PREDICTING DISEASES USING MULTI-BACKPROPAGATION

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

by

Wan Hussain Bin Wan Ishak

PERMISSION TO USE

In presenting this thesis in fulfillment of the requirements for a post graduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis 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 thesis 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 thesis.

Request for permission to copy or to make other use of materials in this thesis, in whole of in part, should be addressed to:

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

ABSTRAK

Pada masa kini, sistem perubatan berkomputer memainkan peranan besar dalam amalan perubatan. Pada peringkat awal komputer digunakan bagi menyimpan dan menguruskan maklumat secara efektif. Peranan komputer kemudiannya menjadi lebih penting sejak pengenalan kepada sistem perkomputeran pintar. Sistem perubatan pintar meningkatkan keupayaan pengamal perubatan bagi membuat diagnosis dan peramalan. Rangkaian neural merupakan salah satu daripada teknik kepintaran buatan yang menyamai fungsi neuron biologi manusia. Rangkaian neural membolehkan komputer "belajar" dan "berfikir" seperti mana manusia. Walau bagaimanapun, lazimnya pembelajaran oleh rangkaian melibatkan jumlah data yang banyak. Lebih banyak data digunakan, rangkaian akan menjadi lebih kompleks. Rangkaian yang kompleks lebih susah untuk dilatih dan akan mengambil masa yang lama untuk mencapai tahap generalisasi.

Kajian ini mencadangkan pendekatan berbilang rangkaian (*multi network*) berbanding pendekatan satu rangkaian (*single network*). Pendekatan berbilang rangkaian tidak memerlukan sebarang perubahan dalam algoritma pembelajaran. Sebaliknya, set data yang besar dipecahkan kepada beberapa kategori atau rangkaian yang lebih kecil. Kedua-dua pendekatan tersebut dikaji dan dibuat perbandingan. Dapatan kajian menunjukkan anggaran masa bagi rangkaian yang mempunyai 26 pembolehubah untuk 100 peratus pencapaian teritlak berdasarkan 7,466 set data ialah lebih kurang 1,037,472,836 millisaat. Sebaliknya berdasarkan 256 set data rangkaian mengambil masa 2,459,172,864 millisaat untuk melengkapkan pembelajaran. Jumlah pengulangan (*epoch*) dianggarkan 359,544 dan 26,214,400 bagi kedua-dua set data.

Bagi pendekatan berbilang rangkaian, lima rangkaian yang berbeza dan satu rangkaian gabungan telah dibina. Eksperimen menunjukkan keenam-enam rangkaian dapat belajar atau menghafal corak data dengan lengkap dalam beberapa pengulangan sahaja. Masa yang diambil oleh rangkaian adalah masing-masing 281, 197, 32, 440, 83 dan 22 bagi rangkaian *Risk Factor*, *Medication*, *Investigation*, *ECG*, *Complication* dan *Integration*. Secara purata pendekatan ini mengambil masa 175.833 millisaat dan 7.66667 pengulangan untuk rangkaian belajar. Secara keseluruhan, jumlah masa yang diambil oleh keenam-enam rangkaian ialah 1055 millisaat dengan 46 pengulangan.

Walaupun beberapa rangkaian terpaksa dibina dan dilatih secara berasingan, pendekatan berbilang rangkaian telah mengurangkan kekompleksan rangkaian yang besar dan mengatasi kelemahan pendekatan satu rangkaian. Ini kerana rangkaian-rangkaian yang dihasilkan dalam pendekatan ini mewakili kesemua kombinasi data dan kesemua data tersebut digunakan bagi melatih rangkaian. Dengan kata lain, melalui pendekatan berbilang rangkaian kesemua set data digunakan dalam proses pembelajaran rangkaian. Pengetahuan (pemberat) yang dihasilkan oleh rangkaian boleh diaplikasikan bagi kesemua kemungkinan set data.

ABSTRACT

A Computer-based medical system plays an important role in the current practice of medicine. Initially, computer is used to store and manage information effectively. The computer becomes more important with the introduction of the intelligent system. The intelligent medical system increases the ability of medical practitioners in providing diagnosis and prognosis. Neural network is one of the artificial intelligence techniques that emulate the human neuron function. Neural network enable the computer to "learn" and "think" like human. However, learning usually involves a large amount of data. If more data is used, the network complexity will be increased. Complex network is hard to learn and take more time to generalize.

