

ชื่อเรื่องวิทยานิพนธ์ การสกัดกัญสำหรับการจำแนกโดยใช้วิธีนิวโรฟัซซีแบบใหม่

ผู้เขียน นางสาวนริศรา เอี่ยมคณิตชาติ

ปริญญา วิศวกรรมศาสตรดุษฎีบัณฑิต (วิศวกรรมไฟฟ้า)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์

รศ. ดร. นิพนธ์ ธีรอำพน	ประธานกรรมการ
รศ. ดร. สุทธิชัย เปรมฤดีปรีชาชาญ	กรรมการ
ผศ. ดร. เสริมศักดิ์ เอื้อตรงจิตต์	กรรมการ

บทคัดย่อ

การคัดเลือกกลุ่มย่อยของคุณลักษณะที่ดี เพื่อใช้สำหรับทำนายการแยกประเภทได้อย่างมีประสิทธิภาพนั้นเป็นที่ต้องการอย่างมากสำหรับการทำงานกับชุดข้อมูลทุกประเภท เครื่องมือที่คุณสมบัติในการสร้างคุณลักษณะข้อมูลแบบข้อความ และสามารถเลือกเฉพาะคุณลักษณะข้อมูลที่สำคัญ ทั้งยังสามารถใช้ได้กับชุดข้อมูลทั่วไปไม่เฉพาะเจาะจง จึงเป็นที่ต้องการอย่างยิ่ง งานวิจัยชิ้นนี้เป็นการนำองค์ประกอบด้านการสร้างคุณลักษณะข้อมูลแบบข้อความที่สามารถปรับค่าได้ การเลือกคุณลักษณะที่สำคัญ และการแยกประเภทแบบการสร้างกฎ รวมกันเอาไว้ในโครงข่ายเดียว เซตฟัซซีมักใช้เป็นตัวจัดการกับข้อมูลที่มีขอบเขตข้อมูลไม่คงที่หรือยืดหยุ่นได้ เพื่อให้โครงข่ายสามารถปรับค่าสำหรับการสร้างคุณลักษณะข้อมูลแบบข้อความได้ จึงนำฟังก์ชันการเป็นสมาชิกแบบฟัซซีมาใช้ โดยฝังตัวในโครงข่ายแทนการใช้คุณลักษณะข้อมูลดั้งเดิม หลังจากจบขั้นตอนการสอนแล้ว คุณลักษณะข้อมูลที่สำคัญจะถูกเลือก จากนั้นโครงข่ายจะทำการแยกประเภททั้งแบบการคำนวณ และแบบการสร้างกฎ โครงสร้างของโครงข่ายถูกออกแบบเป็นลักษณะโครงข่ายประสาทเทียมแบบ 3 ชั้น โดยเป็นการออกแบบที่คำนึงถึงการหาข้อมูลที่สำคัญระหว่างขั้นตอนการสอน

โดยที่ชั้นซ่อนจะมีการปรับคุณลักษณะของข้อมูลฝังตัว เพื่อใช้สำหรับการสกัดกฎการแยกประเภท ได้มีการพิสูจน์และนำเสนอชุดของสมการใหม่สำหรับการปรับค่าเพื่อให้บรรลุวัตถุประสงค์ที่กล่าวมาข้างต้นในระหว่างขั้นตอนการสอน

ในการทดลองกับข้อมูล 3 5 ชุดซึ่งประกอบด้วยชุดข้อมูลมาตรฐาน ชุดข้อมูล ไมโครอาร์เรย์ และชุดข้อมูลสังเคราะห์นั้น ได้ผลการทำงานเป็นที่น่าพอใจ ทั้งผลจากการคำนวณโดยตรงและจากกฎการแยกประเภท ซึ่งเป็นผลการทดลองจากการทำครอสวาเลดิเคชัน 10 รอบ (10-fold cross validation) ผลการทดลองแสดงให้เห็นว่า ชุดคุณลักษณะข้อมูลที่ถูกเลือกโดยระบบเป็นคุณลักษณะที่สำคัญ และผลการทดลองยังแสดงให้เห็นว่าวิธีที่นำเสนอในงานวิจัยนี้มีประสิทธิภาพดี เมื่อเปรียบเทียบกับวิธีการในงานวิจัยอื่นอีกหลากหลายวิธี

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

Thesis Title	Rule Extraction for Classification Using a Novel Neuro-Fuzzy Method	
Author	Miss Narissara Eiamkanitchat	
Degree	Doctor of Philosophy (Electrical Engineering)	
Thesis Advisory Committee	Assoc. Prof. Dr. Nipon Theera-Umpon	Chairperson
	Assoc. Prof. Dr. Suttichai Premrudeepreechacharn	Member
	Asst. Prof. Dr. Sermsak Uatrongjit	Member

ABSTRACT

The efficient feature subset selection for predictive and accurate classification is highly desirable in data analysis for any datasets. The generalized classification tool that can create proper linguistic features and can select informative features is extremely required. Our research is the combination of embedded linguistic feature creation and tuning algorithm, feature selection and rule-based classification in one neural network mechanism. The adjustable linguistic features are embedded in the network structure. Fuzzy sets are used for modeling data with uncertainties or flexible boundaries. Fuzzy membership functions are embedded in the neural network for linguistic feature creation instead of using ordinary features. After the training process, the informative features are selected. The network performs the classification task either by the direct calculation or by the rule-based approach. The structure of the three-layer feedforward neural network is designed with the consideration of useful information during the training process. The hidden layer is embedded with the linguistic feature tuning and mechanism for rule extraction. A set of new update equations are derived to achieve the tasks during the training process.

In the experiment, 35 datasets are used including standard datasets, microarray datasets and synthetic datasets. Good results from both direct calculation and from logical rules are achieved using the 10-fold cross validation. The results demonstrate the importance of the linguistic features selected by the network. The results also show that the proposed method achieves better classification performance than the other previously proposed methods.