# THE IMPACT OF NORMALIZATION TECHNIQUES ON PERFORMANCE BACKPROPAGATION NETWORKS

**NORLIDA BINTI HASSAN**

**UNIVERSITI UTARA MALAYSIA**
**2004**

# THE IMPACT OF NORMALIZATION TECHNIQUES ON PERFORMANCE BACKPROPAGATION NETWORKS

A thesis submitted to the Faculty of Information Technology in partial
Fulfillment of the requirements for the degree
Master of Science (Intelligent Knowledge Based System)
Universiti Utara Malaysia

By

Norlida binti Hassan

**JABATAN HAL EHWAL AKADEMIK**
***(Department of Academic Affairs)***
**Universiti Utara Malaysia**

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

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

**NORLIDA BINTI HASSAN**

calon untuk Ijazah
(*candidate for the degree of*) **MSc. (IKBS)**

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

**THE IMPACT OF NORMALIZATION TECHNIQUES ON PERFORMANCE BACKPROPAGATION NETWORKS**

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 filed is covered by the project paper*).

Nama Penyelia Utama
(*Name of Main Supervisor*) : **PROF. MADYA FADZILAH SIRAJ**

Tandatangan
(*Signature*) : Tarikh (Date): 28/03/2004

Nama Penyelia Kedua
(*Name of 2nd Supervisor*) : **ENCIK AZIZI AB AZIZ**

Tandatangan
(*Signature*) : Tarikh (Date): 28/03/2004

# PERMISSION TO USE

In presenting this thesis in partial 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 part, for scholarly purpose may be granted by my supervisor(s) or, in thesis absence, by the Dean of the Faculty of Information Technology. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without any 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 part, should be addressed to:

**Dean of Faculty of Information Technology**
**Department of Computer Science**
**Universiti Utara Malaysia**
**06010 Sintok**
**Kedah Darul Aman**

# ABSTRACT (BAHASA MELAYU)

Rangkaian neural (NN) adalah merupakan suatu model perkomputeran yang mempunyai keupayaan untuk belajar, pengitlakkan, dan model yang lazim digunakan ialah perseptron berlapis (MLP). Pembangunan NN yang baik bergantung kepada beberapa aspek seperti proses perolehan, pembentukan model, dan pemilihan model yang sesuai. Data perlu ditukarkan dalam bentuk yang boleh diterima pakai sebagai input kepada rangkaian MLP. Data yang telah melalui proses transformasi biasanya lebih berkesan dan mungkin mampu menghasilkan keputusan yang lebih tepat daripada rangkaian. Kajian yang dijalankan merangkumi beberapa jenis teknik normalisasi dengan penggunaan pembelajaran rambatan balik. Teknik normalisasi yang digunakan dalam eksperimen ialah teknik *Min-Max, Z-Score, Decimal Scaling, Sigmoidal* dan *Softmax* atau *Logistic*. Untuk mengkaji kesan teknik normalisasi terhadap prestasi NN, set data perubatan dengan nilai target Boolean dipraproses, dilatih, diukur kesahihan, dan pengujian menggunakan algoritma pembelajaran rambatan balik. Kriteria pemilihan model yang terbaik adalah berdasarkan peratus ramalan benar yang tertinggi. Tiga set data dan lima teknik normalisasi telah dilaksanakan pada fasa praproses data didalam proses pembangunan model NN. Keputusan bagi setiap teknik telah dikaji dan dipersembahkan, kemudian dibandingkan dengan pendekatan statistik. Keputusan menunjukkan bahawa penggunaan teknik normalisasi yang berbeza menghasilkan prestasi yang berbeza. Kajian ini juga mendapati bahawa kelima-lima teknik normalisasi yang dikaji keatas pendekatan statistik (*Logistic regression*) menunjukkan peratus ramalan benar yang rendah berbanding dengan pendekatan NN. Dapatan kajian ini tidak hanya menyumbang kepada peningkatan prestasi algoritma pembelajaran rangkaian rambatan balik tetapi juga dapat membantu didalam membuat keputusan bagi pemilihan teknik normalisasi yang sesuai bagi data tertentu.