Thus this study proposed a multi-network approach as oppose to the single network approach. Multi-network approach does not require any changes in neural network learning algorithm. Instead, the large data is divided into several smaller categories or network. Both approaches are tested and compared. The results show that the estimation time for the single network with 26 variables based on 7466 data set is approximately 1,037,472,836 milliseconds to complete the learning with 100 percent generalization performance. On the other hand, based on 256 data sets the network takes 2,459,172,864 milliseconds to complete the learning. The epochs are estimated as 359,544 and 26,214,400 respectively.

In the multi-network approach, five different networks and one integration network were constructed. The experiments showed that all six networks managed to learn the data completely in only several epochs. The time taken by the networks are 281, 197, 32, 440, 83 and 22 respectively for the risk factor, medication, investigation, ECG, complication and integrating network. On average, this approach takes 175.833 milliseconds and 7.66667 epochs to complete the learning. The total training time for all networks to learn is 1055 milliseconds with 46 epochs.

Although many networks have to be constructed and trained separately, the multi-network approach has reduced the complexity of network with large data set and has overcome the limitation of the single network approach. This is because the networks represent all the possible combination of data, which were all used to train them respectively. That is in the multi network approach all data sets are used in training. The knowledge (weight) produced by the network can be applied for all possible data sets.

ACKNOWLEDGEMENTS

It is a great relief to submit this thesis. Words cannot describe the magnitude of this relief. Many thanks to the Most Gracious and Most Merciful, Allah for giving me the patience and consistency in completing this research. To my wife Fadhilah Mat Yamin, this work would never finish without her support and patience even though she is also in the same battle!

My deepest appreciation and gratitude to both my supervisors Professor Dr. Abu Talib Othman and Associate Professor Fadzilah Siraj for their guidance, support and ideas in the completion of this research and thesis.

I acknowledge Ministry of Health and General Hospital Alor Setar for allowing the access of medical records. This work also would not have been possible without the advice and assistance provided by Associate Professor Dr. Ismail Yaacqb from Kedah Medical Centre, Kedah, Dr. Hasmannizar Abdul Manap and Mr. Hilmi Harun from General Hospital Alor Setar, Kedah. For the financial support, I am grateful to *Universiti Utara Malaysia* for the scholarship, study leave and allowances.

My gratitude to Associate Professor Dr. Wan Rozaini Sheikh Osman for spending her precious time reading and editing this thesis before it being submitted to thesis committee. Not to forget the help from Dr. Azuraliza Abu Bakar in editing and preparing the final version of the thesis.

Special appreciation to my parents for their loves and prayers. Sincere thanks to friends and colleagues, sharing experiences throughout the years.

Wan Hussain Wan Ishak
December 2002

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

High-risk diseases such as diabetes, heart disease and pneumonia are the main cause of deaths every year. Heart disease has been recorded as the number one killer in Malaysia (Appendix A). Many studies and reports had shown that heart disease is the leading cause of death for most countries such as United States and developing countries (*such as in* Hennekens *et al.*, 1997). A compilation of reports from 1990 to 1998 of Malaysia Ministry of Health (MOH) indicates that heart disease and disease of pulmonary caused 4175, 4146, 3967, 3873, 4038, 4241, 4395, 4446 and 4248 numbers of deaths in 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997 and 1998 respectively (*see* Figure 1.1). Each number of deaths due to heart disease contributes to more than 10% to the total number of deaths in Malaysia in a particular year.

In conjunction to this problem, the current need for medical expertise has increased. However, the ratio of doctors compared to the number of patients is not proportionate. A study by Sulong and Mulyadi (1992) reveals that the number of General Practitioners, Neurologists and Cardiologists in 1990s to the beginning of the 21st century in Malaysia will be insufficient. The study showed that the employment of General Practitioners, Neurologists and Cardiologists from 1992 to the year of 2000 were reported as 4.5%, 7.6% and 6.5% (Sulong and Mulyadi, 1992).

The contents of the thesis is for internal user only

Bibliography

Abdi, H., Valentin, D., and Edelman, B. (1999). *Neural Networks*. US: Sage Publications.

Abidi, S. S. R., and Goh, A. (1998). Neural Network Based Forecasting of Bacteria-Antibiotic Interactions for Infectious Disease Control. *In 9th World Congress on Medical Informatics (MedInfo'98), Seoul August 18-22.*

AHA-American Heart Asociation (2000). *2001 Heart and Stroke Statistical Update.* Dallas: American Heart Association.

