การทำเหมืองข้อมูลด้วยแบบจำลอง RFM (Recency, Frequency and Monetary) ในการแบ่งกลุ่มผู้ใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสาร ซึ่งวิธีการดังกล่าวเป็นวิธีที่นิยมใช้ในการจัดกลุ่มลูกค้าตามพฤติกรรมการบริโภคเพื่อขยายผลต่อทางด้านการตลาด แต่ในงานวิจัยนี้เป็นการประยุกต์เพื่อนำมาใช้ศึกษาข้อมูลพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ โดยนำเสนอวิธีการเตรียมข้อมูลและขั้นตอนการสร้างแบบจำลองการแบ่งกลุ่มผู้ใช้งานเพื่อค้นหาข้อมูลที่สามารถนำไปใช้ประโยชน์ในการตรวจสอบความเหมาะสมในการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารต่อไป

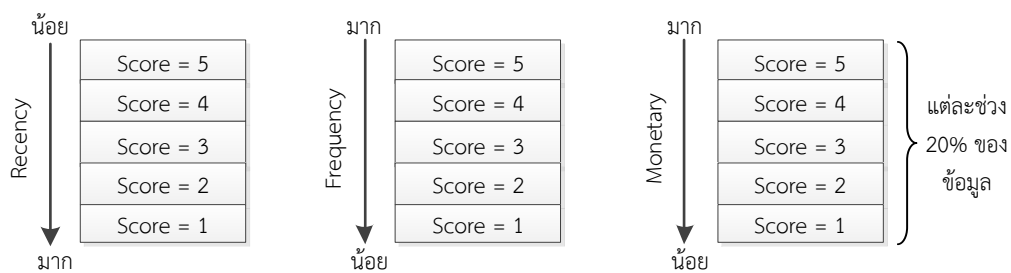
1. วัตถุประสงค์การวิจัย

- 1.1 เพื่อวิเคราะห์พฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรในองค์กร
- 1.2 เพื่อทำการจัดกลุ่มพฤติกรรมการใช้งานออกเป็นกลุ่มและนำไปใช้ในการตรวจสอบการทำงานต่อไป

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

การค้นหาคำถามความรู้จากฐานข้อมูลขนาดใหญ่ (Knowledge discovery in database: KDD) เป็นการค้นหาข้อมูลที่มีประโยชน์ที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ โดยนำผลที่ได้ไปใช้ประโยชน์ต่อหน่วยงาน เช่น การศึกษาพฤติกรรมของลูกค้าเพื่อใช้สำหรับการรักษาความสัมพันธ์ที่ดีกับลูกค้า เป็นวิธีการในการรักษาลูกค้าหรือช่วยให้สามารถขายสินค้าได้ การทำเหมืองข้อมูลจึงเป็นขั้นตอนหนึ่งของการค้นหาคำถามความรู้จากระบบฐานข้อมูลขนาดใหญ่ เทคนิคที่ใช้ในการทำเหมืองข้อมูลที่สำคัญมีอยู่มากมาย เช่น การจัดกลุ่มข้อมูล การค้นหารูปแบบความสัมพันธ์ของข้อมูล การจำแนกประเภทข้อมูล [2] เป็นต้น โดยการทำงานของ KDD จะนำเอาผลลัพธ์ที่ได้จากการทำเหมืองข้อมูลมาประเมินเพื่อเลือกเอาความรู้ที่มีประโยชน์นำไปใช้งานต่อไป [3], [4], [5] การวิเคราะห์แบบ RFM (Recency, Frequency and Monetary) เป็นเทคนิคหนึ่งที่นิยมใช้มาเป็นเวลานาน โดยประยุกต์ใช้ด้านการตลาดแบบทางตรง เช่น การแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้าของลูกค้า [6], [7] โดยพิจารณาจากข้อมูล 3 ค่า ดังนี้ Recency หมายถึง ความเป็นปัจจุบัน โดยพิจารณาจากระยะเวลา (จำนวนวัน) ที่ซื้อสินค้าล่าสุด Frequency หมายถึง ความถี่ โดยพิจารณาจากความถี่ของการซื้อสินค้าว่ามีมากน้อยเพียงใด และ Monetary หมายถึง มูลค่าในการซื้อ หรือปริมาณการใช้จ่ายของลูกค้า โดยจะทำการแบ่งค่า Recency (R), Frequency (F) และ Monetary (M) ออกเป็นส่วนๆ เท่าๆ กัน เช่น ถ้าแบ่งออกเป็น 5 ส่วน โดยเลขที่มีค่ามากที่สุด คือ เลข 5 จะมีความสำคัญมากที่สุด และใช้ค่าตัวเลข 3 หลักมาแทนค่าข้อมูลในแต่ละกลุ่ม เช่น กลุ่ม 555 หมายถึง กลุ่มที่มีค่า R

