

การเปรียบเทียบประสิทธิภาพระหว่าง FP-Growth และ Apriori ในการวิเคราะห์ รูปแบบการซื้อสินค้าในร้านกาแฟ

กัญญ์ณัฐ สุริยันต์^{1*}

Received : February 6, 2025

Revised : June 18, 2025

Accepted : June 21, 2025

บทคัดย่อ

การเปรียบเทียบประสิทธิภาพระหว่าง FP-Growth และ Apriori ในการวิเคราะห์รูปแบบการซื้อสินค้าในร้านกาแฟ มีวัตถุประสงค์เพื่อศึกษากฎความสัมพันธ์ (Association Rules) และเปรียบเทียบประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth ในการวิเคราะห์ธุรกรรมการซื้อสินค้าของผู้บริโภคในร้านกาแฟ โดยใช้ข้อมูลธุรกรรมจากเว็บไซต์ Kaggle จำนวน 1,000 รายการ ผ่านกระบวนการเหมืองข้อมูลด้วยภาษา Python และไลบรารี mlxtend ผลการวิจัยพบว่า ทั้งสองอัลกอริทึมสามารถสร้างกฎความสัมพันธ์ที่มีค่า Confidence สูง เช่น กาแฟ (Latte) → (Croissant) มีค่า Confidence 85% อย่างไรก็ตาม FP-Growth มีประสิทธิภาพสูงกว่า Apriori โดยใช้เวลาในการประมวลผล 0.01 วินาที และหน่วยความจำ 30.34 KB ขณะที่ Apriori ใช้เวลา 0.00 วินาที และหน่วยความจำ 30.33 KB ซึ่งเหมาะกับชุดข้อมูลขนาดเล็ก ทั้งนี้ FP-Growth มีข้อได้เปรียบในด้านโครงสร้างข้อมูลแบบ FP-Tree ที่สามารถลดการสแกนข้อมูลซ้ำจึงเหมาะสมกับข้อมูลขนาดใหญ่และซับซ้อนมากกว่า

คำสำคัญ: กฎความสัมพันธ์ การวิเคราะห์ตะกร้าสินค้า เหมืองข้อมูล อัลกอริทึม Apriori
อัลกอริทึม FP-Growth

¹ อาจารย์ประจำหลักสูตรบริหารธุรกิจบัณฑิต สาขาวิชาธุรกิจดิจิทัล คณะดิจิทัล วิทาลัยเทคโนโลยีสยาม

อีเมล: kunyanuts@siamtechno.ac.th

* ผู้นิพนธ์หลัก อีเมล: kunyanuts@siamtechno.ac.th

COMPARING THE PERFORMANCE BETWEEN FP-GROWTH AND APRIORI ALGORITHMS FOR ANALYZING SHOPPING PATTERNS IN A COFFEE SHOP

Kanyanut Suriyan^{1*}

Abstract

The current study, namely Comparison of FP-Growth and Apriori Algorithms for Analyzing Shopping Patterns in a Coffee Shop, was aimed to investigate association rules and compare the performance of the Apriori and FP-Growth algorithms in analyzing customer purchase transactions in a coffee shop. The dataset, consisting of 1,000 transactions, was obtained from the Kaggle website then analyzed using data mining techniques with the Python programming language and the mlxtend library. The research results were revealed that both algorithms could generate association rules with high confidence values. For example, the rule (Latte) → (Croissant) achieved a confidence of 85%. However, FP-Growth algorithms demonstrated the better performance than Apriori's, spending 0.01 seconds of processing time and using 30.34 KB of memory. On the other hand, Apriori algorithms spending 0.00 seconds and using 30.33 KB. Although Apriori algorithms are suitable for small datasets, FP-Growth offers a structural advantage through its use of the FP-Tree, which reduces redundant data scans. FP-Growth algorithms are more suitable for large and complex datasets.

Keywords: Association Rules, Market Basket Analysis, Data Mining, Apriori Algorithm, FP-Growth Algorithm

¹ Lecturer of Digital Business Program in Digital Faculty, Siam Technology College,
e-mail: kunyanuts@siamtechno.ac.th

* Corresponding author, e-mail: kunyanuts@siamtechno.ac.th

บทนำ

ในสภาพแวดล้อมทางธุรกิจที่มีการแข่งขันสูง ผู้ประกอบการจำเป็นต้องพัฒนากลยุทธ์ และปรับปรุงการดำเนินงานอย่างต่อเนื่อง เพื่อรักษาความได้เปรียบในตลาด หนึ่งในแนวทางสำคัญ คือการใช้ข้อมูลธุรกรรม (Transaction Data) ซึ่งเป็นแหล่งข้อมูลที่สะท้อนพฤติกรรมของผู้บริโภค และสามารถนำมาวิเคราะห์เพื่อเพิ่มประสิทธิภาพด้านการขาย และการตลาด (Sun et al., 2024) เทคนิคเหมืองข้อมูล (Data Mining) เป็นกระบวนการที่ช่วยให้สามารถค้นหาสารสนเทศที่ซ่อนอยู่ในข้อมูลขนาดใหญ่ เพื่อนำไปใช้ในการตัดสินใจ เช่น การวิเคราะห์แนวโน้ม การคาดการณ์ หรือการจัดกลุ่มลูกค้า หนึ่งในเทคนิคที่สำคัญในการวิเคราะห์ข้อมูลลูกค้า คือ การวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis) ซึ่งช่วยค้นหาความสัมพันธ์ระหว่างสินค้าที่ลูกค้ามักซื้อพร้อมกัน โดยข้อมูลที่ได้นำมาใช้ในการวางแผนโปรโมชั่น การจัดเรียงสินค้า หรือการกำหนดราคาที่เหมาะสม (Istrat & Lalic, 2017; Chusna et al., 2021) การขาดการประยุกต์ใช้อัลกอริทึม FP-Growth และ Apriori กับข้อมูลธุรกรรมจริงของธุรกิจร้านค้าในประเทศไทย โดยเฉพาะการเปรียบเทียบด้านประสิทธิภาพเชิงเทคนิค ได้แก่เวลาในการประมวลผล การใช้หน่วยความจำ และคุณภาพของกฎที่ได้ เช่นค่า Confidence และ Lift ซึ่งมีความสำคัญต่อการใช้งานจริงของผู้ประกอบการ อัลกอริทึมที่นิยมใช้ในกระบวนการนี้ ได้แก่ Apriori และ FP-Growth โดย Apriori ทำงานโดยสร้าง Candidate Itemsets และสแกนฐานข้อมูลซ้ำหลายครั้ง ซึ่งอาจใช้ทรัพยากรจำนวนมากหากข้อมูลมีขนาดใหญ่ (Mustakim et al., 2018) ขณะที่ FP-Growth ใช้โครงสร้าง Frequent Pattern Tree (FP-Tree) เพื่อลดจำนวนครั้งในการสแกนข้อมูล ช่วยเพิ่มประสิทธิภาพ โดยเฉพาะกับฐานข้อมูลขนาดใหญ่ (Hunyadi, 2011) จากเหตุผลข้างต้น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษา และเปรียบเทียบประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth ในการวิเคราะห์รูปแบบการซื้อสินค้าของลูกค้า โดยมีมุ่งหวังให้นำผลที่ได้ไปประยุกต์ใช้กับกลยุทธ์ธุรกิจค้าปลีกได้อย่างมีประสิทธิภาพ

