200638

การสร้างหลักเกณฑ์เชื่อมโยงเป็นเทคนิคหนึ่งที่สำคัญในการค้นความรู้จากข้อมูล ซึ่งถูกใช้วิเคราะห์ข้อมูลมาเก็ตบาสเก็ต วัตถุประสงค์หลักคือการดึงหลักเกณฑ์ที่แสดงความสัมพันธ์ที่เชื่อมโยงระหว่างสินค้าภายในการซื้อหนึ่งใบเสร็จ วิธีการแบบตรงไปตรงมาที่สร้างหลักเกณฑ์ที่เป็นไปได้ทั้งหมดเป็นเรื่องที่ทำได้ยากเนื่องจากปัญหาการระเบิดออกของความเป็นไปได้ของหลักเกณฑ์ ดังนั้นนักวิจัยจึงใช้ตัววัดความน่าสนใจเพื่อลดปริมาณการพิจารณาหลักเกณฑ์ที่ไม่น่าสนใจตามค่าสนับสนุนต่ำสุด และค่าความเชื่อมั่นต่ำสุด ในที่นี้ค่าสนับสนุนของหลักเกณฑ์แสดงถึงการนำหลักเกณฑ์ไปใช้ได้กับข้อมูลเมื่อเทียบกับข้อมูลทั้งหมด โดยคำนวณสัดส่วนของข้อมูลที่มีสินค้าที่ปรากฏในหลักเกณฑ์กับปริมาณข้อมูลการซื้อทั้งหมด ในอีกมุมมองหนึ่งคือค่าความเชื่อมั่นของหลักเกณฑ์แสดงถึงความน่าเชื่อถือของหลักเกณฑ์ ในงานวิจัยนี้นำเสนอตัววัดความน่าสนใจอีกหนึ่งตัวคือค่าสนับสนุนแบบอ่อน จากพื้นฐานทางตรรกศาสตร์จริงหรือเท็จ ข้อมูลที่ขัดแย้งกับหลักเกณฑ์เชื่อมโยงจะเหมือนกับข้อมูลที่ขัดแย้งกับประพจน์ถ้า-แล้ว ดังนั้นเรานิยามค่าสนับสนุนแบบอ่อนคือค่าที่พิจารณาเฉพาะระเบียนที่ไม่ขัดแย้งกับหลักเกณฑ์เปรียบเทียบกับระเบียนทั้งหมด สภาพไวถูกใช้ในการเปรียบเทียบหลักเกณฑ์ที่ได้จากค่าสนับสนุนแบบอ่อนบวกค่าความเชื่อมั่นกับค่าสนับสนุนดังเดิมบวกค่าความเชื่อมั่น การเลือกใช้สภาพไวก็เนื่องมาจากกลุ่มของหลักเกณฑ์ที่ถูกสร้างขึ้นจากทั้งสองรูปแบบมีความแตกต่างกันมาก ผลลัพธ์ที่ได้จากการเปรียบเทียบสภาพไว พบว่ากลุ่มของหลักเกณฑ์ที่เกิดขึ้นจากค่าสนับสนุนแบบอ่อนและค่าความเชื่อมั่นให้ค่าที่ดีกว่าค่าที่ได้จากกลุ่มของหลักเกณฑ์ที่เกิดจากค่าสนับสนุนและค่าความเชื่อมั่น

Association rule mining is one of a knowledge discovery methodology that has been used to analyze the market basket data. The aim of this methodology is to extract rules that address a strong association among purchased items appearing within transactional data. The naïve method of searching for all possibilities are prohibitive because of the combinatorial explosions of rules. Therefore, researchers use two interesting measures to narrow the search based on the minimum support and minimum confidence. The support of the rule exhibits the applicable of the rule with respect to the whole transaction by computing the ratio of transactions containing related items versus the total number of transactions. The confidence, on the other hand, demonstrates the reliability of the rule. This thesis suggests another numerical interesting measure called the weak support. Based on the true-false logic, the contradiction of association rule is the same as the contradiction of if-then rules and hence we define the weak support measure as the non-contradictory of the rule from the whole transactions. The sensitivity is used to compare rules from the weak support together with confidence against rules from the original support with confidence. The sensitivity is selected due to distinct rules and unequal number of generated rules. The results from the sensitivity comparison between the group of rules from weak support and confidence show better value than the sensitivity from the group of rules from support and confidence.