= 5, F = 5 และ M = 5 ซึ่งหมายถึงกลุ่มของลูกค้าที่มีพฤติกรรมการซื้อสินค้าไม่นานนี้ มีการซื้อบ่อย และมีการใช้จ่ายที่สูง โดยการแบ่งกลุ่มลูกค้าด้วยวิธี RFM มีการจัดเรียงลำดับข้อมูล เช่น Recency ทำการเรียงจากน้อยไปหามาก ค่าน้อย หมายถึง ลูกค้ามาซื้อสินค้าเมื่อไม่นานมานี้ ทำให้ทราบว่าเป็นลูกค้าที่มาซื้อสินค้าล่าสุดในช่วงเวลานานแล้วหรือไม่ Frequency ทำการเรียงค่าจากมากไปหาน้อย ลูกค้าที่มีค่า Frequency สูง แสดงว่ามีความถี่ในการซื้อสินค้าบ่อยหรือซื้อสินค้าเป็นประจำ และเรียงค่า Monetary จากมากไปหาน้อย ลูกค้าที่มีค่า Monetary สูง แสดงว่ามียอดการใช้จ่ายในการซื้อสินค้าในปริมาณมาก โดยในการคำนวณค่าคะแนน RFM ในแต่ละกลุ่มของตัวอย่างนี้จะแบ่งออกเป็น 5 ระดับ ดังภาพที่ 1



ภาพที่ 1 การแบ่งกลุ่มระดับคะแนนของ RFM

FADER และคณะ [8] ได้ทำการศึกษาและนำเสนอแบบจำลองแบบใหม่ที่เชื่อมโยงเข้าด้วยกันกับ RFM เพื่อใช้ในการประเมินคุณค่าของลูกค้าตลอดช่วงชีวิต (Customer Lifetime Value: CLV) ซึ่งเป็นงานวิจัยเพื่อทำการจัดกลุ่มลูกค้าโดยใช้ตัวแบบในการวิเคราะห์แบบ RFM และใช้ข้อมูลลูกค้าจากร้านขายปลีกในประเทศได้หันมาทำการวิเคราะห์ช่วยให้การจัดกลุ่มของลูกค้าแต่ละรายที่มีประวัติการซื้อที่มีความแตกต่างกัน นอกจากนี้ยังมีการวิจัยที่นำเสนอการประยุกต์ใช้แบบจำลอง RFM ในธุรกิจประกันภัย เรียกว่าแบบจำลอง LRFMC [9] ซึ่งแตกต่างจาก RFM แบบเดิม คือ ผู้วิจัยได้เสนอวิธีการเพิ่มพารามิเตอร์ที่สำคัญอีก 2 ตัว คือ ระยะเวลาที่เป็นลูกค้า (Length of time) และค่าใช้จ่ายของลูกค้า (Customer costs) โดยใช้กรณีศึกษาของธุรกิจประกันภัยจากข้อมูลลูกค้าจำนวน 120 คน มาทำการวิเคราะห์ด้วยแบบจำลองดังกล่าวเพื่อทำการแบ่งกลุ่มลูกค้า ผลของการเปรียบเทียบประสิทธิภาพของแบบจำลอง RFM และแบบจำลอง LRFMC จากการใช่วิธีการหาความสัมพันธ์ (Association rules) พบว่า LRFMC มีประสิทธิภาพดีกว่าเล็กน้อย ส่วนอีกงานวิจัยเป็นการประยุกต์ใช้การวิเคราะห์แบบ RFM เพื่อทำการวิเคราะห์หาโปรไฟล์ของลูกค้าโรงแรมที่ทำกำไร ด้วยการวิเคราะห์แบบ RFM ในการวิเคราะห์ความถี่และตัวชี้วัดทางการเงินสำหรับการค้นหาจากคุณลักษณะของลูกค้า โดยใช้ข้อมูลลูกค้าจากระบบบริหารลูกค้าสัมพันธ์ (Customer Relationship Management: CRM) ที่เกิดขึ้นจริงของโรงแรมระดับ 5 ดาว จำนวน 3 โรงแรมในประเทศไทยก็มาวิเคราะห์ ผลการวิเคราะห์ลูกค้าจำนวน 369 คน พบว่าสามารถแบ่งกลุ่มลูกค้าโรงแรมได้อย่างมีประสิทธิภาพ โดยแบ่งออกเป็น 8 กลุ่ม และนำผลการแบ่งกลุ่มลูกค้าที่ได้ไปใช้ในการสร้างกลยุทธ์ใหม่ๆ เพื่อเพิ่มขีดความสามารถของธุรกิจในการดูแลและสร้างความสัมพันธ์กับลูกค้าของโรงแรมได้ [10]