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษากฎความสัมพันธ์ (Association Rules) ของอัลกอริทึม Apriori, และ FP-Growth
2. เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth โดยมีตัวแปรในการประเมิน

Processing Time Memory Usage Iterations

วิธีดำเนินการวิจัย

การวิจัยนี้ใช้กระบวนการแสดงผลโดยอ้างอิงจากชุดข้อมูล (Dataset) ซึ่งได้มาจากเว็บไซต์ Kaggle.com โดยเป็นข้อมูลธุรกรรมจากร้านขายกาแฟ และดำเนินการตามขั้นตอนดังนี้

1. เริ่มต้นโดยการนำชุดข้อมูลจากเว็บไซต์ Kaggle.com ที่เกี่ยวข้องกับธุรกรรมการซื้อกาแฟและสินค้าเสริม โดยชุดข้อมูลมีจำนวนทั้งสิ้น 1,000 ธุรกรรม โดยแต่ละแถวในชุดข้อมูลแสดงถึงธุรกรรมเดียวที่ลูกค้าทำการซื้อสินค้าจากร้านกาแฟ

2. การเตรียมข้อมูล (Data Preprocessing) ก่อนการนำชุดข้อมูลไปวิเคราะห์ ผู้วิจัยได้ดำเนินการทำความสะอาดข้อมูล (Data Cleaning) โดยมีขั้นตอนดังนี้

2.1 การจัดการค่าที่สูญหาย (Missing Values) ตรวจสอบรายการที่ไม่มีข้อมูลสินค้า แล้วลบรายการที่มีช่องว่างหรือไม่มีรายการสินค้าใดเลยออกจากชุดข้อมูล

2.2 การจัดการข้อมูลผิดปกติ (Outliers) พิจารณาความถี่ในการซื้อของสินค้าที่เกิดขึ้นน้อยมาก เช่น ปรากฏในไม่ถึง 1% ของธุรกรรมทั้งหมด แล้วตัดออกจากชุดวิเคราะห์เพื่อป้องกันการรบกวนค่าความสัมพันธ์

2.3 การปรับโครงสร้างข้อมูล แปลงข้อมูลให้อยู่ในรูปแบบ Itemsets โดยให้แต่ละ Transaction แสดงรายการสินค้าที่ซื้อ เพื่อให้เหมาะสมกับการประมวลผลโดยอัลกอริทึม Apriori และ FP-Growth

3. เกณฑ์การประเมินประสิทธิภาพของอัลกอริทึม งานวิจัยนี้ใช้เกณฑ์มาตรฐาน 3 ตัวเพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึม ได้แก่

3.1 ค่า Support ใช้เกณฑ์ขั้นต่ำที่ 0.05 หมายถึงรูปแบบรายการต้องปรากฏอย่างน้อย 5% ของธุรกรรมทั้งหมด จึงจะถือว่า “เกิดบ่อย”

3.2 ค่า Confidence ตั้งเกณฑ์ขั้นต่ำที่ 0.6 เพื่อให้มั่นใจว่าความสัมพันธ์ที่พบมีความน่าเชื่อถือพอสมควร

3.3 ค่า Lift ใช้เพื่อประเมินว่าความสัมพันธ์ที่พบมีความสัมพันธ์เชิงบวก (ค่า > 1) หรือเป็นเพียงความบังเอิญ (ค่า ≈ 1)

เหตุผลในการตั้งค่าเกณฑ์เหล่านี้เพื่อให้ได้กฎความสัมพันธ์ที่มีความน่าเชื่อถือ ปริมาณไม่มากเกินไป และสามารถนำไปใช้งานในเชิงกลยุทธ์ได้จริง โดยอิงตามแนวทางในงานวิจัยของ Mustakim et al. (2018) และ Chusna et al. (2021)

ตารางที่ 1 แสดงจำนวน Attribute ของการหากฎความสัมพันธ์

Attribute name	Type	ความหมาย
TransactionID	Integer	รหัสธุรกรรมที่แสดงถึงการซื้อขายแต่ละครั้ง
Spring	Boolean	ฤดูใบไม้ผลิ ระบุว่าธุรกรรมเกิดในฤดูใบไม้ผลิหรือไม่
Summer	Boolean	ฤดูร้อน ระบุว่าธุรกรรมเกิดในฤดูร้อนหรือไม่
Fall	Boolean	ฤดูใบไม้ร่วง ระบุว่าธุรกรรมเกิดในฤดูใบไม้ร่วงหรือไม่