# ABSTRACT (ENGLISH)

Neural networks (NN) are computational models with the capacity to learn, generalize, and the most used are multi-layer perceptrons (MLP). Building successful NN applications depends on several aspects such as the process of acquiring, modeling and selecting the appropriate model. The data needs to be transformed into a form that is acceptable as input to the MLP network. The transform data often determines the efficiency and possibly the accuracy of results from the network. This study explored several normalization techniques using backpropagation learning. The normalization techniques used in the experiments are Min-Max, Z-Score, Decimal Scaling, Sigmoidal, and Softmax or Logistic techniques. To explore the impact of normalization techniques on the performance on NN, medical datasets with Boolean target were preprocessed, trained, validated and tested using backpropagation learning algorithm. The criterion of choosing the best model is based on the highest percentage of correct prediction. Three datasets and five normalization techniques were employed in the data preprocessing phase of building the NN models. The results of each normalization techniques are presented and compared with statistical approach. The results reveal that the utilization of different normalization techniques produces different NN performance. The experiments also indicate that all five normalization techniques of logistic regression achieve lower percentage of correct prediction than the results produced using NN. The findings will not only contribute towards enhancing the performance of backpropagation nets but it will also assist in making decision to the choice of normalization techniques to be applied to a particular dataset.

# ACKNOWLEDGEMENT

*By the Name of Allah, the Most Gracious and the Most Merciful*

I owe a debt of my sincere gratitude and deep appreciation to my supervisors, Associate Professor Fadzilah Siraj and Mr. Azizi Ab Aziz, Faculty of Information Technology, Universiti Utara Malaysia (UUM) whom I have learned much about neural networks. I wish to acknowledge their assistance and time, provided excellent facilities, support and guidance throughout the project also for their advice during this project.

I would also like to express thanks to my beloved parents for their support, devotion and encouragement during my study. Last but not least, special thanks to my dear friends for their encouragement throughout the study. There is a tremendous sense of achievement in completing this study. It is certain that the project would not successful it did without their help and support.

**Norlida Binti Hassan**
**Faculty of Information Technology**
**Department of Computer Science**
**Universiti Utara Malaysia**
**March 2004**

# TABLE OF CONTENT

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

In this chapter, the first section describes the context of the study that presents an introduction to data preparation in Neural Network (NN), followed by the problem statement, the objectives of the study and the significance of the study. Finally, the scope of the study that includes the limitations of the study is also discussed.

## 1.1 Preprocessing

Data in the real world are often dynamic, incomplete which are lacking attribute values, lacking certain attributes of interest, or containing only aggregate data, containing errors or outliers that is known as noisy data (Han and Kamber, 2001). It also an inconsistent which are containing discrepancies in codes or names, and most of the problem which appear in the data mining step within the knowledge discovery process are due to an inadequate preprocessing of the data present in these databases (Gomez-Skarmeta, 1997).

Before applying any of the built-in functions for training, it is important to check that the data is *reasonable* (Wolfram, 2003). Naturally, it cannot expect to obtain good models from poor or insufficient data. Unfortunately, there is no standard procedure that can be used to test the quality of the data. Even if they appear to be quite reasonable, it might be a good idea to consider preprocessing the data before initiating training. Preprocessing is a transformation, or conditioning, of data designed to make modeling easier and more robust (Wolfram, 2003). For example, a known nonlinearity in some given data could be removed by an appropriate transformation, producing data that conforms to a linear model that is easier to work with.

The contents of the thesis is for internal user only

# BIBLIOGRAPHY


Abdullah, C.S., Siraj, F. and Abu Bakar, M.D. (2001) Design of Normal Concrete Mixes Using Neural Network Model. In Proceeding of the 2nd Conference on Information Technology in Asia, in Collaboration with Global Information & Telecommunication Institute. October 17-19, 2001.

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.

Adali, S., Sapino, M. L., V. S., and Subrahmanian, V. S., (1999) A Multimedia Presentation Algebra. SIGMOD Conference 1999: 121-132.

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.

Alkharouf, N., Cervone, G., El-Askary, H., Tang, J. and Nefissi, S. (2002) A *Multistrategy Approach for the Investigation of Gene Behaviour in Space.* URL: http://www.scs.gmu.edu/~nalkhar3/CSI801/project.html Accessed date: 6 December 2003.

Armoni, A., (1998) Use of neural networks in medical diagnosis, MD Computing, Mac-Apr;15(2):100-4

Aronson, A. R., Bodenreider, O., Chang, H. C., Humphrey, S. M., Mork, J. G., Nelson, S. J., Rindflesch, T. C. and Wilbur, W. J. (1999) The Indexing Initiative, A Report to the Board of Scientific Counselors of the Lister Hill National Center for Biomedical Communications.

Beale, R. and Edwards, A. D. N. (1992) Recognising postures and gestures using neural networks. in R. Beale and J. Finlay (ed.) *Neural Networks and Pattern Recognition in Human-Computer Interaction.* New York: Ellis Horwood. pp. 163-169.

