MULTILEVEL LEARNING IN KOHONEN SOM NETWORK
FOR CLASSIFICATION PROBLEMS

NORFADZILA BINTI MOHD YUSOF

UNIVERSITI TEKNOLOGI MALAYSIA

MULTILEVEL LEARNING IN KOHONEN SOM NETWORK FOR
CLASSIFICATION PROBLEMS

NORFADZILA BINTI MOHD YUSOF

A project report submitted in partial fulfillment
of the requirements for the award of the
degree of Master of Science (Computer Science)

Faculty of Computer Science and Information System
Universiti Teknologi Malaysia

JUNE 2006

To my beloved family…

ACKNOWLEDGEMENTS

Although the writing of a project report is ultimately falls to the hand of single student, it is by no means a solitary act. During the writing of this project report, I have been fortunate to have the counsel and friendship of a great many people. This project report would have come to fruition were it not for these individuals.

Firstly, I would like to thank my supervisor, Assoc. Prof. Dr. Siti Mariyam Shamsudin for spending his valuable time to give me many helpful suggestions and encouragements. I would also like to thank Dr. Ali Selamat, Dr. Siti Zaiton and Assoc. Prof. Abd. Manan Ahmad for providing me with crucial advice and information.

Finally my family and friends are also deserving of my thanks. My parents and my entire family members were instrumental in establishing the love of learning which has culminated in this thesis; for this, plus their emotional and financial support over the years. I shall be eternally grateful. A special thank to my beloved twin sister, Norfadzlia binti Mohd Yusof that always being there for me, give me support and encouragement to do the best. Lastly I would to dedicate this project report to the memory of my beloved grandfather that passed away on 15 May 2006, as I was nearing completion of this thesis. Although I never met him, spoke with him or directly sought his advice again, he is singularly responsible for my love of readings. May his soul rest in peace.

ABSTRAK

Pengelasan merupakan satu salah satu bidang kajian dan aplikasi rangkaian neural yang giat dijalankan. Peta swa-organisasi (PSO) ialah rangkaian neural yang mengaplikasikan pembelajaran tanpa seliaan telah membuktikan kemampuannya dalam menyelesaikan masalah pengelasan dan pengecaman pola. PSO tidak memerlukan sebarang pengetahuan mengenai corak taburan pola seperti kaedah-kaedah statistik yang sedia ada. Di dalam kajian ini, kaedah pembelajaran multiaras telah dicadangkan untuk diimplentasikan ke atas rangkaian neural PSO. Keupayaan dan keberkesanan kaedah ini dalam menyelesaikan masalah berkaitan pengelasan pola dianalisa. Keadah pembelajaran PSO yang dicadangkan dan kaedah pembelajaran PSO piawai dianalisa dengan menggunakan beberapa jenis sukatan jarak atau ketakserupaan yang digunakan bagi mengukur keserupaan antara pola. Penilaian dibuat terhadap kualiti maklumat yang dipersembahkan di atas peta output yang dihasilkan melalui proses pembelajaran menggunakan beberapa jenis sukatan ketidakserupaan ini. Hasil yang diperolehi melalui kedua-dua kaedah pembelajaran ini digunakan untuk membuat peramalan dan pengelasan ke atas sampel pola yang baru. Eksperimen ini dijalankan bertujuan untuk membuat perbandingan terhadap keupayaan algoritma PSO menggunakan kaedah pembelajaran multiaras dengan pembelajaran piawai. Keberkesanan kedua-dua kaedah ini dapat dibuktikan dengan mengimplementasikannya ke atas lima set data. Hasil kajian menunjukkan bahawa kaedah yang dicadangkan berupaya menjadi rangka alternatif bagi masalah pengelasan data. Ini adalah ekoran daripada keupayaannya memberi persembahan yang baik dari aspek pengelasan data dan mengurangkan masa pemprosesan berbanding pembelajaran PSO piawai terutamanya bagi data yang bersaiz kecil dan sedarhana. Walaupun begitu, bagi masalah pengelasan yang melibatkan data yang bersaiz besar, ia masih didominasi oleh kaedah pembelajaran PSO piawai.

ABSTRACT

Classification is one of the most active research and application areas of neural networks. Self-organizing map (SOM) is a feed-forward neural network approach that uses an unsupervised learning algorithm has shown a particular ability for solving the problem of classification in pattern recognition. Classification is the procedure of recognizing classes of patterns that occur in the environment and assigning each pattern to its relevant class. Unlike classical statistical methods, SOM does not require any preventive knowledge about the statistical distribution of the patterns in the environment. In this study, an alternative classification of self organizing neural networks, known as multilevel learning, is proposed to solve the task of pattern separation. The performance of standard SOM and multilevel SOM are evaluated with different distance or dissimilarity measures in retrieving similarity between patterns. The purpose of this analysis is to evaluate the quality of map produced by SOM learning using different distance measures in representing a given dataset. Based on the results obtained from both SOM learning methods, predictions can be made for the unknown samples. This study aims to investigate the performance of standard SOM and multilevel SOM as supervised pattern recognition method. The multilevel SOM resembles the self-organizing map (SOM) but it has several advantages over the standard SOM. Experiments present a comparison between a standard SOM and multilevel SOM for classification of pattern for five different datasets. The results show that the multilevel SOM learning gives good classification rate, however the computational times is increased compared over the standard SOM especially for medium and large scale dataset.