ตารางที่ 1 แสดงจำนวน Attribute ของการหากฎความสัมพันธ์ (ต่อ)

Attribute name	Type	ความหมาย
Winter	Boolean	ฤดูหนาว ระบุว่าธุรกรรมเกิดในฤดูหนาวหรือไม่
Latte	Boolean	ลูกค้าซื้อเมนู Latte หรือไม่
Hazelnut Latte	Boolean	ลูกค้าซื้อเมนู Hazelnut Latte หรือไม่
Vanilla Latte	Boolean	ลูกค้าซื้อเมนู Vanilla Latte หรือไม่
Creamy Aren Latte	Boolean	ลูกค้าซื้อเมนู Creamy Aren Latte หรือไม่
Caramel Latte	Boolean	ลูกค้าซื้อเมนู Caramel Latte หรือไม่
Matcha Latte	Boolean	ลูกค้าซื้อเมนู Matcha Latte หรือไม่
Cappuccino	Boolean	ลูกค้าซื้อเมนู Cappuccino หรือไม่
Caramel Macchiato	Boolean	ลูกค้าซื้อเมนู Caramel Macchiato หรือไม่
Americano	Boolean	ลูกค้าซื้อเมนู Americano หรือไม่
Matcha Bun	Boolean	ลูกค้าซื้อขนม Matcha Bun หรือไม่
Cheese Toast	Boolean	ลูกค้าซื้อเมนู Cheese Toast หรือไม่
Chocolate Toast	Boolean	ลูกค้าซื้อเมนู Chocolate Toast หรือไม่
Butter Toast	Boolean	ลูกค้าซื้อเมนู Butter Toast หรือไม่
Cookies	Boolean	ลูกค้าซื้อเมนู Cookies หรือไม่
Croissant	Boolean	ลูกค้าซื้อเมนู Croissant หรือไม่
Egg Salad Toast	Boolean	ลูกค้าซื้อเมนู Egg Salad Toast หรือไม่
Fries	Boolean	ลูกค้าซื้อเมนู Fries หรือไม่
Crispy Chicken	Boolean	ลูกค้าซื้อเมนู Crispy Chicken Nuggets หรือไม่
Nuggets		

จากตารางที่ 1 ตารางแสดงจำนวน attribute ของการทำธุรกรรมของลูกค้าร้านกาแฟประกอบด้วย attribute name ประเภทข้อมูล (Type) Integer ใช้สำหรับรหัสธุรกรรม Boolean ใช้สำหรับแสดงว่าลูกค้าซื้อสินค้านั้นหรือไม่ (0 = ไม่ซื้อ, 1 = ซื้อ) การใช้งาน สามารถใช้ Attribute เหล่านี้ในการวิเคราะห์ Association Rules เพื่อดูว่าลูกค้ามักซื้อสินค้าร่วมกันอย่างไร เช่น ลูกค้าที่ซื้อ Latte มักจะซื้อ Croissant ด้วยหรือไม่

4. การสร้างกฎความสัมพันธ์ (Association Rules) มีขั้นตอนดังนี้

4.1 การกำหนด Frequent Itemsets

1. ใช้อัลกอริทึม Apriori และ FP-Growth เพื่อตรวจหาชุดรายการที่เกิดขึ้นบ่อย (Frequent Itemsets) จากข้อมูลธุรกรรม

เช่น เมนูเครื่องดื่มและอาหารที่ลูกค้าสั่งซื้อในแต่ละธุรกรรม ตัวอย่างชุดข้อมูลที่เกิดขึ้นบ่อย ["Latte", "Croissant"] ["Cappuccino", "Brownie"]

2. การคำนวณกฎความสัมพันธ์ (Association Rules) ใช้ไลบรารี mlxtend เพื่อสร้างกฎความสัมพันธ์จาก Frequent Itemsets ที่ได้มา กฎ ถ้าลูกค้าซื้อ "Latte" มีโอกาส 70% ที่จะซื้อ "Croissant" ด้วย

5. การประเมินผล (Evaluation)

5.1 ประสิทธิภาพการทำงาน (Performance) วัดประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth ในการค้นหาชุดรายการที่เกิดขึ้นบ่อย (Frequent Itemsets) และสร้างกฎความสัมพันธ์ (Association Rules) ใช้ตัวชี้วัด เช่น จำนวน Iterations หรือ Processing Steps ที่ใช้ในแต่ละอัลกอริทึม การประยุกต์ในร้านกาแฟตรวจสอบว่าการวิเคราะห์พฤติกรรมการซื้อในแต่ละฤดูกาล เช่น การจับคู่ระหว่าง "Latte" และ "Croissant" สามารถทำได้อย่างรวดเร็วและแม่นยำเพียงใด

5.2 การใช้หน่วยความจำ (Memory Usage) ตรวจสอบการใช้ทรัพยากรหน่วยความจำขณะประมวลผลอัลกอริทึม วัดปริมาณหน่วยความจำที่ใช้สำหรับการจัดเก็บ Candidate Itemsets และ Frequent Itemsets การประยุกต์ในร้านกาแฟ เปรียบเทียบว่าอัลกอริทึม Apriori หรือ FP-Growth ใช้หน่วยความจำอย่างมีประสิทธิภาพมากกว่าในการประมวลผลข้อมูล เช่น เมนูที่ลูกค้าสั่งในแต่ละฤดูกาล

5.3 เวลาในการประมวลผล (Processing Time) วัดระยะเวลาที่ใช้ในการประมวลผลข้อมูลธุรกรรมทั้งหมดจนได้ผลลัพธ์ที่ต้องการ ใช้ตัวชี้วัด เช่น Total Execution Time หรือ Time per Iteration ประเมินว่าอัลกอริทึมใดเหมาะสำหรับการประมวลผลข้อมูลขนาดใหญ่ เช่น รายการซื้อที่มีทั้ง "Latte", "Cappuccino", และ "Muffin" ในช่วงเวลาต่าง ๆ

