

ชื่อโครงการ	โปรแกรมจำแนกข้อมูลแบบหลายโมเดล Development of Ensemble Classifier
โดย	นายกานต์นิธิ แซ่ลี้ม นายศักดิ์ดา กรรัมย์
ชื่อปริญญา	วิทยาศาสตร์บัณฑิต
สาขาวิชา	วิทยาการคอมพิวเตอร์
อาจารย์ที่ปรึกษา	อาจารย์ศุภฤกษ์ ฉัตรธนาโชติ
ปีการศึกษา	2560

บทคัดย่อ

การจำแนกประเภทข้อมูลนั้นมีการพัฒนาเทคนิคในการจำแนกประเภทอย่างต่อเนื่อง โดยค่าความถูกต้องในการจำแนกประเภทข้อมูลถือเป็นค่าที่นำมาใช้ในการวัดประสิทธิภาพของตัวจำแนกประเภทอย่างแพร่หลายในแต่ละชุดข้อมูลที่มีความแตกต่างกันมีความเหมาะสมในการเลือกใช้ตัวจำแนกประเภทที่แตกต่างกันซึ่งเป็นความยากต่อผู้ใช้โดยทั่วไป งานวิจัยนี้ จึงนำเสนอแนวคิดในการพัฒนาโปรแกรมจำแนกประเภทข้อมูลแบบหลายโมเดล โดยคัดเลือกตัวจำแนกที่เป็นที่รู้จัก ห้าชนิด ได้แก่ นาอ์ฟเบย์, โครงข่ายประสาทเทียม, เคเนียร์สเนเบอร์, ซัพพอร์ตเวกเตอร์แมชชีน และ ต้นไม้ตัดสินใจมาใช้งาน ร่วมกันเพื่อให้โปรแกรมคัดเลือกโมเดลที่ดีที่สุดสำหรับชุดข้อมูลนั้น ๆ มาใช้งาน ผลการทดสอบกับชุดข้อมูล แอปชุด พบว่าใน แต่ละชุดข้อมูลจะมีตัวจำแนกประเภทที่เหมาะสมและให้ค่าความถูกต้องสูง กระจายตัวกันไป และตัวโปรแกรมได้เลือกใช้โมเดล ที่มีค่าความถูกต้องสูงที่สุดให้ผู้ใช้โดยอัตโนมัติ

คำสำคัญ: การจำแนกประเภทข้อมูล ออนซัมเบิล การเรียนรู้ของเครื่อง

Project	Development of Ensemble Classifier
Author	Prefix Kannithi Saelim Prefix Sukda Kokram
Major	Computer Science
Advisor	Academic Position, Supharoek Chattanachot
Academic Year	2020

Abstract

Data Classification have been developed for improve the accuracy value because it is used as a measurement of classification performance. Selected classifier from users might be not suitable for the dataset. In this work, we proposed the program that make the selected classifier for dataset task is easier than usual by convert the manual process to automatic selected classifier. The proposed program creates several model from Neural network, Decision Tree, Naïve Bayes, Support Vector Machine, and K-Nearest Neighbor prior selected the best one based on accuracy value. Finally, the test results are represented that matching of suitable classifier and dataset are selected by proposed program. Therefore, this program can be used by user easily without concerning about the suitable of classifier and dataset.

Keywords: data classification, ensemble, machine learning.