

ปิยธิดา รุจะศิริ 2552: การเปรียบเทียบเทคนิคการแบ่งกลุ่มข้อมูลสำหรับการวิเคราะห์
กลุ่ม ปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ อาจารย์ที่ปรึกษา
วิทยานิพนธ์หลัก: ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมทิ, Ph.D. 185 หน้า

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบความสามารถในการแบ่งกลุ่มของ
เทคนิคการแบ่งกลุ่ม 5 วิธี ได้แก่ วิธีการจัดกลุ่มแบบลำดับชั้น อัลกอริทึมจัดกลุ่มแบบเคมีน
วิธีจัดระบบด้วยตัวเองของโคโฮเนน อัลกอริทึมจัดกลุ่มแบบเคมคอยด์ และการใช้อัลกอริทึม
จัดกลุ่มแบบเคมคอยด์ร่วมกับวิธีวัดระยะห่างแบบไดนามิกไทม์วอร์ปปีง โดยทำการศึกษากับ
ข้อมูลจริงและข้อมูลจำลอง ซึ่งมีการแจกแจงแบบปกติหลายตัวแปร ในการจำลองข้อมูลใช้เทคนิค
มอนติคาร์โลสร้างข้อมูลจำนวน 1,000 ซ้ำ สำหรับแต่ละชุดข้อมูลซึ่งมีขนาด 25, 100 และ 300
ตัวอย่าง และกำหนดจำนวนกลุ่มข้อมูลขนาด 2-5, 2-8 และ 2-15 กลุ่มตามลำดับ กำหนดจำนวน
ตัวแปรที่ศึกษาเป็น 3, 5 และ 7 ตัวแปร โดยพิจารณาจากเกณฑ์ค่าความแตกต่างของข้อมูลภายใน
กลุ่ม (RMSSTD) ซึ่งถ้าค่า RMSSTD ต่ำ แสดงว่าเทคนิคการแบ่งกลุ่มนั้นเป็นวิธีที่ดี และพิจารณา
จากเกณฑ์ค่าความแตกต่างของข้อมูลระหว่างกลุ่ม (RS) ซึ่งถ้าค่า RS สูง แสดงว่าเทคนิคการ
แบ่งกลุ่มนั้นเป็นวิธีที่ดี ผลการวิจัยพบว่า ทั้งการศึกษการแบ่งกลุ่มจากข้อมูลจำลองและจาก
ข้อมูลจริงให้ผลสอดคล้องกัน กล่าวคือ อัลกอริทึมจัดกลุ่มแบบเคมีนให้ค่า RMSSTD ต่ำที่สุดและ
ค่า RS สูงที่สุดในเกือบทุกกรณี ซึ่งสรุปได้ว่า อัลกอริทึมจัดกลุ่มแบบเคมีนให้ผลการวิเคราะห์
กลุ่มที่ดีที่สุด ยกเว้นกรณีที่แบ่งกลุ่มข้อมูลขนาด 25 ตัวอย่างเป็น 4 หรือ 5 กลุ่ม ซึ่งวิธีการจัดกลุ่ม
แบบลำดับชั้นด้วยวิธี Ward linkage ให้ผลการวิเคราะห์กลุ่มดีที่สุด นอกจากนี้เมื่อจำนวนกลุ่ม
เพิ่มขึ้นประสิทธิภาพในการแบ่งกลุ่มของทั้ง 5 เทคนิควิธี มีแนวโน้มเพิ่มขึ้น แต่จำนวนตัวแปร
และขนาดของชุดข้อมูลที่เพิ่มขึ้นไม่มีผลต่อประสิทธิภาพในการแบ่งกลุ่มของเทคนิคการแบ่งกลุ่ม
ทั้ง 5 วิธี