6. สิ้นสุดกระบวนการ (End) สรุปผลการวิจัยและเปรียบเทียบประสิทธิภาพของอัลกอริทึมทั้งสาม เพื่อระบุอัลกอริทึมที่เหมาะสมที่สุดสำหรับการวิเคราะห์ธุรกรรมการซื้อสินค้าในร้านขายกาแฟ

ผลการวิจัยและอภิปรายผล

ในขั้นตอนนี้ ผู้วิจัยได้ดำเนินการสร้างแบบจำลองข้อมูลด้วยเทคนิคเหมืองข้อมูลโดยใช้วิธีการหากฎความสัมพันธ์ (Association Rules) ด้วยอัลกอริทึม FP-Growth และ Apriori สำหรับเครื่องมือที่ใช้ในการ

วิเคราะห์ข้อมูลครั้งนี้ ประกอบด้วยโปรแกรม Microsoft Excel สำหรับการจัดการและตรวจสอบข้อมูลเบื้องต้น, Google Colab สำหรับการประมวลผลข้อมูล และ Python ร่วมกับไลบรารี mlxtend สำหรับการสร้าง Frequent Itemsets และ Association Rules โดยมีเป้าหมายเพื่อให้ได้ผลการวิเคราะห์ที่ครอบคลุมและมีประสิทธิภาพสูงสุดในการค้นหาความสัมพันธ์ระหว่างสินค้าต่าง ๆ สามารถอธิบายกฎความสัมพันธ์ของอัลกอริทึม Apriori, และ FP-Growth ได้จากตารางต่อไปนี้

1. ศึกษากฎความสัมพันธ์ (Association Rules) ของอัลกอริทึม Apriori, และ FP-Growth ได้ผลการศึกษาในตารางที่ 2 และตารางที่ 3

ตารางที่ 2 กฎความสัมพันธ์ระหว่างรายการสินค้าที่ลูกค้ามักซื้อพร้อมกัน โดยใช้อัลกอริทึม Apriori

Antecedents รายการสินค้าที่ เกิดขึ้นก่อน	Consequents รายการสินค้าที่ เกิดขึ้นตามมา	Support ความถี่ที่ Antecedents และ Consequents ปรากฏร่วมกัน ในข้อมูล	Confidence ความน่าจะเป็น ที่ลูกค้าจะซื้อ Consequents เมื่อซื้อ Antecedents แล้ว	Lift ความแข็งแรง ของความสัมพันธ์ (> 1 คือสัมพันธ์ กันมากกว่าความ บังเอิญ)	Leverage ความแตกต่าง ระหว่างความถี่ จริงกับความถี่ ที่คาดว่าจะเกิด	Conviction ตัวชี้วัดความ น่าเชื่อถือ ของกฎ (ยิ่งมากยิ่ง สัมพันธ์แน่น)
(Spring_0)	(Latte_0)	0.460	0.868	1.007	0.003140	1.044857
(Spring_0)	(Hazelnut Latte_0)	0.439	0.828	1.002	0.000690	1.007582
(Spring_0)	(Vanilla Latte_0)	0.447	0.843	0.991	-0.004030	0.951446
(Spring_0)	(Creamy Aren Latte_0)	0.452	0.853	1.020	0.008920	1.114359
(Spring_0)	(Caramel Latte_0)	0.435	0.821	0.975	-0.011260	0.881474
(Butter Toast_0, Americano_0, Chocolate Toast_0)	(Egg Salad Toast_0, Latte_0, Cappuccino_0)	0.301	0.622	1.047	0.013504	1.073792
(Butter Toast_0, Chocolate Toast_0, Cappuccino_0)	(Americano_0, Egg Salad Toast_0, Latte_0)	0.301	0.613	1.034	0.009837	1.051774

ตารางที่ 2 กฎความสัมพันธ์ระหว่างรายการสินค้าที่ลูกค้ามักซื้อร่วมกัน โดยใช้อัลกอริทึม Apriori (ต่อ)

Antecedents รายการสินค้าที่ เกิดขึ้นก่อน	Consequents รายการสินค้าที่ เกิดขึ้นตามมา	Support ความถี่ที่ Antecedents และ Consequents ปรากฏร่วมกัน ในข้อมูล	Confidence ความน่าจะเป็น ที่ลูกค้าจะซื้อ Consequents เมื่อซื้อ Antecedents แล้ว	Lift ความแข็งแรง ของความสัมพันธ์ (> 1 คือสัมพันธ์ กันมากกว่าความ บังเอิญ)	Leverage ความแตกต่าง ระหว่างความถี่ จริงกับความถี่ ที่คาดว่าจะเกิด	Conviction ตัวชี้วัดความ น่าเชื่อถือ ของกฎ (ยิ่งมากยิ่ง สัมพันธ์แน่น)
(Americano_0, Chocolate Toast_0, Latte_0, Cappuccino_0)	(Butter Toast_0, Egg Salad Toast_0, Matcha Latte_0)	0.301	0.609	1.047	0.013492	1.069907
(Americano_0, Chocolate Toast_0, Latte_0, Cappuccino_0)	(Butter Toast_0, Egg Salad Toast_0, Cappuccino_0)	0.301	0.602	1.031	0.009000	1.045226
(Americano_0, Chocolate Toast_0, Cappuccino_0, ...)	(Butter Toast_0, Egg Salad Toast_0, Latte_0)	0.301	0.622	1.025	0.007212	1.039410

จากตารางที่ 2 สรุปผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules) ระหว่างสินค้าต่าง ๆ โดยเฉพาะรายการสินค้าที่ลูกค้ามักซื้อร่วมกัน โดยใช้ อัลกอริทึม Apriori ดังนี้

1. ความสัมพันธ์ระหว่างฤดูใบไม้ผลิ (Spring_0) กับเครื่องดื่มต่าง ๆ

1.1 เมื่อลูกค้าเข้าร้านในช่วงฤดูใบไม้ผลิ มักจะซื้อเครื่องดื่มประเภทลาเต้ เช่น ลาเต้ (Latte), เฮเซลนัทลาเต้, วานิลลาลาเต้, ครีมมีอาเรนลาเต้, คาราเมลลาเต้ ยกตัวอย่างเช่น

