



**Sekolah Siswazah
(Graduate School)
Universiti Utara Malaysia**

**PERAKUAN KERJA KERTAS PROJEK
(Certification of Project Paper)**

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certify that)

MOHD KHALID BIN AWANG

calon untuk Ijazah _____ Sarjana Sains (Teknologi Maklumat)
(candidate for the degree of)

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project paper of the following title)

AN EVALUATION OF ARTIFICIAL NEURAL NETWORK

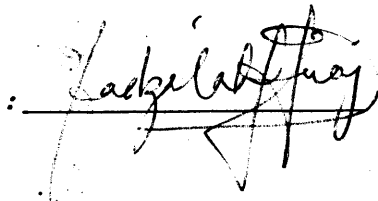
IN PREDICTING THE PRESENCE OF HEART DISEASE

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project paper)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan,
dan meliputi bidang ilmu dengan memuaskan.
*(that the project paper acceptable in form and content, and that a satisfactory
knowledge of the field is covered by the project paper).*

Nama Penyelia
(Name of Supervisor): Puan Fadzilah Siraj

Tandatangan
(Signature)

: 

Tarikh
(Date)

: 27 September 2000

AN EVALUATION OF ARTIFICIAL NEURAL NETWORK IN
PREDICTING THE PRESENCE OF HEART DISEASE

A project submitted to the Graduate School in partial
fulfilment of the requirements for the degree
Master of Science (Information Technology),
Universiti Utara Malaysia

by
Mohd Khalid Bin Awang

© Mohd Khalid Bin Awang, October 2000. All rights reserved

PERMISSION TO USE

In presenting this project in partial fulfilment of the requirements for a postgraduate degree from the Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for copying of this project in any manner, in whole or in part, for scholarly purposes may be granted by my supervisor(s) or, in their absence, by the Dean of the Graduate School. It is understood that any copying or publication or use of this project or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my project paper.

Requests for permission to copy or to make other use of material in this project in whole or in part should be addressed to:

**Dean of Graduate School
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman**

ABSTRACT (BAHASA MALAYSIA)

Projek **ini** bertujuan untuk menilai keupayaan model rangkaian neural dalam meramal penyakit jantung terutama kewujudan 'angina' pada pesakit yang telah dikenalpasti menghidap 'myocardial infarction'. **Ramalan tentang** kewujudan 'angina' adalah penting dalam menentukan kaedah rawatan yang bersesuaian untuk pesakit tersebut. Tambahan pula, diagnostik dan pengurusan 'angina' adalah penting untuk mencegah pengulangan serangan 'myocardial infarction'. Pembangunan aplikasi **ini** melibatkan tiga fasa utama. Fasa pertama ialah pembangunan Sistem Pengurusan Maklumat Myocardial Infarction (MIMIS) yang bertujuan untuk mengumpul dan **mengurus** maklumat pesakit. Kemudian diikuti oleh fasa kedua iaitu pembangunan Simulasi Rangkaian Neural (NNS) dengan menggunakan model perambatan balik untuk melatih dan menguji rangkaian. Fasa terakhir ialah pembangunan Sistem **Ramalan** (PS) untuk membuat **ramalan** terhadap pesakit yang **baru**. Kesemua sistem **tersebut** dibangunkan dengan menggunakan perisian Microsoft Visual Basics. Data untuk sesi latihan dan ujian diperolehi daripada Hospital Besar Alor Setar, Kedah. Model rangkaian neural terbaik yang dihasilkan mampu mencapai ketepatan **ramalan** sehingga 88.89 peratus. **Kajian ini** bukan sahaja membuktikan kemampuan rangkaian neural dalam mendiagnosis penyakit, malahan **juga** berjaya mengabungkan rangkaian neural dengan sistem pengurusan maklumat. Sebagai projek perintis, aplikasi **ini** dapat digunakan sebagai model untuk membangunkan sistem sokongan keputusan perubatan, terutama didalam mengdiagnosis penyakit jantung.

