

เป็น large itemset หรือในทางกลับกันข้อมูลที่เคยเป็น large itemset อาจจะเป็นหรือไม่เป็น large itemset ในฐานข้อมูลที่ได้รับการปรับปรุงแล้วอีกต่อไป ซึ่งมีผลต่อการเปลี่ยนแปลงกฎความสัมพันธ์ทำให้ต้องหา large k-itemset ใหม่ ลักษณะปัญหาของการหาความสัมพันธ์ใหม่สามารถสรุปได้ดังนี้คือ

- จำนวนครั้งที่ใช้ในการสแกนฐานข้อมูลเดิมและข้อมูลใหม่ที่ต้องการปรับปรุงใหม่ ซึ่งใช้เวลามากในการค้นหาความสัมพันธ์ด้วยขั้นตอนการหาความสัมพันธ์ใหม่ทั้งหมด โดยไม่นำความรู้ที่ได้จากการทำคาด้าไมน์นิ่งก่อนที่จะทำการปรับปรุงมาใช้

- ปัญหาในการหา large k-itemset ที่เกิดในฐานข้อมูลที่ปรับปรุง (L_{DB+db}) ซึ่งจะต้องมีค่า support ที่เกิดขึ้นมากกว่าค่า minimum support ที่กำหนด เพื่อนำมาสร้างกฎความสัมพันธ์ใหม่ที่เกิดขึ้น

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

งานวิจัยนี้มีวัตถุประสงค์เพื่อค้นคว้าและพัฒนาอัลกอริทึมที่ใช้ในการเพิ่มขยายการค้นหา กฎความสัมพันธ์เมื่อมีการปรับเปลี่ยน transaction ในฐานข้อมูลหรือการปรับเปลี่ยนค่า minimum support ซึ่งค่า minimum support คือ ค่าความน่าจะเป็นทางสถิติที่ใช้ในการวัดความสัมพันธ์ระหว่างข้อมูลที่มีโอกาสจะเกิดขึ้นด้วยกัน โดยมีผลต่อการเปลี่ยนแปลงกฎความสัมพันธ์ที่ได้ทำการค้นหาไว้แล้วจากการทำคาด้าไมน์นิ่งในฐานข้อมูลเดิม (original database) และนำแนวความคิดที่จะหลีกเลี่ยงหรือใช้จำนวนครั้งน้อยที่สุดในการสแกนฐานข้อมูลเดิมที่มีจำนวนของข้อมูลที่จัดเก็บไว้จำนวนมาก รวมถึงการจัดการฐานข้อมูลใหม่ (incremental database) ที่มีผลต่อการเปลี่ยนแปลงกฎความสัมพันธ์ใหม่ที่เกิดขึ้นภายหลังจากการปรับปรุงฐานข้อมูล (updated database)

เพื่อให้การค้นหาความสัมพันธ์ระหว่างข้อมูลครอบคลุมข้อมูลทั้งหมด ด้วยอัลกอริทึมที่มีประสิทธิภาพในการสร้าง candidate itemset ที่สามารถนำมาใช้ในการหา large k-itemset ได้ อย่างถูกต้องและรวดเร็ว โดยใช้ทรัพยากรที่มีอยู่ให้เกิดประโยชน์มากที่สุด

1.3 ทฤษฎีหรือแนวคิดที่ใช้ในการวิจัย

งานวิจัยนี้ได้นำทฤษฎีและแนวความคิดต่างๆมาใช้ คือการประยุกต์ใช้ในการหาความสัมพันธ์ โดยมีทฤษฎีและแนวคิดที่สามารถนำมาประยุกต์ใช้ได้ดังนี้คือ

1. Knowledge discovery in database
2. Association rule
3. Incremental association rule

1.4 ขอบเขตของการวิจัย

งานวิจัยที่นำเสนอนี้เป็นโมเดลที่ประยุกต์ใช้กับการเพิ่มขยายการค้นหาความสัมพันธ์เมื่อฐานข้อมูลถูกปรับปรุงให้สอดคล้องกับข้อมูลที่ต้องการหาความสัมพันธ์ได้

1.5 ขั้นตอนของการศึกษา

ในขั้นตอนของการศึกษานี้ ได้แสดงลำดับการทำงานตั้งแต่เริ่มต้นจนถึงสิ้นสุดการทำงาน ดังรายละเอียดต่อไปนี้

- 1.5.1 ศึกษาทฤษฎีและงานวิจัยจากเอกสารบทความต่าง ๆ ที่เกี่ยวข้องกับการทำงานวิจัย
- 1.5.2 กำหนดหัวข้อ วัตถุประสงค์ และขอบเขตการทำงานวิจัย
- 1.5.3 วิเคราะห์อัลกอริทึมและออกแบบโมเดลใหม่
- 1.5.4 พัฒนาโปรแกรม โดยใช้ซอฟต์แวร์ MATLAB พร้อมทั้งแก้ไขข้อผิดพลาดและทดสอบการทำงานของโมเดลกับข้อมูลที่กำหนดขึ้นเอง
- 1.5.5 เตรียมข้อมูลที่ใช้งานจริง เพื่อนำมาทดสอบการทำงานของโมเดล โดยในที่นี้จะใช้ synthetic dataset ที่ได้มีการคิดค้นโดย Agrawal et al [2] ด้วยหลักการสร้าง dataset จากหลักการทางสถิติที่เลียนแบบมาจากลักษณะการซื้อขายสินค้าจริง และได้มีการนำมาใช้ในการทดสอบอย่างแพร่หลายในงานวิจัยต่างๆ
- 1.5.6 รวบรวมผลการทดลองจากการทำงานของโมเดล
- 1.5.7 วิเคราะห์และสรุปผลการทดลอง

การค้นหาความสัมพันธ์เป็นงานวิจัยหนึ่งที่ได้รับคามนิยมนเพื่อค้นหาความรู้หรือ information ที่ซ่อนอยู่ในฐานข้อมูล และนำความรู้ที่ได้มาแสดงอยู่ในรูปของความสัมพันธ์เพื่อใช้ในงานด้านต่างๆ เช่น การบริหาร, การตัดสินใจ และการวางแผน เป็นต้น ซึ่งการค้นหาความสัมพันธ์มีการทำวิจัยในหลายรูปแบบแตกต่างกันตามลักษณะของข้อมูล

ในบทที่ 2 จะกล่าวถึง ทฤษฎีและงานวิจัยต่างๆ ที่เกี่ยวข้องกับการทำงานวิจัยที่นำเสนอ