

222773

วิทยานิพนธ์นี้ได้นำเสนอวิธีการสร้างกลุ่มก้อนของตัวจำแนกประเภทกำหนดการพันธุกรรมสำหรับการจำแนกประเภทข้อมูลไมโครอาร์เรย์ ซึ่งเป็นข้อมูลที่มีจำนวนข้อมูลน้อย ขณะที่จำนวนคุณลักษณะมีจำนวนมาก ในการสร้างสมาชิกของกลุ่มก้อนนั้น จะมุ่งสร้างตัวจำแนกประเภทที่มีประสิทธิภาพในการจำแนกประเภทข้อมูลที่ดี ในขณะที่ตัวจำแนกประเภทแต่ละตัวจะต้องมีความแตกต่างจากสมาชิกตัวอื่น ๆ ในกลุ่มก้อน วิธีการที่นำเสนอจะใช้เทคนิคการจัดกลุ่มข้อมูลแบบ K-Means เพื่อจัดกลุ่มของคุณลักษณะของข้อมูลที่มีลักษณะคล้ายกันให้อยู่ในกลุ่มเดียวกัน และการเลือกคุณลักษณะแบบ SNR (Signal-to-Noise Ratio) โดยจะนำคุณลักษณะที่มีค่า SNR สูงที่สุดลำดับที่ i ของแต่ละกลุ่ม มาสร้างเป็นเซตย่อยของคุณลักษณะเพื่อใช้ในการสอน เพื่อสร้างตัวจำแนกประเภทกำหนดการพันธุกรรมตัวที่ i ซึ่งวิธีการนี้สามารถสร้างตัวจำแนกประเภทกำหนดการพันธุกรรมที่มีประสิทธิภาพที่ดี และมีความแตกต่างจากตัวจำแนกประเภทตัวอื่น ๆ เนื่องจากการใช้คุณลักษณะที่แตกต่างกัน ทำให้ประสิทธิภาพของกลุ่มก้อนดีขึ้นตามไปด้วย

222773

This thesis presents an algorithm for generating an ensemble of Genetic Programming classifiers for microarray data. The number of data is small and it has high dimensions. In order to construct an ensemble, each classifier must have high efficiency and at the same time it must be different from other classifiers. The proposed method uses K-Means clustering for grouping the features of data which are similar into the same group. The SNR (Signal-to-Noise Ratio) feature selection is used to select informative features. The feature with the i^{th} best SNR score in each group is selected to form a set of features. This feature set is used to train the i^{th} Genetic Programming classifier. The proposed method creates a good Genetic Programming classifier where each classifier is different from the others. They contain different set of features. As a result, the performance of the ensemble is improved.