

จากรวบรวม กาญจนศุภวรรณ 2554: การแบ่งกลุ่มผู้เชี่ยวชาญในมหาวิทยาลัยของรัฐบาล โดยใช้เทคนิคการจัดกลุ่มแบบอัลกอริทึม 2 ขั้นตอนและการหาหน้าหนึ่งของคุณลักษณะ ปริญาวิทยาศาสตร์ มหาวิทยาลัย (วิทยาการคอมพิวเตอร์) สาขาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: รองศาสตราจารย์อนงค์นาถ ศรีวิหค, Ph.D. 104 หน้า

มหาวิทยาลัยเป็นองค์กรการศึกษาที่มีขนาดใหญ่ที่ประกอบด้วยผู้เชี่ยวชาญมากมายหลายสาขา ข้อมูลผู้เชี่ยวชาญมักถูกเก็บไว้โดยไม่ได้นำมาใช้ประโยชน์อย่างเต็มที่ การจะค้นหาผู้เชี่ยวชาญเป็นเรื่องยาก เนื่องจากในองค์กรขนาดใหญ่มีผู้เชี่ยวชาญจำนวนมาก ดังนั้นการศึกษานี้จึงนำเทคนิคการทำเหมืองข้อมูลมาช่วยในการจัดการข้อมูลผู้เชี่ยวชาญที่มีอยู่ (ซึ่งผู้เชี่ยวชาญต่างสาขากันอาจมีความเชี่ยวชาญที่เหมือนกันหรือคาบเกี่ยวกัน) โดยพิจารณาจากคำสำคัญซึ่งได้มาจากการกำหนดไว้แล้วจากผู้เชี่ยวชาญ ข้อมูลในเบื้องต้นได้จาก (1) ข้อมูลบุคลากร (2) ข้อมูลผลงานวิจัยและ (3) ข้อมูลวิทยานิพนธ์ เมื่อรวมข้อมูลดังกล่าวพบว่ามีจำนวนคุณลักษณะขนาดใหญ่ (จำนวน 971 คำสำคัญ) จึงจำเป็นต้องนำเทคนิคการคัดเลือกคุณลักษณะมาช่วยลดขนาดคุณลักษณะลง ซึ่งเทคนิคที่นำมาใช้ คือ การผสมผสานระหว่างวิธีการ CFS (Correlation-based Feature Subset Selection) และ วิธีการเชิงพันธุกรรม (Genetic Search) เทคนิคการจัดกลุ่มที่ใช้ คือ เทคนิคการจัดกลุ่มแบบอัลกอริทึม 2 ขั้นตอน ในขั้นตอนแรกหาจำนวนกลุ่มที่เหมาะสมโดยใช้อัลกอริทึม SOM (Self-Organizing Maps) และในขั้นตอนที่สองเป็นการจัดกลุ่มโดยใช้อัลกอริทึม K-Means และอัลกอริทึม Fuzzy C-Means ร่วมกับการให้น้ำหนักคุณลักษณะด้วยวิธีการ TF-IDF (Term Frequency-Inverse Document Frequency) วิธีการ Logarithm Weight และ วิธีการ Augmented Weight

จากผลการศึกษาพบว่าการลดขนาดคุณลักษณะให้ผลการจัดกลุ่มที่ดีขึ้น (เดิม 971 คุณลักษณะเหลือ 258 คุณลักษณะ) สำหรับตัววัดประสิทธิภาพการจัดกลุ่มที่นำมาใช้ ได้แก่ F-Statistic, R-Squared และ Silhouette จำนวนกลุ่มที่ดีที่สุด คือ 7 กลุ่มและจัดกลุ่มโดยใช้อัลกอริทึม K-Means ร่วมกับการ Logarithm Weight ให้ผลดีที่สุด เมื่อพิจารณาผลการจัดกลุ่มและคำสำคัญที่มีความถี่สูงสุดต่อพบว่ากลุ่มที่ (1) ผู้เชี่ยวชาญ 50 คน คำสำคัญ “เศรษฐศาสตร์” (2) ผู้เชี่ยวชาญ 249 คน คำสำคัญ “ป่า” (3) ผู้เชี่ยวชาญ 63 คน คำสำคัญ “Rice” (4) ผู้เชี่ยวชาญ 11 คน คำสำคัญ “Corn” (5) ผู้เชี่ยวชาญ 43 คน คำสำคัญ “สอน” (6) ผู้เชี่ยวชาญ 2,671 คน เป็นกลุ่มใหญ่ มีคำสำคัญที่ปะปนกันสูง (7) ผู้เชี่ยวชาญ 107 คน คำสำคัญ “ไฟฟ้า” สำหรับในกลุ่มที่ 6 มีผู้เชี่ยวชาญปะปนมาก จึงนำกลุ่มนี้มาจัดกลุ่มต่อโดยทดลองกับอัลกอริทึม K-Means พบว่าผลการจัดกลุ่มรอบนี้ให้แนวโน้มที่ไม่ดีนัก จึงหยุดการจัดกลุ่มไว้เท่านี้ จากการศึกษาครั้งนี้จะเป็นแนวทางในการต่อยอดพัฒนาวิธีการจัดกลุ่มอื่นที่มีประสิทธิภาพมากขึ้นและสามารถนำมาพัฒนาเป็นระบบสารสนเทศในมหาวิทยาลัยได้