กระบวนการทำเหมืองข้อมูลประกอบด้วยเจ็ดขั้นตอนหลักคือ (๑) การรวบรวมข้อมูล, (๒) การแปลงรูปแบบข้อมูล, (๓) การปรับปรุงข้อมูล, (๔) การคัดเลือกข้อมูล, (๕) การค้นหาเพื่อคัดแยกแพทเทิร์นของข้อมูล หรือเรียกว่าการทำเหมืองความรู้, (๖) การประเมินคุณภาพของแพทเทิร์น, และ (๗) การนำเสนอความรู้ ขั้นตอนที่ ๑ ถึง ๔ จัดเป็นขั้นตอนก่อนการทำเหมืองข้อมูล ในขณะที่ขั้นตอนที่ ๖ และ ๗ อาจจัดได้ว่าเป็นขั้นตอนหลังการทำเหมืองข้อมูล ดังนั้นทั้งเจ็ดขั้นตอนสามารถจัดรวมเป็นขั้นก่อนการทำเหมืองข้อมูล ขณะทำเหมืองข้อมูล และหลังการทำเหมืองข้อมูล งานวิจัยนี้เน้นการศึกษาที่ขั้นหลังการทำเหมืองข้อมูล ระบบทำเหมืองข้อมูลส่วนมากจะสิ้นสุดกระบวนการที่ขั้นตอนนำเสนอความรู้ แต่ในงานวิจัยนี้ได้นำเสนอขั้นตอนต่อจากนั้น คือ ขั้นตอนการใช้ประโยชน์จากความรู้ที่ได้จากการทำเหมืองข้อมูล งานวิจัยนี้แสดงวิธีการนำความรู้ในรูปแบบของกฎการจำแนกมาใช้ประโยชน์ โดยความรู้ในรูปแบบของกฎการจำแนกจะถูกนำมาประเมินคุณภาพด้วยเกณฑ์ความครอบคลุมข้อมูล กฎที่ครอบคลุมข้อมูลได้มากจะได้รับการคัดเลือกมาใช้ในระบบผู้เชี่ยวชาญ การทดสอบความสามารถของระบบผู้เชี่ยวชาญที่ใช้การสร้างฐานความรู้แบบอัตโนมัตินี้ ใช้วิธีวัดความถูกต้องของคำแนะนำที่ให้โดยระบบผู้เชี่ยวชาญ จากนั้นเปรียบเทียบกับความถูกต้องของการทำนายโดยโปรแกรมการทำเหมืองข้อมูลแบบจำแนก รวมสามโปรแกรมได้แก่ โปรแกรมที่ใช้หลักการสร้างต้นไม้ตัดสินใจ (ID3) โปรแกรมสร้างกฎการจำแนก (PRISM) และโปรแกรมข่ายงานประสาท (Neural network) ผลการทดสอบยืนยันความถูกต้องในการให้คำแนะนำของระบบผู้เชี่ยวชาญ

224290

The process of data mining comprises of seven major steps: (1) data integration, (2) data transformation, (3) data cleaning, (4) data selection, (5) pattern extraction or knowledge mining, (6) pattern evaluation, and (7) knowledge presentation. Steps 1 to 4 are pre-data mining, whereas steps 6 and 7 may be viewed as post-data mining. Therefore, the seven major steps can be grouped into pre-data mining, mining, and post-data mining. This research studies the post-data mining processing. Most data mining systems finish their processing at the knowledge presentation step. Our work further the post-data mining processing to the step of knowledge deployment. This research illustrates the knowledge deployment step in which its input is the induced knowledge, in the formalism of classification rules. These rules are evaluated and filtered on the basis of coverage measurement. High coverage rules are transformed into decision rules to be used by the inference engine of the expert system. The accuracy of recommendation given by the expert system is evaluated and compared to other three classification systems, i.e. decision-tree induction (ID3), rule induction (PRISM), and neural network. The experimental results confirm the high accuracy of our expert system and the induced knowledge base.