Bennett, K. P., and Mangasarian, O. L., (1992) Robust linear programming discrimination of two linearly inseparable sets, Optimization Methods and Software 1, 23-34, Gordon & Breach Science Publishers.

Bennett, P. (1996) A neural net-based weather prediction system. Project Report, Knowledge-Based Systems MSc. School of Cognitive and Computing Sciences, University of Sussex. URL: http://freepages.pavilion.net/users/pjb/wps0170.htm Accesed date: 6 December 2003.

Bigus, J.P. (1996) *Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support*, McGraw Hill, New York.

Bishop, C. (1995). *Neural Networks for Pattern Recognition.* Oxford: University Press.

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

Byun, H. and Ko, B. C. (2003) Robust face detection and tracking for real-life applications. *International Journal of Pattern Recognition and Artificial Intelligence.*

Caruana, R., Baluja, S., and Mitchell, T. (1996) Using the Future to "Sort Out" the Present: Rankrop and Multitask Learning for Medical Risk Evaluation. *Advances in Neural Information Processing Systems 8*, The MIT Press, Cambridge, pp. 959-965.

Cestnik, G., Konenenko,I, & Bratko,I. (1987) Assistant-86: A Knowledge-Elicitation Tool for Sophisticated Users. *In I.Bratko & N.Lavrac (Eds.) Progress in Machine Learning*, 31-45, Sigma Press.

Chowdhury, A., Aljlayl, M., Jensen, E., Beitzel, S., Grossman, D. and Frieder, O. (2002) Linear Combinations Based on Document Structure and Varied Stemming for Arabic Retrieval. *Proceedings of the $10^{th}$ Text Retrieval Conference (TREC-2002).*

Clark, P. and Niblett, T. (1987) Induction in Noisy Domains. In Progress in Machine Learning, In *Proceedings of the $2^{nd}$ European Working Session on Learning*, 11-30, Bled, Yugoslavia: Sigma Press.

DARPA Neural Network study (1988) AFCEA International Press, p. 60.

Deboeck, Guido J. & Cader, Masud. (Ed.) (1994). *Pre- and Postprocessing of Financial Data.* Canada: John Wiley & Sons. pp. 27-44.

DeLurgio, S. A. (2000) *Forecasting Principle and Applications.* McGraw-Hill International Editions.

Demuth, H., and Beale, M., (1998) *Neural Network Toolbox User's Guide* Ver.3 For Use With Matlab. Massachusetts: The MathWorks Inc.

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. Springer, pp. 26-36.

Engels, R., and Theusinger, C., (1998) Using a Data Metric for Preprocessing Advice for Data Mining Applications. 19th European Conference on Artificial Intelligence. John Wiley & Sons, Ltd.

Feng, C., Sutherland, A., King, S., Muggleton, S. and Henery, R. (1993). Comparison of Machine Learning Classifiers to Statistics and Neural Networks. AI & Stats Conf. 93.

Fielden, M. R., Halgren, R. G., Dere, E., and Zackarewski, T. R., (2002) GP3: GenePix Post-processing Program for Automated Analysis of Raw Microarray Data. *Bioinformatics Application Notes*. Vol. 18. No. 5, pg 771-773.

Gomez-Skarmeta, A. F., Jimenez, F., and Ibanez, J., (1997) Data Preprocessing in Knowledge Discovery with Fuzzy-Evolutionary Algorithm. Departamento de Informatica, Universidad de Murcia.

Han, J., and Kamber, M., (2001) *Data Mining: Concept and Techniques*. Simon Fraser University, Canada.

Haykin, S. (1999) *Neural Networks: A Comprehensive Foundation*, 2nd Edition, New Jersey: Prentice Hall.

He, D., Park, H. R., Murray, G. C., Subotin, M. and Oard, D. W., (2002) Topic Tracking at the University of Maryland, Institute for Advanced Computer Studies, University of Maryland.

Heden, B., Ohlsson, M., Rittner, R., Pahlm, O., Haisty, W. K., Peterson, C., and Denbrandt, L. (1996) Agreement Between Artificial Neural Networks and Human Expert for the Electrocardiographic Diagnosis of Healed Myocardial Infarction. *Journal of the American College of Cardiology*, Vol. 28, pp. 1012-10s16.

Hoong, N. K. (1988) Medical Information Science - Framework and Potential. *International Seminar and Exhibition Computerization for Development-the Research Challenge*, Universiti Pertanian Malaysia: Kuala Lumpur, pp. 191 - 198.