วิธีดำเนินการวิจัย

1. เครื่องมือการวิจัย

- 1.1 โปรแกรมวิเคราะห์ข้อมูล RapidMiner Studio เวอร์ชัน 7.4
- 1.2 โปรแกรม Microsoft Excel เวอร์ชัน 2007 ในการเตรียมข้อมูล
- 1.3 ขั้นตอนการวิเคราะห์แบบ RFM

2. ขั้นตอนการดำเนินการวิจัย

2.1 การเตรียมข้อมูล เริ่มจากการคัดเลือกข้อมูลการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ จำนวน 84,764 รายการ ตั้งแต่เดือนสิงหาคม พ.ศ. 2554 ถึงเดือนมกราคม พ.ศ. 2559 จากผู้ใช้งานจำนวนทั้งหมด 78 คน ข้อมูลประกอบด้วย 1) การพิมพ์งานด้วยเครื่องพิมพ์ 2) การพิมพ์งานด้วยเครื่องถ่ายเอกสาร 3) การถ่ายเอกสาร ซึ่งได้นำไฟล์ดังกล่าวมาจากรฐานข้อมูลและแปลงข้อมูลให้อยู่ในรูปแบบของ Excel files จำนวน 3 ไฟล์ โดยใช้แอตทริบิวต์ของการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสาร ดังตารางที่ 1 จากนั้นจึงทำการกลั่นกรองและรวบรวมข้อมูลการใช้งานที่มีอยู่ 3 ไฟล์ให้อยู่ในไฟล์เดียวกัน และทำการสร้างโปรเซสเพื่อวิเคราะห์ตามอัลกอริทึมของการทำเหมืองข้อมูลตามขั้นตอนของแบบจำลอง RFM โดยใช้โปรแกรม RapidMiner Studio เวอร์ชัน 7.4 ซึ่งเป็นซอฟต์แวร์สำหรับการทำเหมืองข้อมูลที่เป็นซอฟต์แวร์รหัสเปิด (Open source software) สามารถดาวน์โหลดมาใช้งานได้โดยไม่มีค่าใช้จ่าย [11], [12]

ตารางที่ 1 แอตทริบิวต์การใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารที่ใช้ในการวิเคราะห์

ชื่อแอตทริบิวต์	คำอธิบาย
ลำดับที่	หมายเลขรายการ
วันที่ใช้งาน	วันที่และเวลาในการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสาร
ชื่อผู้ใช้งาน	ชื่อผู้ใช้งาน
ขนาดเอกสาร	ขนาดกระดาษในการพิมพ์หรือถ่ายเอกสาร: A3, A4, A5, A6, B5, C5, Custom, Legal
จำนวนหน้า	จำนวนเอกสารที่ใช้บริการในการถ่ายเอกสารหรือพิมพ์เอกสาร
ประเภทการใช้งาน	ประเภทการใช้บริการ: การถ่ายเอกสาร (copy), การพิมพ์เอกสารด้วยเครื่องถ่ายเอกสาร (copyprint), การพิมพ์เอกสารด้วยเครื่องพิมพ์ (print)

