

ชื่อวิทยานิพนธ์ การสกัดความรู้โดยใช้แผนที่การจัดกลุ่มเอง
ผู้เขียน นางสาวพุทธิพร ธนธรรมเมธี
สาขาวิชา วิทยาการคอมพิวเตอร์
ปีการศึกษา 2549

บทคัดย่อ

ในปัจจุบันข้อมูลที่เก็บอยู่ในฐานข้อมูลนั้นมีเพิ่มมากขึ้น ทำให้การนำข้อมูลในฐานข้อมูลมาใช้ประโยชน์เป็นไปได้ยาก วิทยานิพนธ์นี้ได้นำเสนอแบบจำลองการสกัดความรู้ซึ่งนำมาช่วยในการค้นหาความรู้จากฐานข้อมูลที่มีข้อมูลจำนวนมาก เพื่อให้ได้มาซึ่งความรู้ที่เป็นประโยชน์และสามารถเข้าใจลักษณะของข้อมูลได้ง่าย แบบจำลองการสกัดความรู้ที่ได้นำเสนอมี 2 แบบจำลอง คือ แบบจำลองการสกัดความรู้โดยใช้แผนที่การจัดกลุ่มเอง ความรู้ที่สกัดได้อยู่รูปกฎทั่วไป และแบบจำลองการสกัดความรู้จากแผนที่การจัดกลุ่มเองโดยใช้หลักเอ็นโทรปีค่าต่ำสุด ความรู้ที่สกัดได้อยู่ในรูปกฎภาษาธรรมชาติ โดยแบบจำลองทั้งสองนี้ได้นำหลักการของแผนที่การจัดกลุ่มเองมาใช้ ซึ่งแผนที่การจัดกลุ่มเองเป็นการจัดกลุ่มข้อมูลที่มีลักษณะเหมือนกันให้อยู่ในกลุ่มเดียวกันโดยไม่ต้องใช้ข้อมูลตัวอย่างในการสอน ทำการลดมิติข้อมูลเพื่อให้สามารถเข้าใจลักษณะของข้อมูลในภาพรวมได้ สำหรับแบบจำลองการสกัดความรู้โดยใช้แผนที่การจัดกลุ่มเองนั้นประกอบด้วย 4 ขั้นตอน คือ 1) กระบวนการแผนที่การจัดกลุ่มเอง 2) การวิเคราะห์ Component Plane และ Label 3) กระบวนการสกัดคุณลักษณะสำคัญ และ 4) การสร้างความรู้สำหรับแบบจำลองการสกัดความรู้จากแผนที่การจัดกลุ่มเองโดยใช้หลักเอ็นโทรปีค่าต่ำสุด ประกอบด้วย 5 ขั้นตอน คือ 1) กระบวนการแผนที่การจัดกลุ่มเอง 2) ขั้นตอนวิธีหลักเอ็นโทรปีค่าต่ำสุด 3) การสร้างฟิชชีเซต 4) การสร้างช่วงค่าภาษาธรรมชาติจากฟังก์ชันความเป็นสมาชิก และ 5) การสกัดความรู้โดยใช้ทฤษฎีกราฟเซต การทดลองของทั้งสองแบบจำลองใช้ฐานข้อมูลมาตรฐานจากประเทศสหรัฐอเมริกา ประกอบด้วยฐานข้อมูลดอกไม้ไอริส ฐานข้อมูลโรคมะเร็งเต้านม ฐานข้อมูลโรคหัวใจ และฐานข้อมูลโรคเบาหวาน ผลการทดลองพบว่าแบบจำลองการจัดกลุ่มเองโดยใช้แผนที่การจัดกลุ่มเองได้กฎทั่วไป มีค่าความถูกต้องสูง จำนวนกฎน้อย ส่วนแบบจำลองการสกัดความรู้จากแผนที่การจัดกลุ่มเองโดยใช้หลักเอ็นโทรปีค่าต่ำสุดได้กฎภาษาธรรมชาติที่เข้าใจได้ง่าย มีค่าความถูกต้องสูง และมีจำนวนกฎน้อยเช่นกัน

Thesis Title Knowledge Extraction Using Self-Organizing Map
Author Miss Putthiporn Thanathamthee
Major Program Computer Science
Academic Year 2006

ABSTRACT

At present, the increasing of tremendous data in database brings difficulty to user for employing data in the database. This thesis proposes knowledge extraction model for assisting the search of knowledge from the tremendous of data in the database. The knowledge extraction model composes of 2 models that are Knowledge Extraction Using Self-Organizing Map (KESOM) model, which generates crisp rules and Knowledge Extraction from Self-Organizing Map Using Minimization Entropy Principle Algorithm (KESOM_MEP) model, which generates linguistic rules. The Self-Organizing Map is the clustering technique or unsupervised learning technique. A key characteristic of the Self-Organizing Map is to reduce multi-dimensional input data into two-dimensional data that is suitable for visualization. The Knowledge Extraction Using Self-Organizing Map (KESOM) model composes of 4 steps, which are 1) Self-Organizing Map process, 2) component plane and label analysis, 3) feature extraction process, and 4) generated knowledge based. Knowledge Extraction from Self-Organizing Map Using Minimization Entropy Principle Algorithm (KESOM_MEP) model composed of 5 steps, which are 1) Self-Organizing Map process, 2) minimization entropy principle algorithm, 3) creating fuzzy set, 4) creating linguistic term using membership function, and 5) knowledge extraction using rough set. The experiments of both models uses benchmark data from USA, which are iris database, breast cancer database, heart disease and diabetes database. The experimental results indicate that the KESOM model will receive crisp rules with high accuracy and few numbers of rules. However, the KESOM_MEP model will receive linguistic rules that are easy to understand with high accuracy and few numbers of rules as well.