ABSTRACT (ENGLISH)

The purpose of this study is to evaluate the application of artificial neural network in predicting the presence of heart disease, particularly the angina in patients that already diagnosed with myocardial infarction. The prediction and detection of angina is important in determining the most appropriate form of treatment for these patients. Furthermore, diagnosis and management of angina is important since it can lead to the recurrent of myocardial infarction. The development of the application involves three main phases. The first phase is the development of Myocardial Infarction Management Information System (MIMIS) for data collection and management. Then followed by the second phase, which is the development of Neural Network Simulator (NNS) using back propagation for network training and testing. The final phase is the development of Prediction System (PS) for prediction on new patient's data. All systems had been developed using Microsoft's Visual Basics. The data used to train and test the network was provided by Alor Setar General Hospital, Kedah. The best network model produced prediction accuracy of 88.89 percents. Apart from proving the ability of neural network technology in medical diagnosis, this study also shown how the neural network could be integrated into a management information system as a prediction tools. As the pilot project, the application developed could be used as the starting point in building a medical decision support system, particularly in diagnosing the heart disease.

ACKNOWLEDGEMENTS

The development of this project has been a long journey. Throughout this journey, I was fortunate to have had the help and contributions of my supervisors, Puan Fadzilah Siraj. I would like to extend my thanks to my beloved wife, Azizah, my daughters, Nur Hadiyah and Nur Hazwani for the courage and understanding. This project would not have been possible without their encouragement, support and guidance.

I would also like to thank the Alor Setar General Hospital staffs, especially Dr Nizar for supplying the data of myocardial infarction. The data and advices have been tremendously useful in the development of the project.

TABLE OF CONTENTS

	Page
PERMISSION TO USE	i
ABSTRACT (BAHASA MALAYSIA)	ii
ABSTRACT (ENGLISH)	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER ONE : INTRODUCTION	
1.1 Problem Statement	2
1.2 Objectives	5
1.3 Scope of the Study	6
1.4 Significance of the Study	7
CHAPTER TWO : NEURAL NETWORK	
2.1 Neural Network Learning Method	10
2.2 Back Propagation Neural Network	11
2.3 Back Propagation Architecture	12
2.4 Back Propagation Algorithm	13
2.4.1 Back Propagation Training Algorithm	15
2.4.2 Back Propagation Activation Function	17
CHAPTER THREE : LITERATURE REVIEW	
3.1 Neural Network Application in Medical	20
3.2 Neural Network Application in Cardiovascular and Heart Disease	21

3.2.1	Diagnosis of Coronary Heart Disease	22
3.2.2	Diagnosis of Myocardial Infarction	22
3.2.3	Mortality Prognosis in Heart Failure	23
CHAPTER FOUR : PROJECT METHODOLOGY		
4.1	The Development of the Myocardial Infarction Management Information System (MIMIS)	27
4.1.1	Patient Registration Module	27
4.1.2	Maintenance Module	28
4.1.3	Reporting Module	29
4.2	Data Preparation	29
4.2.1	Data Selection	30
4.2.2	Data Preprocessing	34
4.2.3	Representation	35
4.3	Selection of Neural Network Model and Architecture	36
4.3.1	Selection of Neural Network Learning Paradigms	37
4.3.2	Selection of Neural Network Topology	37
4.3.3	Selection of Neural Network Learning Algorithm	38
4.3.4	Selection of Neural Network Model	38
4.3.5	Definition of Input and Output Nodes	39
4.4	The Development of Neural Network Simulator (NNS)	41
4.4.1	Selection of Activation Function	41
4.5	Neural Network Training and Testing	42
4.5.1	Selection of Hidden Unit	43
4.5.2	Selection of Learning Rate	44

4.5.3	Selection of Momentum Rate	45
4.5.4	Selection of Activation Function	46
4.5.5	Selection of Stopping Criteria	46
4.5.6	Use of UCI Database for Comparison	47
4.6	The development of Prediction System (PS)	48
4.7	The System Integration (MIMIS, NNS and PS)	50
CHAPTER FIVE : FINDINGS AND RESULTS		52
CHAPTER SIX : CONCLUSION AND RECOMMENDATION		60
BIBLIOGRAPHY		62
APPENDICES		
Appendix A	Explanation on Symbols Used in Main Texts	67
Appendix B	Questionnaire	68
Appendix C	Trained Neural Network Weights File	71
Appendix D	Sample Raw Data	73
Appendix E	Sample Preprocess Data	74
Appendix F	User Manual	75
Appendix G	Programming Manual	104