ตัวอย่างข้อมูลที่อยู่ในรูปแบบของไฟล์ Excel จำนวน 6 แอตทริบิวต์ 84,764 รายการ แสดงดังตารางที่ 2

ตารางที่ 2 ข้อมูลตัวอย่าง (บางส่วน) ของการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสาร

ID	Date	User*	Page_type	Pages	Type
1	3/08/2554 8:20 AM	User1	A4	1	print
2	8/08/2555 6:50 PM	User2	A4	157	copyprint
3	8/08/2555 6:55 PM	User2	A4	471	copyprint
4	3/08/2554 8:22 AM	User3	A4	1	print
5	8/08/2555 9:20 PM	User6	A3	220	copy

หมายเหตุ * ชื่อผู้ใช้งาน โดยในงานวิจัยนี้ขอใช้ชื่อผู้ใช้งานสมมติ เป็น user1, 2, 3 แทนข้อมูลผู้ใช้งานจริง

2.2 การวิเคราะห์ข้อมูลการใช้งานในภาพรวม ด้วย RapidMiner มีขั้นตอนดังนี้

2.2.1 สร้างโปรเซสโดยใช้ฟังก์ชัน Aggregate เพื่อทำการคำนวณค่าของจำนวนครั้งในการใช้บริการของผู้ใช้แต่ละคน (Frequency) จำนวนปริมาณที่ใช้บริการของผู้ใช้แต่ละคน (Monetary) และคำนวณหาจำนวนวันตั้งแต่วันที่ใช้บริการล่าสุดของผู้ใช้แต่ละคนไปจนถึงวันที่กำหนด (Recency) โดยในงานวิจัยนี้กำหนดให้เป็นวันที่ 1 กุมภาพันธ์ 2559 ซึ่งเป็นวันที่หลังจากข้อมูลการใช้งานสิ้นสุด ตัวอย่างผลที่ได้ (บางส่วน) แสดงดังตารางที่ 3

ตารางที่ 3 ข้อมูลตัวอย่าง (บางส่วน) แสดงผลที่ได้จากการใช้ฟังก์ชัน Aggregate ในการหาค่า RFM เบื้องต้น

ID	User	Recency	Frequency	Monetary	max_date
1	User1	25.35	1963	76359	Wed Jan 06 15:33:00 ICT 2016
2	User2	25.33	308	16814	Wed Jan 06 16:12:00 ICT 2016
3	User3	60.33	25	196	Wed Dec 02 16:08:00 ICT 2015
4	User4	25.52	4651	42034	Wed Jan 06 11:26:00 ICT 2016
5	User5	44.60	764	4043	Fri Dec 18 09:37:00 ICT 2015
6	User6	25.50	894	13610	Wed Jan 06 12:00:00 ICT 2016
7	User7	53.39	18	139	Wed Dec 09 14:44:00 ICT 2015
8	User8	33.52	1515	50256	Tue Dec 29 11:32:00 ICT 2015

2.2.2 แบ่งกลุ่มข้อมูลจำนวนครั้งในการใช้บริการ (Frequency) จำนวนปริมาณที่ใช้บริการ (Monetary) และจำนวนวันตั้งแต่วันที่ใช้บริการล่าสุดจนถึงวันที่กำหนด (Recency) โดยในงานวิจัยนี้ขอเสนอการแบ่งระดับการใช้งานออกเป็น 3 กลุ่ม กลุ่มละ 33.5% ของข้อมูลเพื่อไม่ให้เกิดการแบ่งกลุ่มออกเป็นกลุ่มย่อยมากเกินไป ข้อมูลตัวอย่าง (บางส่วน) แสดงผลที่ได้จากการแบ่งระดับ แสดงดังตารางที่ 4

