

องค์กรสมัยใหม่จะมีการสร้างข้อมูลอิเล็กทรอนิกส์ปริมาณมหาศาล เก็บบันทึกไว้ในฐานข้อมูลของหน่วยงาน ข้อมูลเหล่านี้เป็นวัตถุดิบที่มีค่าสำหรับกระบวนการอัตโนมัติที่เรียกว่าการค้นหาคำถามจากฐานข้อมูลหรือการทำเหมืองข้อมูล มีวัตถุประสงค์เพื่อสำรวจหาความรู้ที่มีประโยชน์ในการสนับสนุนการตัดสินใจระดับสูง ความรู้ที่สำรวจได้อาจจะเป็นตัวแทนของข้อมูลที่แสดงในรูปแบบสรุป ความสัมพันธ์ระหว่างข้อมูลที่แสดงในลักษณะกฎ หรือตัวแทนของกลุ่มข้อมูลที่แสดงด้วยค่ากลางของกลุ่ม ในหลายทศวรรษที่ผ่านมาได้มีความพยายามประยุกต์ใช้เทคโนโลยีเหมืองข้อมูลและการเรียนรู้ของเครื่องจักร เพื่อสังเคราะห์ความรู้จากข้อมูลที่มีอยู่ด้วยวิธีการ โปรแกรมเชิงคำสั่ง และเชิงวัตถุ ในงานวิจัยนี้ได้เสนอกรอบแนวคิดเชิงตรรกะเพื่อการพัฒนากระบวนการเหมืองข้อมูลที่ทำหน้าที่จำแนกประเภทของข้อมูล การพัฒนาระบบใช้วิธีการ โปรแกรมเชิงประกาศที่มีพื้นฐานจากตรรกศาสตร์ เนื่องจากการเป็นโปรแกรมระดับสูงที่เขียนคำสั่งสั้น ทำให้สามารถลดเวลาการพัฒนาโปรแกรมของโปรแกรมเมอร์ และการโปรแกรมในเชิงตรรกะมีความเหมาะสมสำหรับงานที่ต้องเกี่ยวข้องอย่างมากกับเรื่องของความรู้ ผู้วิจัยได้พัฒนาอัลกอริทึมที่สามารถสร้างความรู้ที่มีค่าความน่าจะเป็นกำกับ เทคนิคการเรียนรู้ใช้วิธีสร้างต้นไม้ตัดสินใจที่สังเคราะห์ขึ้นจากข้อมูล ความรู้ที่ได้จะถูกแปลงให้อยู่ในรูปแบบกฎการตัดสินใจที่มีค่าความน่าจะเป็นกำกับ รายงานการวิจัยนี้แสดงวิธีการพัฒนาระบบเหมืองข้อมูลด้วยภาษาโปรล็อก และจากผลการทดสอบโปรแกรมด้วยข้อมูลมาตรฐานแสดงให้เห็นถึงความเป็นไปได้ที่จะเชื่อมต่อผลลัพธ์ที่อยู่ในรูปแบบกฎการตัดสินใจที่มีค่าความน่าจะเป็นกำกับ ประกอบเข้ากับระบบฐานความรู้

Modern organizations normally generate huge amount of data in electronic form stored in databases. These data are a valuable resource for automatic discovering of useful knowledge, known as knowledge discovery in databases or data mining, to support high-level decisions. Discovered knowledge may be patterns of data represented in summarized form, relationships among data represented as rules, or representatives of data subgroups represented by mean values. During the past decades there has been an increasing interest in devising database and machine learning technologies to automatically induce knowledge from stored data using imperative and object-oriented programming styles. In this research, we propose a data mining system based on a logical framework. The proposed logic-based system performs a data classification task. Declarative programming based on logic can greatly reduce the burden of programmers as it is a very high-level programming scheme suitable for the development of knowledge intensive tasks. We devise algorithms to generate probabilistic knowledge from the induced decision tree and infer decision from the induced probabilistic rules. The implementation of the proposed algorithms is demonstrated via a Prolog programming language. Experimental results on several data domains emphasize the simple form of knowledge representation and the potential of incorporating learning results as probabilistic knowledge in the knowledge-base system.