(Spring_0) → (Latte_0) มีค่า ความน่าจะเป็น = 86.8% แปลว่าลูกค้า 100 คนที่มาซื้อของในฤดูนี้ จะมีประมาณ 87 คนที่ซื้อลาเต้ด้วย

1.2 ความสัมพันธ์ส่วนใหญ่ในกลุ่มนี้มีค่า Lift ใกล้ 1 ซึ่งหมายถึงพฤติกรรมที่ซื้อที่ "พบบ่อย" แต่ไม่ใช่ความสัมพันธ์ที่แข็งแกร่งมากนัก

2. กลุ่มขนมปังและกาแฟ ที่มักซื้อเป็นชุด

2.1 ลูกค้าที่ซื้อ ขนมปังเนย (Butter Toast), ช็อกโกแลตโทสต์, อเมริกาโน่ มักจะซื้อพร้อมกับลาเต้, คาปูชิโน และไข่ขนมปัง (Egg Salad Toast) ยกตัวอย่าง กฎความสัมพันธ์

(Butter Toast, Americano, Chocolate Toast) → (Egg Salad Toast, Latte, Cappuccino)

มีค่า Confidence = 62.2% และ Lift = 1.047 แปลว่าชุดเมนูนี้มีโอกาสสูงที่จะถูกซื้อร่วมกันจริง ไม่ใช่แค่ความบังเอิญ

ตารางที่ 3 กฎความสัมพันธ์ระหว่างรายการสินค้าที่ลูกค้ามักซื้อร่วมกัน โดยใช้อัลกอริทึม FP-Growth

Antecedents รายการสินค้าที่ เกิดขึ้นก่อน	Consequents รายการสินค้าที่ เกิดขึ้นตามมา	Support ความถี่ที่ Antecedents และ Consequents ปรากฏร่วมกันใน ข้อมูล	Confidence ความน่าจะเป็นที่ ลูกค้าจะซื้อ Consequents เมื่อซื้อ Antecedents แล้ว	Lift ความแข็งแรง ของความสัมพันธ์ (> 1 คือสัมพันธ์ กันมากกว่าความ บังเอิญ)	Leverage ความ แตกต่าง ระหว่าง ความถี่จริง กับความถี่ที่ คาดว่าจะ เกิด	Conviction ตัวชี้วัดความ น่าเชื่อถือของ กฎ (ยิ่งมากยิ่ง สัมพันธ์แน่น)
(Butter Toast_0)	(Latte_0)	0.726	0.861	0.999	-0.000666	0.994308
(Latte_0)	(Butter Toast_0)	0.726	0.842	0.999	-0.000666	0.995103
(Butter Toast_0)	(Vanilla Latte_0)	0.711	0.843	0.991	-0.006393	0.951568
(Vanilla Latte_0)	(Butter Toast_0)	0.711	0.835	0.991	-0.006393	0.954336
(Butter Toast_0, Vanilla Latte_0)	(Latte_0)	0.607	0.854	0.990	-0.005882	0.943442
(Vanilla Latte_0, Latte_0, Fall_1)	(Matcha Bun_0)	0.322	0.856	1.023	0.007288	1.134963
(Matcha Bun_0, Fall_1)	(Vanilla Latte_0, Latte_0)	0.322	0.747	1.026	0.008232	1.075523
(Latte_0, Fall_1)	(Matcha Bun_0, Vanilla Latte_0)	0.322	0.744	1.056	0.017168	1.154667
(Vanilla Latte_0, Fall_1)	(Matcha Bun_0, Latte_0)	0.322	0.740	1.021	0.006625	1.058628
(Fall_1)	(Matcha Bun_0, Latte_0, Vanilla Latte_0)	0.322	0.645	1.065	0.019606	1.110768

จากตารางที่ 3 สรุปผลการวิเคราะห์ข้อมูลด้านพฤติกรรมการซื้อขายสินค้าที่มักเกิดขึ้นร่วมกันของลูกค้าของกฎความสัมพันธ์ Association Rules (FP-Growth)

1. ขนมปังเนยกับลาเต้

- 1.1 ลูกค้ามักซื้อ ขนมปังเนย (Butter Toast) คู่กับ ลาเต้ (Latte) เป็นประจำ
- 1.2 พบว่ามากกว่า 72% ของลูกค้าซื้อสองรายการนี้พร้อมกัน
- 1.3 โอกาสที่จะซื้อทั้งสองอย่างคู่กันอยู่ที่ประมาณ 84–86%

2. ขนมปังเนยกับวานิลลาลาเต้ (Vanilla Latte)

- 2.1 ลูกค้ามีแนวโน้มซื้อขนมปังเนยคู่กับวานิลลาลาเต้เช่นกัน
- 2.2 ซื้อพร้อมกันในกว่า 71% ของกรณี และมีโอกาสเกิดขึ้น 83%

3. ความสัมพันธ์กับฤดูใบไม้ร่วง (Fall_1)

3.1 ในฤดูใบไม้ร่วง ลูกค้ามักซื้อเมนูพิเศษ เช่น มัทฉะบัน (Matcha Bun) ร่วมกับลาเต้ และวานิลลาลาเต้ เช่น เมื่อมีฤดูใบไม้ร่วง รวมกับเครื่องดื่ม → ลูกค้ามีโอกาสซื้อมัทฉะบันสูงถึง 85.6%

จากการวิเคราะห์ พบว่าความสัมพันธ์นี้ “แข็งแรง” กว่าคู่ทั่วไป (ค่าที่เรียกว่า Lift > 1 หมายถึงสัมพันธ์กันจริง ไม่ใช่แค่บังเอิญ)

4. ซื้อสลับกันหลายทาง ซื้ออะไรก่อนได้ ลูกค้าอาจเริ่มจากลาเต้ หรือวานิลลาลาเต้ หรือมัทฉะบันก่อน แล้วค่อยซื้ออีกอย่างตาม ยกตัวอย่างกฎความสัมพันธ์