Jankowski, N. (1999) Approximation and Classification in Medicine with IncNet Neural Networks. *Machine Learning and Applications: Machine Learning in Medical pplications*. Chania, Greece, pp. 53-58.

Kaastra, I. and Boyd, M. (1996) Designing a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing* 10: 215-236. *Elsevier Science B.V.*

Karkanis, S. A., Magoulas, G. D., Grigoriadou, M., and Schurr, M. (1999) Detecting Abnormalities in Colonoscopic Images by Textual Description and Neural Networks. *Machine Learning and Applications: Machine Learning in Medical Applications*. Chania, Greece, pp. 59-62.

Katz, J. O. (1990) Developing neural network forecasters for trading, *Technical Analysis of Stocks and Commodities*. April 1990. 58-70.

Keally, R. (1999) *Artificial Neural Network: An Introductory Course*. URL: http://www.maths.uwa.edu.au/~rkeally/ann_all/ Accessed date: 9 January 2002.

Kivinen, J. and Warmuth, M. K. (2001) Relative Loss Bounds For Multidimensional Regression Problems. Kluwer Academic Publishers. Machine Learning, 45, 301–329.

Klaussen, K. L. and Uhrig, J. W. (1994) Cash soybean price prediction with neural networks; *NCR-134 Conference on Applied Commodity Analysis, Price Forecasting, and Market Risk Management Proceeding*, Chicago, 56-65.

Salchenberger, L. M., Mine Cinar, E., and Lash, N. A. (1992) Neural Network: A new tool for predicting thrift failures; *Decision Sciences* 23 (July/Aug. 1992) 899-916.

Klerfors, D. (1998) Artificial Neural Networks, School of Business and Adminstration, Saint Louis University.

Lapuerta, P., Azen, S.P., and La, B. (1995) Use of Neural Networks in predicting the risk of coronary artery disease, Comput Biomed Res, Feb;28(!):38-52

Lapuerta, P., Rajan, S., and Bonacini, M. (1997) Neural networks as predictors of outcomes in alcoholic patients with severe liver disease, Hepatology. Feb;25(2):302-6

Lawrence, J., (1991) Data Preparation for a Neural Network. *AI Expert,* Vol. 6, No.11, 34-41.

LiMin Fu. (1994). *Neural Networks in Computer Intelligence.* Singapore: Mc-Graw Hill. pp. 18-19, 31, 80-82.

Lippmann, R. P., Kulkolich, L., Shahian, D. (1995) Predicting the Risk of Complications in Coronary Artery Bypass Operations Using Neural Networks. *Advances in Neural Information Processing Systems* 7, The MIT Press, Cambridge, pp. 1053-1062.

Mangasarian, O. L., and Wolberg, W. H., (1990) Cancer diagnosis via linear programming, SIAM News, Vol. 23, Number 5, September 1990, pp 1 & 18.

Mangasarian, O. L., Setiono, R., and Wolberg, W.H. (1990) Pattern recognition via linear programming: Theory and application to medical diagnosis, in: Large-scale numerical optimization, Thomas F. Coleman and Yuying Li, editors, SIAM Publications, Philadelphia 1990, pp 22-30.

McCulloch, W., and Pitts, W. (1943) "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133.

Mendelsohn, L. (1993) *Technical Analysis of Stocks & Commodities: Preprocessing Data For Neural Networks.* URL: http://www.day-trading-commodities.com/preprocessing_data.asp. Accessed Date: 3 September 2003.

Michalski, R. S., Mozetic, I., Hong, J., and Lavrac, N. (1986) The Multi-Purpose Incremental Learning System AQ15 and its Testing Application to Three Medical Domains. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, 1041-1045, Philadelphia, PA: Morgan Kaufmann.

Minsky, M. and Papert, S., (1969) *Perceptrons: An introduction to Computational Geometry*, MIT Press.

Naeher, L. P., Holford, T. R., Beckett, W. S., Belanger, K., Triche, E. W., Bracken, M. B., and Leaderer, B. P. (1993) Healthy Women's PEF Variations with Ambient Summer Concentrations of $PM_{10}$, $PM_{2.5}$, $SO_4^2$, $H^+$, and $O_3$ . *American Journal Respiratory Critical Care Medicine*, V. 160, Number 1, 117-125.

Nelson, M. M. and Illingworth, W.T. (1991) *A Practical to Guide to Neural Nets,* Addison Wesley, MA.

Nigrin, A., (1993) Neural Network for pattern recognition, Cambridge, MA: The MIT press, p.11.

Nikki, M., and Morgan, N., (1999) Combining Connectionist Multi-Band And Full-Band Probability Streams For Speech Recognition Of Natural Numbers. International Computer Science Institute, University of California.