LIST OF TABLES

		Page
Table 4.1	Data Type and Presentation	36
Table 5.1	Selection of the Hidden Units	52
Table 5.2	The Best Hidden Unit	53
Table 5.3	The Selection of Learning Rates	53
Table 5.4	The Best Learning Rate	54
Table 5.5	The Selection of Momentum Rates	55
Table 5.6	The Best Momentum Rate	55
Table 5.7	Selection of Stopping Criteria (EPOCH)	56
Table 5.8	The Final Model	58

LIST OF FIGURES

		Page
Figure 1.1	Proposed Diagnosis System for Detecting Angina in Patients	4
Figure 2.1	Human Neuron	9
Figure 2.2	Artificial Neuron	10
Figure 2.3	Back Propagation Neural Network with One Hidden Layer	13
Figure 2.4	Back Propagation phases	14
Figure 4.1	Major Steps in Developing the Application	26
Figure 4.2	Major Steps in Neural Network Training and Testing	43
Figure 4.3	Final System's Data Flow	51
Figure 5.1	Stopping Criteria	57
Figure 5.2	Neural Network Model Implemented	58

Chapter 1 Introduction

Worldwide, cardiovascular disorders and heart disease are considered as the number one killer. Although the death rates from the diseases are now on the decrease as a result of changes in lifestyle, it is still a major cause of death and disability in both developing and developed countries. Cardiovascular disorders claimed 953,110 lives in the United States in 1997 and it contributed 41.2 percent of all deaths or 1 of every 2.4 deaths. A cardiovascular disorder was also as a primary or contributing cause on over 1,406,000 death certificates (American Heart Association, 2000). People in the developing country including Malaysian are also exposed to the same risk. According to our Ministry of Health (1997), heart disease and cardiovascular disorders contributed up to 20.26 percent of all total deaths in Malaysia or equivalent to 8,915 of death yearly.

The main function of the cardiovascular system is to supply vital oxygen as well as nutrients to all body cells and tissues will die within minutes without this supply (Calnan, 1991). A cardiovascular disorder is a medical terminology that describes all kind of diseases that related to our heart and circulatory system. The coronary heart disease (CHD) is one of the most dangerous heart problems, which occurs when the heart muscle receives insufficient oxygen because the coronary arteries fail to maintain a sufficient supply of blood (Calnan, 1991). According to Open University (1985), coronary heart disease threatening people in three main ways by producing:

The contents of
the thesis is for
internal user
only

Bibliography

- Aho, A.V., Hopcroft, J.E. & Ullman, J.D. (1974), *Design and Analysis of Algorithm*, Addison Wesley, Reading, MA
- Akay, M. (1992), *Noninvasive Diagnosis of Coronary Artery Disease Using a Neural Network Algorithm*, Biol. Cybern., 67(4):361-7
- Akay, M., Akay, Y.M. & Welkowitz, W. (1992), *Neural Networks for the Diagnosis of Coronary Artery Disease*, Proceed. Int. Joint Conf. Neural Networks, Baltimore, MD, New York: IEEE Press
- Amari, S.I. & Arbib, M.A. (1982), *Competition and Cooperation in Neural Nets*, Springer-Verlag, New York
- American Heart Association (2000), [www document]. URL.
http://www.americanheart.org/Heart_and_Stroke_A_Z_Guide/cvds.html
- Ballard, D.H. (1986), *Cortical Connections and Parallel Processing*, Behav. Brain Sc., 9(1), 67
- Baxt, W.G. (1991), *Use of an Artificial Neural Network for The Diagnosis of Myocardial Infarction*, Ann. Intern. Med., 115: 843-848
- Bigus, J.P. (1996), *Data Mining with Neural networks: Solving Business Problems - from application development to decision support*, McGraw Hill, New York
- Browner, W.S. (1992), *Myocardial Infarction Prediction by Artificial Neural Networks*, Ann. Intern. Med., 116(8):701-2
- Calnan, M. (1991), *Preventing Coronary Heart Disease: Prospects, Policies and Politics*, Routledge, London
- Cook, D.G., Pocock, S.J., Sharper, A.G. & Kussick, S.J. (1986), *Giving Up Smoking and the Risk of Heart Attacks*, Lancet, 11, 1376-9
- DARPA (1988), *Neural Network Study*, AFCEA International Press, New York
- Delogu, P. (1994), *Neural Network for Automatic Diagnosis of Labeled Monodispersed Aerosols in Obstructive Lung Disease*, Dipartimento di Fisica dell' Universiti and sezione I.N.F.N., Pisa, Italy
- Doig, G.S., Inman, K.J., Sibbald, W.J., Martin, C.M. & Robertson, J.M. (1993), *Modeling mortality in the intensive care unit: comparing the performance of a back-propagation, associative-learning neural network with multivariate logistic regression*, in InSafran C(ed), Proceedings of the Seventeenth Annual