กฎ	Support	Confidence
Latte → Butter Toast	0.726	0.842
Butter Toast → Latte	0.726	0.861

โดยพิจารณาจาก ค่า Support (ความถี่ของการซื้อร่วม) และ Confidence (ความน่าจะเป็นของการซื้ออีกสินค้าตามมา) ที่ใกล้เคียงกัน แสดงให้เห็นว่าลูกค้ามีความชอบที่ “ยืดหยุ่น” ในการเลือกสินค้าแต่ยังอยู่ในกลุ่มเมนูเดียวกัน

จากผลการวิเคราะห์ข้อมูลธุรกรรมของลูกค้าด้วยอัลกอริทึม Apriori และ FP-Growth พบรูปแบบการซื้อสินค้าที่เกิดขึ้นร่วมกันอย่างมีนัยสำคัญ ซึ่งสามารถนำไปประยุกต์ใช้ในเชิงธุรกิจร้านกาแฟได้ดังนี้

1. จัดชุดเมนูส่งเสริมการขาย (Combo Set) จากความสัมพันธ์ระหว่างเครื่องดื่มและขนม เช่น ลาเต้กับขนมปังเนย หรือมัทฉะบันกับลาเต้ในฤดูใบไม้ร่วง สามารถจัดชุดเมนูยอดนิยมเพื่อเพิ่มยอดขายเฉลี่ยต่อใบเสร็จ (Cross-selling)

2. ออกแบบโปรโมชั่นตามฤดูกาล ตารางที่ 2 แสดงให้เห็นว่าในฤดูใบไม้ผลิ ลูกค้านิยมเครื่องดื่มรสหวาน (เช่น วานิลลา ลาเต้, ครีมน้ำตาล) สามารถออกเมนูพิเศษประจำฤดู (Seasonal Menu) เช่น “Spring Latte Set”

3. จัดวางสินค้าภายในร้าน สินค้าที่มักถูกซื้อร่วมกันควรวางใกล้กันเพื่อกระตุ้นการซื้อ เช่น วางขนมปังใกล้จุดส่งลาเต้

4. ปรับกลยุทธ์การแนะนำสินค้า (Recommendation) นำกฎที่ได้มาใช้ในการแนะนำสินค้าอัตโนมัติ (เช่น บนแอปหรือจอ POS) หากลูกค้าสั่งลาเต้ ระบบสามารถแนะนำขนมที่ซื้อร่วมกันบ่อย

5. วางแผนสต็อกสินค้าอย่างมีประสิทธิภาพ รู้ว่าลูกค้าชอบซื้อสินค้าคู่กัน ทำให้สามารถประมาณการสต็อกได้แม่นยำขึ้น เช่น เมื่อมีการสั่งลาเต้เพิ่ม ควรเตรียมขนมปังเนยไว้มากขึ้นด้วย

ตารางที่ 4 เปรียบเทียบความแตกต่างระหว่างผลลัพธ์ของ Association Rules (Apriori) และ Association Rules (FP-Growth)

คุณสมบัติ	Apriori อัลกอริทึม	FP-Growth อัลกอริทึม
จำนวนกฎที่สร้างขึ้น	578,180 กฎ	578,180 กฎ
กระบวนการสร้างกฎ	ใช้การสร้าง Candidate และสแกนฐานข้อมูลซ้ำๆ	ใช้โครงสร้าง FP-Tree เพื่อลดจำนวนการสแกนฐานข้อมูล
จำนวนครั้งที่สแกนฐานข้อมูล	หลายครั้ง	เพียง 2 ครั้ง
การใช้หน่วยความจำ	ใช้หน่วยความจำสูงสำหรับ Candidate Sets	ใช้หน่วยความจำลดลงเนื่องจาก FP-Tree
ประสิทธิภาพ	ใช้เวลามากกว่าในฐานข้อมูลขนาดใหญ่	ประมวลผลได้เร็วกว่าในฐานข้อมูลขนาดใหญ่
ความซับซ้อน	มีความซับซ้อนสูง เนื่องจากต้องสร้างและกรอง Candidate	ความซับซ้อนน้อยกว่าเพราะใช้ FP-Tree
ตัวอย่าง Confidence	Confidence สูงสุด ~ 86.7%	Confidence สูงสุด ~ 86.7%
ตัวอย่าง Lift	สูงสุด ~ 1.056	สูงสุด ~ 1.064
การใช้งานเหมาะสม	เหมาะสำหรับชุดข้อมูลขนาดเล็กถึงปานกลาง	เหมาะสำหรับชุดข้อมูลขนาดใหญ่

จากตารางที่ 4 การวิเคราะห์และเปรียบเทียบคุณสมบัติของอัลกอริทึม Apriori และ FP-Growth พบว่าทั้งสองอัลกอริทึมจะสามารถสร้างกฎความสัมพันธ์ได้ในจำนวนที่เท่ากัน คือ 578,180 กฎ แต่กระบวนการทำงานและประสิทธิภาพของทั้งสองต่างกันอย่างมีนัยสำคัญ อัลกอริทึม Apriori อาศัยการสร้างชุดข้อมูลที่

เรียกว่า Candidate Itemsets และต้องสแกนฐานข้อมูลซ้ำหลายครั้งเพื่อกรองหาชุดรายการที่เกิดบ่อย ซึ่งทำให้มีความซับซ้อนในการประมวลผล และใช้ทรัพยากรหน่วยความจำมาก โดยเฉพาะเมื่อทำงานกับฐานข้อมูลขนาดใหญ่ อย่างไรก็ตาม Apriori เหมาะสมกับชุดข้อมูลขนาดเล็กถึงปานกลาง เนื่องจากสามารถให้ผลลัพธ์ที่แม่นยำและกระบวนการทำงานเข้าใจง่ายในเชิงพื้นฐาน ในทางตรงกันข้าม อัลกอริทึม FP-Growth ใช้โครงสร้างข้อมูลแบบ FP-Tree เพื่อลดจำนวนครั้งในการสแกนฐานข้อมูล ซึ่งส่งผลให้มีประสิทธิภาพในการประมวลผลที่ดีกว่า FP-Growth สามารถประมวลผลได้โดยใช้การสแกนฐานข้อมูลเพียง 2 ครั้ง และไม่ต้องสร้าง Candidate Itemsets จึงช่วยลดการใช้หน่วยความจำและลดความซับซ้อนของกระบวนการได้อย่างมาก ทำให้เหมาะสำหรับการใช้งานกับฐานข้อมูลขนาดใหญ่ที่มีรายการข้อมูลจำนวนมากและซับซ้อน