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)*, Universiti Pertanian Malaysia: Kuala Lumpur, pp. 20 - 23.

Plett, G.L., Takeshi Doi and Don Torrieri, (1996) Present and Future Methods of Mine Detection Using Scattering Parameters and An Artificial Neural Networks, *Proceeding SPIE vol. 2765, Detection and Remediation Technologies for Mines and Minelike Targets,* April 1996.

Pranckeviciene, E. (1999) Finding Similarities Between An Activity of the Different EEG's by means of a Single layer Perceptron. *Machine Learning and Applications: MachineLearning in Medical Applications*. Chania, Greece, pp. 49-52.

Rashid, R., Jamaluddin, H. and Saidina Amin, A. (2003) Application of Multi-Layer Perceptron in Modeling Tapioca Starch Hydrolysis. In Artificial Intelligence Applications in Industries, June 24-25, 2003.

Rosenblatt, F., (1958) The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, *Psychological Review*, v65, No. 6, pp. 386-408.

Rumelhart, D.E., Hinton, G.E., and Williams, R. J., (1986) Learning Internal Representations by Error Propagation, *in Rumelhart, D. E. and McClelland, J. L. (editor), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Cambridge, MA, and London, England.*

Sarle, W. S. (1999) Neural Network FAQ, part 1 of 7: Introduction. *Periodic posting to the Usenet Newsgroup comp.ai.neural-nets*, ftp://ftp.sas.com/pub/neurl/FAQ.html Accessed date: 21 May 2002.

Schalkoff, R. J. (1997) *Artificial Neural Network*, McGraw Hill, New York.

Shanker, M. S., (1996) Using Neural Networks to Predict the Onset of Diabetes Mellitus, Department of Administrative Sciences, Kent State University, Kent, OH 44242, In Journal of Chemical Information and Computer Sciences, 36.

Siraj, F., Zakaria, A., Aziz, A. and Abas, Z., (2003) A Web Based Business Insolvency Classifier using Neural Network. *Proceeding of AIAI 2003.*

Stergiou, C., and Siganos, D., (1996) Neural Networks, Volume 4.

Stone, J.V. and Thorton, C.J. (1995) Can Artificial Neural Networks Discover Useful Regularities?. In Artificial Neural Networks, Conference Publication No. 409 IEE. Pages 201-205. 26-28 June 1995.

Street, W. N., Mangasarian, O. L., and Wolberg, W. H. (1996) Individual and Collective Prognostic Prediction. *Thirteenth International Conference on Machine Learning.*

Tan, M., and Eshelman, L. (1988) Using weighted networks to represent classification knowledge in noisy domains. *Proceedings of the Fifth International Conference on Machine Learning*, 121-134, Ann Arbor, MI.

Tsoukalas, L. H., and Uhrig R. E., (1997) *Fuzzy and Neural Approaches in Engineering,* John Wiley & Sons, Inc.

Turban, E., (1993) *Decision Support and Expert Systems: Management Support Systems*, Macmillan Publishing Company.

Varslot, T. (1997) *Artificial Neural Network and Breast Cancer.* URL: http://www.varslot.net/trond-karsten/dagligliv/publikasjoner/prosjektii/prosjekt.html Accessed date: 18 March 2003.

Wang, H., Chen, H., Wang, Y. and Sun, W., (2001) Constructive Competitive Neural Network for Associative Memories, *Department of Computer Science and Engineering, Fudan University, Shanghai China.*

Wolberg, W. H., and Mangasarian, O. L., (1990) Multisurface method of pattern separation for medical diagnosis applied to breast cytology. *Proceedings of the National Academy of Sciences, U.S.A.*, Vol. 87, December 1990, pp 9193-9196.

Wolfram, S. (2003) Data Pre-Processing: Neural Network Documentation. URL: http://documents.wolfram.com/applications/neuralnetworks/NeuralNetworkTheory/2.2.0.html Accessed date: 24 March 2003.

Yoon, Y., & G. Swales. (1991). Predicting Stock Price Performance: A Neural Network Approach. *Proceedings of the 24th Annual Hawaii International Conference on Systems Sciences*. IEEE Computer Society Press: *vol. 4*, pp. 156-62.

Zhou, X. *Data Mining Semester Project Progress Report*. (2003) URL: http://home.olemiss.edu/~xzhou/progress.html Accessed date: 1 November 2003.

Zurada, J. M., (1992) Introduction To Artificial Neural Systems, Boston: PWS Publishing Company , p. XV.