Symposium on Computer Applications in Medical Care, McGraw-Hill, New York

- Doll, R. & Peto, R. (1976), *Mortality in Relation to Smoking*, British Medical Journal, 2, 1523-36
- Dorffner, G. & Koller, G. (1994), *Toward Improving Exercise ECG for Detecting Ischemic Heart Disease with Recurrent and Feedforward Neural Nets*, Austrian Research Institute for Artificial Intelligence, Vienna, Austria
- Downs, J., Harrison, R.F. & Kennedy, R.L., (1995), *A prototype neural network decision-support tool for the diagnosis of acute myocardial infarction*, Proceedings of the European Society for Artificial Intelligence in Medicine Fifth European Conference on Artificial Intelligence in Medicine, AIME-95, Pavia, Italy, 355-366
- Fausett, L. (1994), *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*, New Jersey: Prentice Hall
- Findler, N.V. (1991), *An Artificial Intelligence Technique for Information and Fact Retrieval: An Application in Medical Knowledge Processing*, MIT Press, England
- Gotto, A.M. (1988), *Co-existent Hypercholesterolemia and Hypertension in CHD Risk: Clinical Aspects and Considerations*, Lipid Review, 2, 3, 17-22
- Harrison, R.F., Marshall, S.J. & Kennedy, R.L. (1991a), *A connectionist aid to the early diagnosis of myocardial infarction*, Proceed. Third European Conf. Artificial Intelligence in Medicine, Maastricht, The Netherlands
- Harrison, R.F., Marshall, S.J. & Kennedy, R.L. (1991), *The early diagnosis of heart attacks: a neurocomputational approach*. Proceedings of the Institute of Electrical and Electronics Engineers and the International Neural Network Society International Joint Conference on Neural Networks, Seattle, 1-5
- Kannel, W.B. (1993), *Hypertension as a risk factor for cardiac events, epidemiologic results of long-term studies*, Journal of Cardiovascular Pharmacology, 21 Suppl 2:S27-37.
- Kannel, W. & McGhee, D. (1979), *Diabetes and Cardiovascular Disease: The Framingham Study*, JAMA, 241:2035
- Kannel, W.B. (1975), *Risk of Blood Pressure in Cardiovascular Disease: The Framingham Study*, Angiol, 26, 1-14

- Kennedy, R.L., Harrison, R.F., Kirklis, K., Moriarty, K.T., Gangi, M.T., Shaukat, H.N., & Young, M.J. (1990), *Early diagnosis of acute myocardial infarction: A novel approach to decision making using neural networks*, *Clinical Science*, 78, 88-88
- Kennedy, R.L., Harrison, R.F., Marshall, S.J. & Hardisty, C.A. (1991), *Analysis of clinical and electrocardiographic data from patients with acute chest pain using a neurocomputer*, *Quarterly Journal of Medicine*, 80, 788-789
- Kohonen, T. (1977), *Associative Memory: System Theoretic Approach*, Springer-Verlag, Berlin
- Marshall, S.J., Harrison, R.F. & Kennedy, R.L. (1991), *Neural classification of chest pain symptoms: A comparative study*. Proceedings of the Institution of Electrical Engineers Second International Conference on Artificial Neural Networks, Bournemouth. 200-204
- McArthur, D., Hamer, W.G., McAllum, R., Harrison, R.F., Kennedy, R.L. & Steedman, D.J. (1993), *Early diagnosis of acute myocardial infarction (AMI): comparison and combination of an artificial neural network and sequential myoglobin measurements*, Proceedings of the Emergency Medicine Research Society Meeting, Cambridge
- McGonigal M.D., Cole J., Schwab C.W., Kauder D.R., Rotondo M.F. & Angood P.B. (1993), *A new approach to probability of survival scoring for trauma quality assurance*. *Journal of Trauma*, 34(6):863-8
- Ministry of Health (1997), *Annual Report*, Kuala Lumpur
- Minsky, M. L. & Papert, S. A. (1988), *Perceptrons*, Expanded Edition, MA: MIT Press, Original edition, 1969
- National Heart Attack Alert Program Working Group (1997), *An Evaluation of Technology for Identifying Acute Cardiac Ischemia in the Emergency Department*, *Ann Emerg Med*, 29,13-87
- National Heart, Lung, and Blood Institute (2000), [www document]. URL. <http://www.nhlbi.nih.gov/health/public/heart/other/hrtfail.htm>
- North, G. (1987), *A Celebration of Connectionism*, *Nature*, 328, 107
- Ohno-Machado L. & Bialek D. (1998), *Diagnosing Breast Cancer from FNAs: Variable Relevance in Neural Network and Logistic Regression Models*. Brigham and Women's Hospital and Harvard Medical School
- Open University (1985), *The Biology of Health and Disease*, Milton Keynes: Open University Press