ตารางที่ 4 ข้อมูลตัวอย่าง (บางส่วน) แสดงผลที่ได้จากการแบ่งกลุ่มข้อมูลออกเป็น 3 ระดับ

ID	User	Recency	Frequency	Monetary
1	User1	3	3	3
2	User2	3	2	2
3	User3	2	1	1
4	User4	3	3	3
5	User5	2	2	2
6	User6	3	2	2
7	User7	2	1	1
8	User8	2	3	3

ในการแบ่งข้อมูลจากค่า Frequency และ Monetary ข้างต้น เป็นการแบ่งโดยค่าน้อยจะกำหนดให้เป็นระดับ 1 ถึงระดับ 3 ที่มีค่ามากที่สุด ในขณะที่การแบ่งกลุ่มของค่า Recency จะตรงกันข้าม คือ ค่าน้อยจะเป็นระดับ 3 ส่วนค่ามากจะเป็นระดับ 1 ดังนั้นจะต้องแปลงข้อมูลดังกล่าวให้ตรงกับความเป็นจริง

2.2.3 คำนวณค่า RFM ให้แก่ผู้ใช้แต่ละคน ให้อยู่ในรูปแบบของเลข 3 หลัก แสดงดังตารางที่ 5

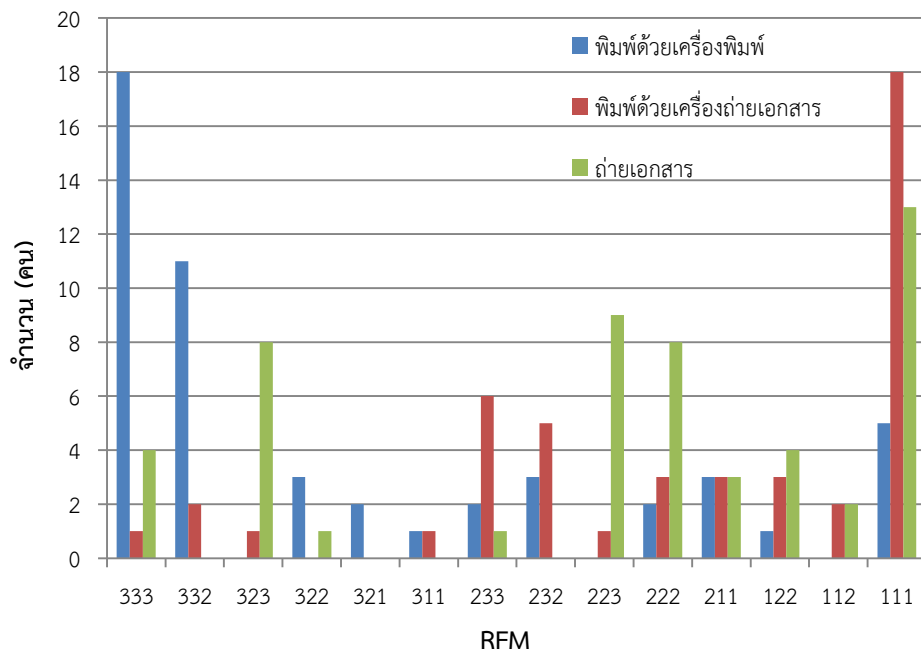
ตารางที่ 5 ข้อมูลตัวอย่าง (บางส่วน) แสดงผลที่ได้หลังจากการแปลงข้อมูลค่า RFM

ID	User	Recency	Frequency	Monetary	RFM
1	User1	3	3	3	333
2	User2	3	2	2	322
3	User3	2	1	1	211
4	User4	3	3	3	333
5	User5	2	2	2	222
6	User6	3	2	2	222
7	User7	2	1	1	211
8	User8	2	3	3	233