สำหรับค่าตัวชี้วัดทางเทคนิค เช่น ค่า Confidence สูงสุดของทั้งสองอัลกอริทึมอยู่ในระดับเดียวกันที่ประมาณ 86.7% แสดงถึงความแม่นยำของกฎความสัมพันธ์ที่สร้างขึ้น ส่วนค่า Lift สูงสุดของ FP-Growth อยู่ที่ประมาณ 1.064 ซึ่งสูงกว่า Apriori เล็กน้อย (ประมาณ 1.056) แสดงให้เห็นว่า FP-Growth มีความสามารถในการระบุความสัมพันธ์เชิงบวกได้ชัดเจนกว่าเล็กน้อย โดยสรุป อัลกอริทึม FP-Growth มีความได้เปรียบในด้านประสิทธิภาพ ความเร็ว และความเหมาะสมสำหรับการใช้งานในบริบทของข้อมูลขนาดใหญ่ ขณะที่ Apriori ยังคงเหมาะสมสำหรับสถานการณ์ที่ต้องการความเข้าใจง่ายและข้อมูลไม่ซับซ้อนมากนัก

2. การเปรียบเทียบประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth ในการวิเคราะห์ธุรกรรม การซื้อสินค้าของผู้บริโภคในร้านกาแฟ

ตารางที่ 5 เปรียบเทียบประสิทธิภาพระหว่าง Apriori และ FP-Growth

เกณฑ์การเปรียบเทียบ	Apriori	FP-Growth
Processing Time	0.00 วินาที	0.01 วินาที
Iterations	4	4
Memory Usage	30.33 KB	30.34 KB

จากตารางที่ 5 แสดงการเปรียบเทียบเชิงประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth โดยพิจารณาจาก 3 เกณฑ์หลัก ได้แก่ เวลาในการประมวลผล (Processing Time), จำนวนรอบการทำงาน (Iterations) และการใช้หน่วยความจำ (Memory Usage) ผลการวิเคราะห์พบว่า ทั้งสองอัลกอริทึมใช้จำนวนรอบการประมวลผลเท่ากัน คือ 4 รอบ และมีการใช้หน่วยความจำที่ใกล้เคียงกันมาก โดย Apriori ใช้ 30.33 KB ขณะที่ FP-Growth ใช้ 30.34 KB ซึ่งแตกต่างกันเพียงเล็กน้อยเท่านั้น ในส่วนของเวลาในการประมวลผล Apriori ใช้เวลา 0.00 วินาที ส่วน FP-Growth ใช้เวลา 0.01 วินาที ซึ่งอาจสะท้อนว่าในกรณีของชุดข้อมูลขนาด

เล็ก Apriori สามารถประมวลผลได้รวดเร็วกว่าเล็กน้อย อย่างไรก็ตาม ความแตกต่างนี้ถือว่าเล็กน้อยมาก และอาจไม่มีนัยสำคัญในเชิงประสิทธิภาพโดยรวม เมื่อพิจารณาโดยภาพรวม ผลการเปรียบเทียบจากตารางนี้แสดงว่าทั้งสองอัลกอริทึมมีประสิทธิภาพใกล้เคียงกันเมื่อใช้กับข้อมูลขนาดเล็ก แต่ในเชิงทฤษฎีและการศึกษาก่อนหน้า FP-Growth มักจะมีข้อได้เปรียบในด้านความเร็วและการจัดการทรัพยากรเมื่อใช้กับฐานข้อมูลขนาดใหญ่

สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ศึกษากฎความสัมพันธ์ (Association Rules) ที่เกิดขึ้นจากข้อมูลธุรกรรมของลูกค้าในร้านกาแฟ และ 2) เปรียบเทียบประสิทธิภาพของอัลกอริทึม Apriori และ FP-Growth ในการประมวลผลข้อมูลดังกล่าว โดยใช้ข้อมูลธุรกรรมจริงจำนวน 1,000 รายการจากเว็บไซต์ Kaggle ผ่านการวิเคราะห์ด้วยภาษา Python และไลบรารี mlxtend

ผลการวิเคราะห์ตามวัตถุประสงค์ข้อที่ 1 พบว่าลูกค้ามีพฤติกรรมการซื้อสินค้าที่มีความสัมพันธ์กันชัดเจน เช่น การซื้อ “ลาเต้” ร่วมกับ “ขนมปังเนย” และการซื้อ “มัทฉะปั่น” ร่วมกับ “วานิลลาลาเต้” โดยกฎบางรายการมีค่า Confidence สูงกว่า 85% และค่า Conviction มากกว่า 1 เช่น 1.11 ซึ่งบ่งชี้ถึงความน่าเชื่อถือของรูปแบบการซื้อที่สามารถคาดการณ์ได้

สำหรับวัตถุประสงค์ข้อที่ 2 เมื่อเปรียบเทียบประสิทธิภาพ พบว่าอัลกอริทึม Apriori ใช้เวลา 0.00 วินาที และหน่วยความจำ 30.33 KB ส่วน FP-Growth ใช้เวลา 0.01 วินาที และหน่วยความจำ 30.34 KB แม้ค่าจะแตกต่างกันเพียงเล็กน้อย แต่ในเชิงโครงสร้าง FP-Growth ใช้แนวทาง FP-Tree ซึ่งสามารถลดจำนวนการสแกนข้อมูล ทำให้เหมาะสมกับชุดข้อมูลขนาดใหญ่หรือซับซ้อนมากกว่า

