

ชื่อเรื่องวิทยานิพนธ์ การจำแนกประเภทของโรคมะเร็ง จากข้อมูลไมโครอาร์เรย์ของ ดีเอ็นเอ ด้วยเทคนิคโครงข่ายประสาท

ผู้เขียน นายทงศักดิ์ เมืองฝั้น

ปริญญา วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.จิรยุทธ ไชยจารุณิช
บทคัดย่อ

โรคลูคีเมียเป็นโรคมะเร็งชนิดหนึ่ง ซึ่งเกิดจากความผิดปกติของเซลล์เม็ดเลือดแบ่งออกเป็น 2 ประเภทคือ AML (Acute Myeloid Leukemia) และ ALL (Acute Lymphoblastic Leukemia) ในการทดลองนี้ใช้ข้อมูลคนไข้โรคลูคีเมียจำนวน 38 คนซึ่งอยู่ในรูปของไมโครอาร์เรย์ดีเอ็นเอ ซึ่งเป็นข้อมูลที่แสดงถึงค่าการแสดงออกของยีนของคนไข้ เพื่อทำนายประเภทของโรคลูคีเมียจากระดับการแสดงออกของยีนที่มีลักษณะแตกต่างกัน

ในการศึกษานี้ได้นำเทคนิคแบกพรวดพราด เนิวรอลเน็ตเวิร์ก มาประยุกต์ใช้โดยมีวัตถุประสงค์เพื่อใช้ในการทำนายประเภทของโรคลูคีเมีย ซึ่งแบกพรวดพราด เนิวรอลเน็ตเวิร์ก เป็นเทคนิคที่สำคัญประเภทหนึ่งของนิวรอลเน็ตเวิร์ก และใช้กันอย่างแพร่หลาย โดยแบ่งออกเป็น 2 วิธีหลักคือ Batch Mode และ Stochastic Mode แต่ละวิธีประกอบด้วยพารามิเตอร์ต่างๆ ที่เกี่ยวข้องเช่น จำนวน Hidden Nodes (N_h), ค่า Convergence Criterion (θ) และค่า Convergence Rate (η) เป็นต้น งานวิจัยนี้ได้พยายามออกแบบโครงสร้าง และทดลองปรับค่าพารามิเตอร์ต่างๆ เพื่อให้ได้ค่าพารามิเตอร์ที่เหมาะสมที่สุดของแต่ละวิธีในการทำนายประเภทของโรคลูคีเมีย ผลจากการศึกษาพบว่า เทคนิคแบกพรวดพราด เนิวรอลเน็ตเวิร์กที่มีฟังก์ชันแบบ Conjugate Gradient Descent (CGD) โดยประกอบด้วยพารามิเตอร์ต่างๆ คือ N_h เท่ากับ 2 และ θ เท่ากับ 0.3 ให้ผลการทำนายที่ถูกต้องมากที่สุด โดยมีค่าเปอร์เซ็นต์ความถูกต้องเฉลี่ยคือ 98.42 เปอร์เซ็นต์

Thesis Title	Classification of Cancer using DNA Microarray Data with Neural Network Technique
Author	Mr. Thanongsak Muangfun
Degree	Master of Science (Computer Science)
Thesis Advisor	Assistant Professor Dr. Jeerayut Chaijaruwanich

ABSTRACT

Leukemia is a type of cancer of blood cells. There are two types of leukemia, i.e. AML (Acute Myloid Leukemia) and ALL (Acute Lymphoblastic Leukemia). The data used in this study are from the DNA microarray experiments of 38 leukemia patients. The differences of gene expression patterns of each patient group are analyzed and further used to predict the types of leukemia of new patients.

This study employs the backpropagation neural network to predict the types of acute leukemia. The backpropagation neural network is widely used and regarded as one of the most important neural network techniques. There are two main backpropagation methods, i.e. Batch Mode and Stochastic Mode. Each method has certain related parameters; such as the number of hidden nodes (N_h), Convergence Criterion value (θ) and Convergence Rate value (η). This study attempts to design the structure and parameters of each backpropagation method that best predicts the types of acute leukemia. The results show that the backpropagation neural network with the Conjugate Gradient Descent (CGD) function and parameters; $N_h = 2$ and $\theta = 0.3$ produces the best prediction results which is 98.42% in average.