ผลการวิจัย

1. ผลการวิเคราะห์พฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากร

ผลการวิเคราะห์ข้อมูลการใช้งานแยกตามประเภทของการจัดทำเอกสาร แบ่งออกเป็น 3 ประเภท คือ การพิมพ์เอกสารทางเครื่องพิมพ์ การพิมพ์เอกสารทางเครื่องถ่ายเอกสาร และการถ่ายเอกสาร จากผลการวิเคราะห์เปรียบเทียบจำนวนผู้ใช้งานในแต่ละประเภทตามแผนภูมิในภาพที่ 2 แสดงให้เห็นถึงพฤติกรรมในการจัดทำเอกสารของบุคลากรในสำนักวิชาสารสนเทศศาสตร์ พบว่ากลุ่มที่มีการใช้งานบ่อยและปริมาณมากซึ่งพิจารณาจากค่า RFM 333 ส่วนใหญ่เป็นผู้ใช้ที่พิมพ์เอกสารทางเครื่องพิมพ์มากกว่าพิมพ์เอกสารด้วยเครื่องถ่ายเอกสารอย่างเห็นได้ชัด โดยจำนวนผู้พิมพ์เอกสารทางเครื่องพิมพ์ 18 คน ในขณะที่จำนวนผู้ที่พิมพ์เอกสารทางเครื่องถ่ายเอกสาร มีเพียง 1 คน เท่านั้น ในขณะที่กลุ่มที่มีการใช้งานในระดับปานกลางและน้อย เช่น RFM 222, 122 และ 111 ส่วนใหญ่มีพฤติกรรมในการพิมพ์เอกสารทางเครื่องถ่ายเอกสารและใช้งานการถ่ายเอกสารมากกว่าการพิมพ์เอกสารทางเครื่องพิมพ์ ส่วนกลุ่มที่ใช้น้อยและไม่บ่อย (RFM 111) นิยมใช้งานเครื่องถ่ายเอกสารมากกว่าการพิมพ์เอกสารด้วยเครื่องพิมพ์



ภาพที่ 2 แผนภูมิเปรียบเทียบจำนวนผู้ใช้งานในแต่ละประเภทจำแนกตามค่า RFM

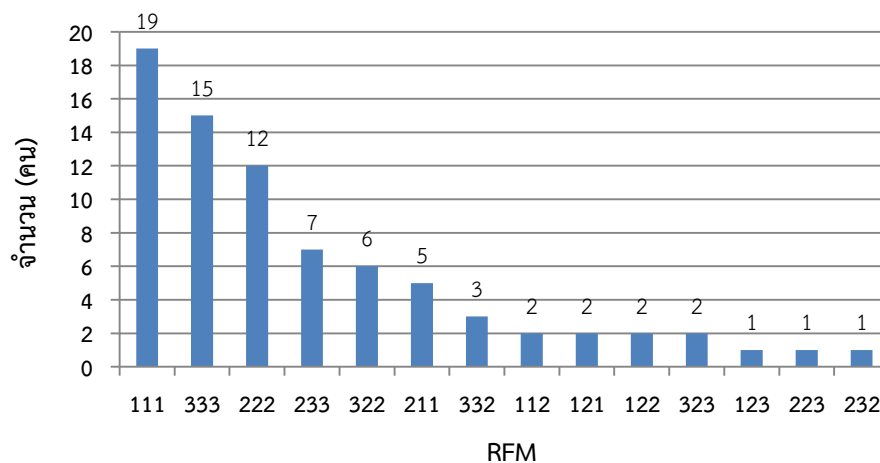
2. ผลการวิเคราะห์การจัดกลุ่มผู้ใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสาร

จากผลการวิเคราะห์ข้อมูลตัวอย่างทั้งหมด พบว่าสามารถแบ่งกลุ่มผู้ใช้งานตามพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารในภาพรวมด้วยแบบจำลอง RFM ออกเป็นกลุ่มที่แตกต่างกันได้จำนวน 14 กลุ่ม ดังตารางที่ 6

ตารางที่ 6 ผลที่ได้จากการจัดกลุ่มผู้ใช้จากการวิเคราะห์แบบ RFM

กลุ่มที่	RFM	จำนวนผู้ใช้ (คน)	Fraction
1	111	19	0.244
2	333	15	0.192
3	222	12	0.154
4	233	7	0.090
5	322	6	0.077
6	211	5	0.064
7	332	3	0.038
8	112	2	0.026
9	121	2	0.026
10	122	2	0.026
11	323	2	0.026
12	123	1	0.013
13	223	1	0.013
14	232	1	0.013