Ahmed, M. N., and Farag, A. A. (1998). Two-stage Neural Network for Medical Volume Segmentation. *Accepted for Publication in the Journal of Pattern Recognition Letters, 1998.*

Awang, Khalid. (2000). An Evaluation of Artificial Neural Network in Predicting the Presence of Heart Disease. *MSc. Thesis.* Universiti Utara Malaysia.

Bakker, B., and Heskes, T. (1999). A neural-Bayesian Approach to survival analysis, *Proceedings ICANN'99*, 832-837.

Bakker, B., Kappen, B., and Heskes, T. (2000). Surival Analysis: A neural-Bayesian approach. *Proceedings ANNIMAB-1.* Downloaded form WWW http://citeseer.nj.nec.com/350957.html.

Bernstam, E. V., Strasberg, H.R., and Rubin, D. L. (2000). Cost Benefit Analysis of Computer-Based Patient Records with Regard to their use in Colon Cancer Screening. *APAMI-MIC2000 Hong Kong.*

Bottaci, L., and Drew, P. J. (1997). Artificial Neural Networks Applied to Outcome Prediction for Colorectal Cancer Patients in Separate Institutions. *Lancet, 350(9076)*, 469-473.

Byrne, S., Cunningham, P., Barry, A., Graham, I., Delaney, T., and Corrigan, O. I. (2000). Using Neural Nets for Dicision Support in Prescription and Outcome Prediction in Anticoagulation Drug Therapy. In N. Lavrac, S. Miksch and B. Kavsek (Eds.), *Proceedings of 5th International Workshop on Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP-2000).* (pp. 9-12) Berlin, Germany.

Cardot, H., Revenu, M., Victorri, B., Revillet, M. J. (1994). A Static Signature Verification System Based on a Cooperating Neural Networks Architecture.

International Journal of Pattern Recognition and Artificial Intelligence. 8(3), pp: 679-692.

Dybowski R, Weller P, Chang R, Gant V. (1996). Prediction Of Outcome In The Critically Ill Using An Artificial Neural Network Synthesized by A Genetic Algorithm. *Lancet, 347*, 1146-1150.

Dybowski, R. (2000). Neural Computation in Medicine: Perspective and Prospects. In Malmgren, H., Borga, M., Niklasson, L. (eds.) *Proceedings of the ANNIMAB-1 Conference (Artificial Neural Networks in Medicine and Biology), Goteborg, 13-16 May 2000.* (pp. 26-36), Springer.

Erenshteyn, R., and Laskov, P. (1996). A multi-stage approach to fingerspelling and gesture recognition. *Proceedings of the Workshop on the Integration of Gesture in Language and Speech*, (pp. 185-194), Wilmington, DE.

Fausett, L. (1994). *Fundamentals of Neural Network: Architectures, Algorithms and Applications.* Englewood Cliffs: Prentice Hall.

Furle, T., and Schikuta, E. (1997). PANNS – A Parallelized Artificial Neural Simulator. *In Proceedings Fourth International Conference on Neural Information Processing (ICONIP'97)*, Dunedin, New Zealand, Springler-Verlag.

Grundy, S. M., Benjamin, I. J., Burke, G. L., Chait, A., Eckel, R. H., Howard, B. V., Mitch, W., Smith, S. C., and Sowers, J. R. (1999). Diabetes and Cardiovascular Disease: A Statement for Healthcare Professionals from the American Heart Association. *Circulation, 100*, 1134-1146.

Hennekens, C. H., Dyken, M. L., and Fuster, V. (1997). Aspirin as a Therapeutic Agent in Cardiovascular Disease. *Circulation, 96*, 2751-2753.

HIVC-Health Ink & Vitality Communications (2001, June). Life After a Heart Attack. *HealthToday Magazine*, 34-36.

Kemsley, D. H., and Martinez, T. R. (1992). A Survey of Neural Network Research and Fielded Applications. *In International Journal of Neural Networks: Research and Applications, 2(2/3/4)*, 123-133

Kroenke, D. M. (1997). *Database Processing: Fundamentals, Design and Implementation (6th ed).* Prentice Hall; USA.

Lawrence, S., Giles, C. L., and Tsoi, A. C. (1997). Lessons in Neural Network Training: Training May be Harder than Expected. *Proceedings of the Fourteenth National Conference on Artificial Intelligence, AAAI-97*, (pp. 540-545), Menlo Park, California: AAAI Press.

