

บทที่ 5

สรุปและข้อเสนอแนะ

การวิเคราะห์ข้อมูลโครงการเรื่องการวิเคราะห์พฤติกรรมการณ์ซื้อของลูกค้าด้วย Machine Learning และ RFM Analysis ผู้วิเคราะห์ข้อมูลได้ดำเนินการ ศึกษาและวิเคราะห์ข้อมูลจาก ผศ. ดร ชัชฎาพร ปุกแก้ว ทางผู้วิเคราะห์ได้ทำการสรุปผลการ ทำโครงการ ข้อจำกัดของเว็บไซต์ ปัญหาและอุปสรรคต่าง ๆ ของการทำโครงการ และ ข้อเสนอแนะเพื่อพัฒนาโครงการ ดังนี้

5.1 สรุปผลการดำเนินงาน

จากการดำเนินโครงการเรื่อง การวิเคราะห์พฤติกรรมการณ์ซื้อของลูกค้าด้วย Machine Learning และ RFM Analysis พบว่าการจัดกลุ่มลูกค้าด้วยอัลกอริทึม K-Means Clustering (K=3) ร่วมกับเกณฑ์ RFM สามารถแบ่งลูกค้าออกเป็น 3 กลุ่มได้อย่างชัดเจน ได้แก่ กลุ่ม A (Can't Lose Them) ลูกค้าชั้นดีที่ซื้อบ่อยและมียอดใช้จ่ายสูง กลุ่ม B (At Risk) ลูกค้าที่มีความเสี่ยงจะสูญเสีย และกลุ่ม C (Hibernating) ลูกค้าที่หลับไหลซึ่งมีจำนวนมากที่สุดในระบบ เมื่อนำข้อมูลไปสร้างแบบจำลองต้นไม้ตัดสินใจ (Decision Tree) เพื่อจำแนกประเภทลูกค้า พบว่า โมเดลมีความแม่นยำ (Accuracy) สูงถึงร้อยละ 100 เนื่องจากชุดข้อมูลมีเงื่อนไขและกฎเกณฑ์ที่ตายตัว นอกจากนี้ อัลกอริทึม FP-Growth ยังค้นพบกฎความสัมพันธ์ที่แข็งแกร่ง 9 กฎ โดยทุกกฎมีค่าความเชื่อมั่น (Confidence) ร้อยละ 100 และค่าความสัมพันธ์ (Lift) มากกว่า 1 ซึ่งช่วยยืนยันความถูกต้องของเกณฑ์การจัดกลุ่มได้อย่างเป็นรูปธรรม

ในส่วนของการประยุกต์ใช้งาน คณะผู้จัดทำได้พัฒนาระบบเว็บไซต์ (Web Application) ที่สามารถจัดการข้อมูลการทำธุรกรรมของลูกค้าได้อย่างครบถ้วน โดยระบบสามารถคำนวณรหัส RFM พร้อมตีป้ายสถานะกลุ่มลูกค้าแยกตามสีให้อัตโนมัติ รวมทั้งมีหน้าแดชบอร์ดสรุปผลเชิงสถิติที่แสดงภาพรวมของข้อมูลได้อย่างเข้าใจง่าย ซึ่งช่วยให้องค์กรสามารถนำผลลัพธ์ไปประยุกต์ใช้ในการบริหารความสัมพันธ์ลูกค้า (CRM) และประกอบการตัดสินใจทางธุรกิจได้อย่างมีประสิทธิภาพ

5.2 ปัญหาและอุปสรรคในการดำเนินงาน

1. คุณภาพและรูปแบบของข้อมูลต้นฉบับในขั้นตอนการเตรียมข้อมูล (Data Preparation) พบว่าข้อมูลวันที่ของการทำธุรกรรมในไฟล์ต้นฉบับมีรูปแบบที่ไม่ตรงกัน
2. ข้อจำกัดของสัญลักษณ์ในข้อมูลในระหว่างการพัฒนาเว็บไซต์ พบข้อผิดพลาดในการบันทึกข้อมูล เช่น กลุ่ม "A (Can't Lose Them)" ลงฐานข้อมูล เนื่องจากมีเครื่องหมายอัฒภาคเดี่ยว (') ซ่อนอยู่ ทำให้ระบบฐานข้อมูลเกิดความสับสน

5.3 ปัญหาและอุปสรรคของโครงการ

1. ข้อมูลดิบมีความไม่สมบูรณ์ (Missing Data) จึงต้องมีการทำ Data Cleaning และ สร้างตัวแปรสังเคราะห์เพิ่มเติม
2. ความซับซ้อนของโมเดลบางประเภท เช่น Decision Tree ทำให้ผลลัพธ์ยากต่อการตีความสำหรับผู้ที่ไม่มีความรู้พื้นฐานด้านการวิเคราะห์ข้อมูล
3. โมเดลบางอันเป็นโมเดลใหม่ที่ไม่เคยศึกษามาก่อน เช่น RFM Analysis ต้องมีการศึกษาทำความเข้าใจใหม่ตั้งแต่ต้น

5.4 ข้อเสนอแนะ

การปรับปรุงแบบจำลองให้ทันเหตุการณ์ เนื่องจากพฤติกรรมของผู้บริโภคในตลาดอีคอมเมิร์ซมีการเปลี่ยนแปลงอยู่ตลอดเวลา ควรมีการออกแบบระบบให้รองรับการนำข้อมูลธุรกรรมชุดใหม่ ๆ เข้ามาสอนแบบจำลองอย่างสม่ำเสมอ