จากตารางที่ 6 แสดงการกระจายตัวของกลุ่มต่างๆ ซึ่งจะเห็นว่ากลุ่มที่มีจำนวนผู้ใช้งานมากที่สุด คือ กลุ่ม 111 ซึ่งเป็นกลุ่มที่ไม่ค่อยมีการใช้บริการคอมพิวเตอร์และเครื่องถ่ายเอกสาร มีจำนวน 19 คน คิดเป็น 24.4% รองลงมาคือกลุ่ม 333 ซึ่งเป็นกลุ่มที่มีการใช้งานคอมพิวเตอร์และเครื่องถ่ายเอกสารบ่อยและมากที่สุด มีจำนวน 15 คน คิดเป็น 19.2% ซึ่งจากผลที่วิเคราะห์ได้นี้หน่วยงานสามารถเข้าไปข้อมูลการใช้งานของผู้ใช้ในแต่ละกลุ่มได้ว่ามีใครบ้าง เพื่อนำไปใช้ประโยชน์ในการตรวจสอบต่อไป ส่วนภาพที่ 3 เป็นแผนภูมิแสดงจำนวนผู้ใช้งานในแต่ละกลุ่มแยกตามค่า RFM



ภาพที่ 3 แผนภูมิแสดงจำนวนผู้ใช้งานในแต่ละกลุ่มตามค่า RFM

อภิปรายผลการวิจัย

จากผลการวิเคราะห์เกี่ยวกับพฤติกรรมในการจัดทำเอกสารของบุคลากรในสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ พบว่ากลุ่มที่มีการใช้งานบ่อยและปริมาณมากเป็นกลุ่มที่น่าสนใจมากที่สุด เพราะเป็นกลุ่มที่ส่งผลกระทบต่อค่าใช้จ่ายขององค์กรมากที่สุด ซึ่งพบว่าในกลุ่มดังกล่าวส่วนใหญ่นิยมทำเอกสารโดยการพิมพ์เอกสารทางเครื่องพิมพ์มากกว่าการพิมพ์เอกสารทางเครื่องถ่ายเอกสาร ซึ่งความรู้ที่ได้รับสามารถนำมาใช้ในการวางแผนการลดค่าใช้จ่าย เช่น ปกติแล้วค่าใช้จ่ายในการพิมพ์เอกสารทางเครื่องพิมพ์จะมีต้นทุนสูงกว่าการพิมพ์ทางเครื่องถ่ายเอกสาร หน่วยงานสามารถนำความรู้ที่ค้นพบไปส่งเสริมให้บุคลากรในกลุ่มดังกล่าวมีการใช้งานโดยพิมพ์ออกเครื่องถ่ายเอกสารให้มากขึ้นเพื่อช่วยลดค่าใช้จ่ายในองค์กรได้ เป็นการนำความรู้ที่ได้จากการวิเคราะห์ไปประยุกต์ใช้ในการทำงานให้มีประสิทธิภาพมากขึ้น ส่วนผลของการแบ่งกลุ่มผู้ใช้ซึ่งสามารถนำไปใช้ในการตรวจสอบพฤติกรรมการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของพนักงานในกลุ่มที่มีการใช้งานบ่อยและมีปริมาณมากเพื่อตรวจสอบว่าเป็นบุคลากรในกลุ่มดังกล่าวที่มีหน้าที่รับผิดชอบในการทำงานที่เกี่ยวข้องกับการจัดทำเอกสารหรือไม่ โดยในงานวิจัยนี้พบผู้ใช้งานบ่อยและมีปริมาณมาก จำนวน 15 คน จากการตรวจสอบรายชื่อพบว่าผู้ใช้งาน 4 คน เป็นเจ้าหน้าที่ธุรการและเจ้าหน้าที่บริหารงานทั่วไป ที่มีหน้าที่จัดทำเอกสาร ส่วนอีก 11 คนเป็นอาจารย์และผู้ประสานงานหลักสูตรที่รับผิดชอบงานโครงการจำนวนมาก ซึ่งไม่พบความผิดปกติในการใช้งานที่เห็นได้ชัดเจน