ผลการวิจัยนี้สามารถนำไปประยุกต์ใช้จริงในธุรกิจร้านกาแฟเพื่อเพิ่มประสิทธิภาพการขาย เช่น การออกแบบชุดโปรโมชั่นเมนูยอดนิยม เช่น ลาเต้ + ขนมปัง, การจัดวางสินค้าให้ตอบโจทย์พฤติกรรมการซื้อหรือการสร้างระบบแนะนำเมนูอัตโนมัติบนแอปพลิเคชัน เพื่อเพิ่มยอดขายและความพึงพอใจของลูกค้า

ข้อเสนอแนะ

1. ข้อเสนอในการนำผลการวิจัยไปใช้ ผลการวิเคราะห์กฎความสัมพันธ์ด้วยอัลกอริทึม Apriori และ FP-Growth สามารถนำไปประยุกต์ใช้ในธุรกิจค้าปลีกหรือร้านกาแฟได้อย่างมีประสิทธิภาพ โดยเฉพาะในด้านต่อไปนี้

1.1 การออกแบบโปรโมชั่นหรือชุดเมนูส่งเสริมการขาย โดยใช้ข้อมูลการซื้อซ้ำของสินค้าช่วยกำหนดรูปแบบชุดเมนูที่ลูกค้ามีแนวโน้มซื้อพร้อมกัน

1.2 การวางผังร้าน (Product Placement) เพื่อให้สินค้าที่มักถูกซื้อร่วมกันอยู่ใกล้กัน เพิ่มโอกาสในการซื้อซ้ำ

1.3 การพัฒนาแอปพลิเคชันหรือระบบแนะนำสินค้าอัตโนมัติ (Recommendation System) เพื่อแนะนำเมนูเพิ่มเติมให้กับลูกค้าแบบเรียลไทม์

1.4 การวางแผนการจัดสต็อกสินค้าให้มีความแม่นยำยิ่งขึ้น โดยใช้รูปแบบการซื้อซ้ำเป็นตัวช่วยในการคาดการณ์ความต้องการสินค้าแต่ละชนิด

2. ข้อเสนอแนะในการทำวิจัยครั้งต่อไป เพื่อให้ผลการวิจัยมีความสมบูรณ์และสามารถประยุกต์ใช้ได้หลากหลายมากยิ่งขึ้น ควรพิจารณาข้อเสนอแนะดังต่อไปนี้

2.1 การปรับแต่งพารามิเตอร์ในการวิเคราะห์ ควรทดลองปรับเปลี่ยนค่าพารามิเตอร์ เช่น ค่า Support และ Confidence ขั้นต่ำ เพื่อศึกษาว่า สามารถส่งผลกระทบต่อจำนวน และคุณภาพของกฎความสัมพันธ์อย่างไร และเพื่อหาเกณฑ์ที่เหมาะสมที่สุดสำหรับข้อมูลธุรกรรมในบริบทของร้านกาแฟ

2.2 การเพิ่มขนาดฐานข้อมูล ควรนำข้อมูลธุรกรรมจากหลายช่วงเวลาหรือจากร้านค้าหลายสาขา มาวิเคราะห์ร่วมกัน เพื่อทดสอบความเสถียรของกฎ และประเมินว่าวิธีการวิเคราะห์ยังคงมีประสิทธิภาพในกรณีที่ข้อมูลมีความหลากหลายหรือซับซ้อนมากขึ้น

2.3 การเปรียบเทียบกับอัลกอริทึมอื่น ๆ แนะนำให้เปรียบเทียบ FP-Growth และ Apriori กับอัลกอริทึมอื่น เช่น ECLAT หรืออัลกอริทึมที่ใช้ deep learning เพื่อวิเคราะห์ความสัมพันธ์ที่ซับซ้อนและมีหลายมิติ เพื่อให้เห็นภาพรวมของประสิทธิภาพและความเหมาะสมในแต่ละกรณี

2.4 การพัฒนาเป็นระบบแนะนำสินค้า (Recommendation System) สามารถต่อยอดผลการวิเคราะห์ไปสู่การพัฒนาโมเดลแนะนำเมนูอัตโนมัติ (Menu Recommendation) เพื่อประยุกต์ใช้จริงในร้านกาแฟ โดยเชื่อมต่อกับระบบ POS หรือแอปพลิเคชันสั่งอาหาร

กิตติกรรมประกาศ

ขอขอบคุณหน่วยงานสาขาวิชาธุรกิจดิจิทัล คณะดิจิทัล วิทยาลัยเทคโนโลยีสยาม ที่ให้การสนับสนุนเวลา และสถานที่สำหรับทำงานวิจัย

เอกสารอ้างอิง

Chusna, N. L., Herwanto, & Sari, I. P. (2021). Application of data mining for product purchase pattern analysis with frequent pattern growth (FP-Growth) algorithm on sales transaction data. *Journal of Science, Technology, and Engineering*, 1(1), 1–6.

- Hunyadi, D. (2011). **Performance comparison of Apriori and FP-Growth algorithms in generating association rules**. Proceedings of the European Computing Conference, 376–381.
- Istrat, V., & Lalić, N. (2017). Association rules as a decision making model in the textile industry. **Fibres & Textiles in Eastern Europe**, 25, 4(124), 8-14.
- Mustakim, M., & Team. (2018). Market basket analysis using Apriori and FP-Growth for analysis consumer expenditure patterns at Berkah Mart in Pekanbaru Riau. **Journal of Physics: Conference Series**, 1114(1), 012131, 1-9.
- Sun, X., Ngueilbaye, A., Luo, K., Cai, Y., Wu, D., & Huang, J. Z. (2024). A scalable and flexible basket analysis system for big transaction data in Spark. **Information Processing and Management**, 61(2), 103577, 1-22.