

การขุดค้นชุดข้อมูลความถี่สูง (Frequent Itemset Mining) เป็นขั้นตอนแรกในการสืบค้นกฎความสัมพันธ์ (Association Rule Mining) ซึ่งเป็นแนวทางหนึ่งสำหรับในการทำเหมืองข้อมูล (Data Mining) ในวิทยานิพนธ์เล่มนี้ นำเสนอเทคนิคใหม่สำหรับในการขุดค้นชุดข้อมูลความถี่สูง โดยการศึกษาเทคนิคทรีโปรเจกชัน (Tree Projection) เดิมที่มีอยู่แล้วทำการปรับปรุงให้มีประสิทธิภาพมากยิ่งขึ้น กลายเป็นทรีโปรเจกชันแบบปรับปรุง (Modified Tree Projection) จากนั้นทำการทดสอบเปรียบเทียบประสิทธิภาพ กับเทคนิค ทรีโปรเจกชันเดิมที่มีอยู่โดยใช้ข้อมูลที่เป็น Public Domain Datasets

ผลจากการทดสอบประสิทธิภาพสรุปได้ว่า เทคนิคใหม่ที่ได้มีประสิทธิภาพมากกว่าเทคนิคเดิม ซึ่งได้มากถึง 39% ในด้านความเร็ว และลดการเขียนข้อมูลทรานแซกชันลงได้มากถึง 60% กับข้อมูลที่มีลักษณะของข้อมูลที่หนาแน่น (Dense Datasets) หรือค่าสนับสนุนขั้นต่ำ (Minimum Support) ที่ต่ำๆ ทำให้ลดระยะเวลาในการประมวลผล และประหยัดเนื้อที่หน่วยความจำที่ใช้ในการประมวลผลอีกด้วย

Frequent Itemset Mining is the first step of the procedure of making Association Rule Mining, which is applied to create Data Mining in this thesis. A new technique of Frequent Itemset Mining is introduced in this thesis by the way in which an existing algorithm of Tree Projection is adapted more efficient, and subsequently, the acquired modified Tree Projection are compared to the previous existent one with an efficiency experiment on the basis of Public Domain Datasets.

According to the consequence of the conducted experiment, the new technique is much more efficient than the previous existent one in both aspects of processing speed and reduction of the amount of transaction data needed to be written. The processing speed of the new technique is thirty tree percent faster than that of the previous existent one and the amount of transaction data needed to be written of new technique is sixty percent less than that of the previous existent one. With denser datasets or minimum supporting value due to new technique, the processing time decreases and the memory, required for data processing, can be saved.

Keywords: Mining / Data Mining / Association Rule / Association Rule Mining / Frequent Itemset / Frequent Itemset Mining / Algorithm / Tree / Tree Projection / Knowledge Discovery / KDD / Database / Public Domain Dataset / Real Dataset / Synthetic Dataset / TP / MTP / Modified / Modified Tree Projection