Lee, A., Ulbricht, C., and Dorffner, G. (1999). Application of Artificial Neural Networks for Detection of Abnormal Fetal Heart Rate Pattern: A Comparison with Conventional Algorithms. *Journal of Obstetrics and Gynaecology, 9(5)*, 482-485

Manamohan, B. E., and Deraman, Aziz. (2000). ICDC: Sistem Diagnostik Penyakit Jantung. *Prosiding Simposium Teknologi Maklumat 2000: Penyelidikan dan Pembangunan Kejuruteraan Maklumat Menjelang Alaf Baru*, (pp. 45-48), Bangi: Universiti Kebangsaan Malaysia.

Mat Isa, N. A., Mashor, M. Y., and Othman, N. H. (2002). Diagnosis of Cervical Cancer using Hierarchical Radial Basis Function (HiRBF) Network. *In Sazali Yaacob, R. Nagarajan, Ali Chekima (Eds.), Proceedings of the International Conference on Artificial Intelligence in Engineering and Technology* (pp: 458-463). June 17-18 2002, Kota Kinabalu, Sabah, Malaysia

Ministry of Health (1990). *Annual Report 1990*. Kuala Lumpur: Ministry of Health.

Ministry of Health (1991). *Annual Report 1991*. Kuala Lumpur: Ministry of Health..

Ministry of Health (1992). *Annual Report 1992*. Kuala Lumpur: Ministry of Health.

Ministry of Health (1993). *Annual Report 1993*. Kuala Lumpur: Ministry of Health.

Ministry of Health (1994). Annual Report 1994. Kuala Lumpur: Ministry of Health.

Ministry of Health (1995). *Annual Report 1995*. Kuala Lumpur: Ministry of Health.

Ministry of Health (1996). *Annual Report 1996*. Kuala Lumpur: Ministry of Health.

Ministry of Health (1997). *Annual Report 1997*. Kuala Lumpur: Ministry of Health.

Ministry of Health (1998). *Annual Report 1998*. Kuala Lumpur: Ministry of Health.

Misra, M. (1996). Parallel Environments for Implementing Neural Networks. *Neural Computing Surveys, 1*, 48-60.

Ockene, I. S., and Miller, N. H. (1997). Cigarette Smoking, Cardiovascular Disease and Stroke. *Circulation, 96*, 3243-3247.

Ohlsson, M., Holst, H., and Edenbrandt, L. (1999). Acute Myocardial Infarction: Analysis of the ECG Using Artificial Neural Networks. *Technical Report LU TP 99-34* (submitted to *ANNIMAB-1; Artificial Neural Networks in Medicine and Biology*).

Ohno-Machado, L. (1996). Medical Applications of Artificial Neural Networks: Connectionist Model of Survival. *Ph.D Dissertation*. Stanford University.

Partridge, D., Abidi, S. S. R., and Goh, A. (1996). Neural Network Applications in Medicine. *Proceedings of National Conference on Research and Development in Computer Science and Its Applications (REDECS'96)*, (pp. 20-23), Kuala Lumpur: Universiti Pertanian Malaysia.

Passold, F., Ojeda, R. G., and Mur, J. (1996). Hybrid Expert System in Anesthesiology for Critical Patients. *In Proceedings of the 8th IEEE*

Mediterranean Electrotecnical Conference - MELECON'96 (ITALIA), III, pp. 1486-1489.

Pofahl, W. E., Walczak, S. M., Rhone, E., and Izenberg, S. D. (1998). Use of an Artificial Neural Network to Predict Length of Stay in Acute Pancreatitis. *American Surgeon, Sep98, Vol. 64 Issue 9*, (pp: 868 – 872)

Prank, K., Jurgens, C., Muhlen, A., and Brabant, G. (1998). Predictive Neural Networks for Learning the Time Course of Blood Glucose levels from the Complex Interaction of Counterregulatory Hormones. *Neural Computation, 10(4)*, 941-954.

Prechelt, L. (1998). Early Stopping-but when? *Neural Networks: Tricks of the trade*, (pp. 55-69). Retrieved March 28, 2002 from World Wide Web: http://wwwipd.ira.uka.de/~prechelt/Biblio/

Sarle, W. (1995). Stopped Training and Other Remedies for Overfitting. *Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics*, (pp. 352-360). Retrieved March 18, 2002 from World Wide Web: ftp://ftp.sas.com/pub/neural/

Sarle, W. S. (1994a). Neural Networks and Statistical Models. *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, (pp. 1538-1550), Cary, NC: SAS Institute. Retrieved March 18, 2002 from World Wide Web: ftp://ftp.sas.com/pub/sugi19/neural/

Sarle, W. S. (1994b). Neural Network Implementation in SAS Software. *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, (pp. 1551-1573), Cary, NC: SAS Institute. Retrieved March 18, 2002 from World Wide Web: ftp://ftp.sas.com/pub/sugi19/neural/

Sarle, W. S. (1997a). Neurak Network FAQ, part 1 of 7: Introduction. Periodic posting to the Usenet newsgroup comp.ai.neural-nets, URL: ftp://ftp.sas.com/pub/neural/FAQ.html Downloaded on 30 Nov. 1999.