ข้อเสนอแนะ

งานวิจัยฉบับนี้นำเสนอวิธีการในการค้นหาความรู้โดยการทำเหมืองข้อมูลด้วยการจัดกลุ่มจากข้อมูลการใช้งานเครื่องพิมพ์และเครื่องถ่ายเอกสารของบุคลากรสำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ เพื่อทำการจัดกลุ่มผู้ใช้งานโดยการประยุกต์ใช้วิธีการวิเคราะห์ด้วยแบบจำลอง RFM กับข้อมูลดังกล่าว การนำเสนอตามวิธีการดังกล่าวสามารถนำไปประยุกต์ใช้ประโยชน์กับการวิเคราะห์ข้อมูลองค์กรอื่นๆ ในลักษณะงานที่คล้ายคลึงกันได้ แต่อย่างไรก็ตามเทคนิคในการแบ่งกลุ่มมีหลายรูปแบบ ซึ่งแต่ละวิธีอาจจะให้ผลลัพธ์ที่แตกต่างกัน ดังนั้นการเลือกใช้วิธีการใดขึ้นอยู่กับวัตถุประสงค์ของการวิเคราะห์และประโยชน์ที่ต้องการเพื่อใช้งานมากกว่า

เอกสารอ้างอิง

- [1] ศูนย์เทคโนโลยีดิจิทัล มหาวิทยาลัยวลัยลักษณ์. (2559). ระบบควบคุมงานพิมพ์และถ่ายเอกสาร. มหาวิทยาลัยวลัยลักษณ์, นครศรีธรรมราช.
- [2] เอกสิทธิ์ พัทธวงศ์ศักดิ์. (2557). *An Introduction to Data Mining Techniques*. กรุงเทพฯ : สำนักพิมพ์เอเชีย ดิจิตอลการพิมพ์.
- [3] Han J. and Kamber M. (2006). *Data Mining Concepts and Techniques*. 2nd ed. San Francisco: Morgan Kaufmann.
- [4] Tan P. N., Steinbach M., and Kumar V. (2005). *Introduction to Data Mining*. Boston: Pearson.
- [5] Stephan K.. (2004). *Managing Data Mining*. London: CyberTechPublishing.
- [6] เอกสิทธิ์ พัทธวงศ์ศักดิ์. (2558). *Introduction to Business Analytics with RapidMiner Studio 6*. กรุงเทพฯ: สำนักพิมพ์เอเชีย ดิจิตอลการพิมพ์.
- [7] Wei Jo-Ting, Lin Shih-Yen and Wu Hsin-Hung. (2010). A review of the application of RFM model. *African Journal of Business Management*, 4(19), 4199-4206.
- [8] Fader P. S., Hardie B. G., and Lee K. L. (2005). RFM and CLV: Using iso-value curves for customer based analysis. *Journal of Marketing Research*, 42(4), 415-430.
- [9] Soeini R. A. and Fathalizade E. (2012). Customer Segmentation based on Modified RFM Model in the Insurance Industry. In *Proceedings of 2012 4th International Conference on Machine Learning and Computing*. 101-104. Singapore: IACSIT Press.

- [10] Dursan A. and Caber M. (2016). Using data mining techniques for profiling profitable hotel customer: An application of RFM analysis. *Journal of Tourism Management*, 18, 153-160.
- [11] เอกสิทธิ์ พัทธวงศ์ศักดิ์. (2558). *Introduction to Business Analytics with RapidMiner Studio 6*. กรุงเทพฯ: สำนักพิมพ์เอเชีย ดิจิตอลการพิมพ์.
- [12] เอกสิทธิ์ พัทธวงศ์ศักดิ์. (2557). *Practical Data Mining with RapidMiner Studio 6*. กรุงเทพฯ: สำนักพิมพ์เอเชีย ดิจิตอลการพิมพ์.