

วิทยานิพนธ์นี้แสดงการใช้การโปรแกรมเชิงพันธุกรรม และข่ายงานระบบประสาท ในการจำแนกประเภทของผู้ป่วยโรคธาลัสซีเมีย ซึ่งต่างจากงานวิจัยก่อนหน้านี้ ที่สนใจปัญหาการจำแนกประเภทของผู้ป่วยโรคโลหิตจางโดยอาศัยเทคนิคเหตุผลอิงกฎ ปัญหาที่ทำการวิจัย นำตัวอย่างทดสอบมาจากคนปกติ ผู้ป่วยและพาหะโรคธาลัสซีเมีย ข้อมูลที่นำมาใช้ในการทดสอบประกอบด้วยคุณลักษณะเฉพาะของเม็ดเลือดแดงและเกร็ดเลือดที่ได้มาจากเลือดตัวอย่าง สำหรับใช้จำแนกประเภทของผู้ป่วยและพาหะโรคธาลัสซีเมียจากคนปกติ มัลติเลเยอร์เพอร์เซ็ปตรอนถูกนำมาใช้เป็นแบบจำลองในการจำแนกความสัมพันธ์ระหว่างข้อมูลเข้าและข้อมูลออกโดยนำขั้นตอนวิธีแพร่ย้อนหลังมาใช้ในกระบวนการเรียนรู้ ในอีกทางหนึ่งเทคนิคการโปรแกรมเชิงพันธุกรรมถูกนำมาสร้างต้นไม้สำหรับตัดสินใจอิงพหุนาม โดยที่จุดปลายถูกสร้างจากข้อมูลเข้า จากผลการทดลองพบว่าผลลัพธ์ของการโปรแกรมเชิงพันธุกรรมดีกว่าข่ายงานระบบประสาทที่มีชั้นซ่อนชั้นเดียว อย่างไรก็ตามเมื่อเปลี่ยนจำนวนชั้นซ่อนเป็นสองชั้น ประสิทธิภาพของข่ายงานระบบประสาทจะสูงกว่าการโปรแกรมเชิงพันธุกรรม นอกจากนี้โครงสร้างของต้นไม้สำหรับตัดสินใจแสดงให้เห็นว่าข้อมูลลักษณะเฉพาะของเกร็ดเลือดไม่มีผลต่อประสิทธิภาพในการจำแนก ด้วยเหตุนี้การลดจำนวนข้อมูลเข้าที่ต้องการสำหรับการจำแนกจะส่งผลให้สามารถปรับปรุงประสิทธิภาพในการจำแนกได้ดียิ่งขึ้น

This thesis presents the use of genetic programming (GP) and a neural network for thalassemic patient classification. In contrast to previous works which concentrate on anemic patient classification problems where a rule-based reasoning technique has been utilized, the interested problem only covers the test samples from normal subjects and that from different types of thalassemic patient and thalassemic gene carrier. The features, which are the characteristics of red blood cell and blood platelet extracted from the blood samples, are used as input to the classifiers while the target output is either the patient/carrier type or a flag indicating that the subject is normal. Multilayer perceptrons have been used to model the classifier input-output relationship where a back-propagation algorithm is used during the learning process. On the other hand, a genetic programming technique has been used to create polynomial-based decision trees where terminals are created from the input features. The results indicate that the GP-generated decision trees have higher classification accuracy in comparison to multilayer perceptrons with one hidden layer in most test cases. However, once the number of hidden layers in the multilayer perceptrons changes to two, the network performance is higher than that of the GP-generated decision trees. Nonetheless, the structure of the decision trees reveals that the features from the characteristics of blood platelet have no effects on the classification performance. This helps to reduce the required input features for the classification task and make further improvements on the classification performance possible.