DATA MINING OF MAGNETOCARDIOGRAMS FOR PREDICTION OF ISCHEMIC HEART DISEASE.

YOSAWIN KANGWANARIYAKUL 4837224 EGTI/M

M.Sc. (TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)

THESIS ADVISORY COMMITTEE: THANAKORN NAENNA, Ph.D. (ENGINEERING SCIENCE), CHANIN NANTASENAMART, Ph.D. (MEDICAL TECHNOLOGY), TANAWUT TANTIMONGCOLWAT, Ph.D. (MEDICAL TECHNOLOGY)

ABSTRACT

Ischemic Heart Disease (IHD) is a major cause of death. Early detection of IHD with an accurate and quick diagnosis is important for reducing the mortality rate. Magnetocardiogram (MCG) is a tool for detecting electro-physiological activity of the myocardium. MCG is a fully non-contact method, which avoids the problems of skin-electrode contact in the Electrocardiogram (ECG) method. However, the interpretation of MCG recordings is time-consuming and requires an expert. Therefore, we propose the use of machine learning for identification of IHD patients. In an automatic classification system, the back-propagation neural network (BPNN), the Bayesian neural network (BNN), the probabilistic neural network (PNN) and the support vector machine (SVM) were applied to develop classification models for identifying IHD patients.

MCG data are acquired by sequential measurement, above the torso, of the magnetic field emitted by the myocardium using a J-T interval of 125 cases. The training and validation data of 74 cases used 10-fold cross-validation methods to optimize support vector machine and neural network parameters. The predictive performance was obtained by prediction on testing data of 51 cases and performance comparison was decided on three parameters (accuracy, sensitivity, and specificity), and area under the receiver operating characteristic (ROC) curve.

The results demonstrated that the BPNN and BNN have the same highest accuracy of 78.43%, but the decision threshold of BPNN is -0.2774, and the area of ROC curve is 0.90596. While the decision threshold and the area of ROC curve of BNN are 0.0470 and 0.84953, respectively. Therefore, the BPNN model is the best classification model. The BNN is the best with 96.65% sensitivity, and the radial basis function (RBF) kernel SVM came out to be the best with 86.36% specificity.

KEY WORDS: ISCHEMIA/ MAGNETOCARDIOGRAPHY/ DATA MINING/ BACK-PROPAGATION NEURAL NETWORK/ BAYESIAN NEURAL NETWORK/ PROBABILISTIC NEURAL NETWORK/ SUPPORT VECTOR MACHINE

72 pages

การทำเหมืองข้อมูลคลื่นแม่เหล็กหัวใจเพื่อพยากรณ์โรคกล้ามเนื้อหัวใจขาดเลือด
DATA MINING OF MAGNETOCARDIOGRAMS FOR PREDICTION OF ISCHEMIC HEART DISEASE.

ยศวิน กังวานอริยะกุล 4837224 EGTI/M

วท.ม. (เทคโนโลยีการจัดการระบบสารสนเทศ)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์ : ธนกรณ์ แน่นหนา, Ph.D. (ENGINEERING SCIENCE), ชนินทร์ นันทเสนามาตร์, Ph.D. (MEDICAL TECHNOLOGY), ธนาวุฒิ ตันติมงคลวัฒน์, Ph.D. (MEDICAL TECHNOLOGY)

บทคัดย่อ

โรคกล้ามเนื้อหัวใจขาดเลือดเป็นหนึ่งในสาเหตุหลักที่ทำให้เกิดการเสียชีวิต ซึ่งถ้าผู้ป่วยได้รับการตรวจและรักษาตั้งแต่ในระยะเริ่มแรกของโรคก็จะช่วยลดอัตราการเสียชีวิตของผู้ป่วย การตรวจวัดคลื่นแม่เหล็กจากกล้ามเนื้อหัวใจมีข้อดีกว่าการตรวจวัดคลื่นไฟฟ้าหัวใจ กล่าวคือการตรวจวัดคลื่นแม่เหล็กจากกล้ามเนื้อหัวใจไม่จำเป็นต้องติดตั้งตัวรับสัญญาณบนร่างกายผู้ป่วย ซึ่งจะช่วยลดปัญหา skin-electrode contact แต่เนื่องจากการแปรผลคลื่นแม่เหล็กหัวใจจำเป็นต้องอาศัยผู้เชี่ยวชาญ และยังใช้เวลานาน ดังนั้นงานวิจัยนี้จึงได้พัฒนาระบบการเรียนรู้อัตโนมัติเพื่อพยากรณ์โรคกล้ามเนื้อหัวใจขาดเลือด โดยใช้เทคนิคการเรียนรู้แบบต่างๆ คือ back-propagation neural network (BPNN), Bayesian neural network (BNN), probabilistic neural network (PNN) และ support vector machine (SVM) ซึ่งข้อมูลคลื่นแม่เหล็กหัวใจในช่วงสัญญาณ J-T ของผู้ป่วยทั้ง 125 คน จะถูกแบ่งออกเป็นชุดข้อมูลในการเรียนรู้และชุดข้อมูลในการตรวจสอบ 74 คน เพื่อใช้ในการหาค่าพารามิเตอร์ที่เหมาะสมใน SVM และ neural network โดยใช้วิธี 10-fold cross-validation ส่วนข้อมูลของผู้ป่วยที่เหลือ 51 คน จะถูกใช้เป็นชุดข้อมูลทดสอบประสิทธิภาพของระบบที่ได้ โดยวัดจาก ค่าความแม่นยำ ค่าความไว ค่าความจำเพาะ และค่าพื้นที่ใต้กราฟ ROC

ผลการวิจัยพบว่า BPNN และ BNN มีค่าความแม่นยำ 78.43% เท่ากัน แต่ BPNN มีค่า decision threshold เท่ากับ -0.2774 และค่าพื้นที่ใต้กราฟ ROC เท่ากับ 0.90596 ในขณะที่ค่า decision threshold และ ค่าพื้นที่ใต้กราฟ ROC มีค่าเท่ากับ 0.0470 และ 0.84953 ตามลำดับ ดังนั้น BPNN จะมีประสิทธิภาพในการพยากรณ์ดีที่สุด ส่วน BNN จะมีค่าความไวสูงที่สุดคือ 96.65% และ RBF kernel SVM จะมีค่าความจำเพาะสูงสุดคือ 86.36%

72 หน้า