Sarle, W.S., (1997b), Neural Network FAQ, part 2 of 7: Learning. *Periodic posting to the Usenet newsgroup comp.ai.neural-nets*, Retrieved March 28, 2002 from World Wide Web: ftp://ftp.sas.com/pub/neural/FAQ.html

Schetinin, V. (2001a). A Neural Network Decision Tree for Learning Concepts from EEG Data. *NIMIA-SC2001 – 2001 NATO Advanced Study Institute on Neural Networks for Instrumentation, Measurement and Related Industrial Applications: Study Cases*, Crema, Italy.

Schetinin, V. (2001b). Polynomial Neural Networks Learnt to Classify EEG Signals. *NIMIA-SC2001 – 2001 NATO Advanced Study Institute on Neural Networks for Instrumentation, Measurement and Related Industrial Applications: Study Cases*, Crema, Italy.

Serbedzija, N. B. (1996). Simulating Artificial Neural Networks on Parallel Architectures. *Computer, 29(3)*, 56-63.

Sheikh Abdulah, Siti Norul Huda., and Surip, Miswan. (1999). Satu Metodologi Perlombongan Data untuk Pesakit AIDS. *Proceedings of Artificial Intelligence Applications in Industry (AIAI'99) Oct. 20-21*, (pp. 57-71), Kuala Lumpur: AISDEL SIRIM.

Shortliffe, E. H. (1987). Computer Programs to Support Clinical Decision Making. *Journal of the American Medical Association, 258(1).*

Silipo, R., and Marchesi, C. (1998). Artificial Neural Networks for automatic ECG analysis, *IEEE Transactions on Signal Processing, 46(5)*, 1417-1425.

Sima, J. (1994). Loading Deep Networks is Hard. *Neural Computation, 6(5)*, 842-850.

Sima, J. (1996). Back Propagation is Not Efficient. *Neural Networks, 9(6)*, 1017-1023.

Sulaiman, M. N., and Evans, D. J. (1996). NEUCOMP2 – Parallel Neural Network Compiler. *Malaysian Journal of Computer Science, 9(2)*, 54-70.

Sulong, Ghazali., and Mulyadi, M. T. (1992). Model Ramalan Keperluan Doktor Umum, Pakar Jantung dan Pakar Saraf Bagi Hospital Kerajaan Di Malaysia, *Jurnal Teknologi Maklumat, 4(1)*, 15 - 30.

Thimm, G., Moerland, P., and Fiesler, E. (1996). The Interchangebility of Learning Rate and Gain in Backpropagation Neural Networks. *Neural Computation, 8(2).*

Tveter, D. R. (1997). The Size of the Network. *Backpropagator's Review*, Retrieved December 6, 2001 from World Wide Web: http://www.dontveter.com/bpr/netsize.html.

Wan Ishak, Wan Hussain & Siraj, Fadzilah (2000). A Model for Web-Based Medical Diagnosis and Prediction. *Proceeding Persidangan Ulang Tahun Ke-30 UKM (5-6 Sept. 2000).* (pp. 238-247) Bangi: Universiti Kebangsaan Malaysia.

Wan Ishak, Wan Hussain., Siraj, Fadzilah., and Othman, Abu Talib. (2000). Pengurusan Pangkalan Data Berpusat Bagi Pesakit Di Hospital. *Presented at Simposium Pengurusan Penyelidikan Siswazah (15-16 November 2000),* Universiti Utara Malaysia.

Watrous, R., and Towell, G. (1995). A Patient-Adaptive Neural Network ECG Patient Monitoring Algorithm. *Computer in Cardiology*, Vienna, Austria.

Weigang, L., and Silva, N. C. (1999). A Study of Parallel Neural Networks. *In Proceedings of International Joint Conference on Neural Networks, 2*, (pp: 1113-1116).

Yoshihara, I., Kamimai, Y., and Yasunaga, M. (2001). Feature Extraction from Genome Sequence using Multi-Model Neural Network. *Genome Informatics 12*, 420-422.