- Ortiz J., Ghefter C.G., Silva C.E. & Sabbatini R.M. (1995), *One-year mortality prognosis in heart failure: a neural network approach based on echocardiographic data*, J Am Coll Cardiol Dec;26(7):1586-93
- Parker, D. (1985), *Learning Logic*: Technical Report TR-87, MA: MIT Press
- Percy D.F. & Hine T.J. (1993), *Multivariate analysis of cholesterol distribution for monitoring the risk of coronary heart disease*, Statistics in Medicine, 12(10):967-74
- Poggio, T., Torre, V. & Koch, C. (1985), *Computational Vision and Regularization Theory*, Nature, 31-319
- Proyor, D.B., Shaw, L., McCants, C.B., Lee K.L., Mark, D.B. & Harrell F.E. (1993), *Value of the History and Physical in Identifying patients at Increased Risk of Coronary Artery Disease*, Ann Int Med, 118:81-90
- Reid, D.D., Hamilton, P.J.S., McCartney, P., Rose, G., Jarrett, R.J. & Keen, H. (1976), *Smoking and Other Risk Factors for CHD in British Civil Servants*, Lancet, ii, 979-84
- Rosenblatt, F. (1959), *Two Theorems of Statistical Separability in the Perceptron, Mechanization of Thought Processes*: Proceeding of a Symposium Held at the National Physical Laboratory, November 1958, 421-456
- Rumelheart, D. E., Hinton, G. E. & Williams R. J. (1986), *Learning Internal Representation by Error Propagation*, Parallel Distributed Processing, 1: 675-695
- Sharper, A.G., Pocock, S.J., Walker, M., Philips, A.W., Whitehead, T.P. & Macfarlane, P.W. (1985), *Risk factors for ischaemic heart disease: The Prospective Phase of the British Regional Heart Study*, Journal of Epidemiology and Community health, 39, 197-209
- Shortliffe, E.H. (1976) *Computer-Based Medical Consultation: MYCIN*, American Elsevier, New York
- Simkus, R. (1994), *Using Neural Networks In The Early Diagnosis Of Acute Myocardial Infarctions*, ITCH-94, Information Technology in Community Health Conference
- Somoza, E. & Somoza, J.R. (1993), *A neural-network approach to predicting admission decisions in a psychiatric emergency room*, Medical Decision Making, 13(4):273-80

- Stensmo, M., & Sejnowski, T.J. (1994), *A Mixture Model System for Medical and Machine Diagnosis*, Tech. Rep., INC-9401, Institute for Neural Computation, University of California, San Diego
- Swales, J.D. (1981), *Clinical Hypertension*, London, Chapman & Hall
- Truswell, A.S. (1985), *Diet and Hypertension*, British Medical Journal, 291, 125-7
- Tsoukalas, L.H. & Uhrig, R.E. (1997), *Fuzzy and Neural Approaches in Engineering*, New York: John Wileys & Sons
- Tu J.V. & Guerriere M.R. (1993), *Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery*, Computers and Biomedical Research, 26(3):220–9
- Werbos, P. (1974), *Beyond Regression: New Tools For Prediction and Analysis in the Behavioral Sciences* (Ph.D. Thesis), MA: Harvard U. Committee on Applied Mathematics