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.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$x_i(k)$	-	Input to node *i* at time *k*
$w_{ij}(k)$	-	Weight from input node *i* to neuron *j* at time *k*
x_i	-	Input to the node *i*
x_j	-	Output node *j*
k	-	Input dimension
d_{ij}	-	Distance between input node *i* and output node *j*
t	-	Number of iteration
c	-	Number of cycle
$\alpha(t)$	-	Learning rate
$h(c,r)$	-	Neighborhood function
r	-	Neighborhood radius
t	-	Number of iteration
c	-	Number of cycle
N	-	Number of input nodes
E_q	-	Quantization Error
$w(t)$	-	weight vector
$w(t+1)$	-	updated weight vector
$d^2_{j,c}$	-	lateral distance of the excited neurons j and the winning neuron c
σ	-	measures the degree to which excited neurons in the vicinity of the winning neuron cooperate in the learning process
X_n	-	New *x* value (after normalization)
X_0	-	Current value of *x* (before normalization)
X_{min}	-	Minimum value of *x* in the sample data

X_{max}	-	Maximum value of x in the sample data
m_c	-	Best matching reference vector
$\lVert x_i - m_c \rVert$	-	Difference (distance) between each data vector and its best matching reference vector
τ_1 τ_1	-	Slope of the graph of $\sigma(t)$ against t

LIST OF ABBREVATIONS

SOM	-	Self-organizing map
SOFM	-	Self-organizing feature map
AQE	-	Average quantization error
BMU	-	Best matching node (neuron)
BP	-	Back-propogation
MSOM	-	Multilevel SOM learning
CT	-	Computation times
A	-	Classification accuracy
P	-	Precision
R	-	Recall

LIST OF TERMINOLOGIES

Back-propogation (BP)	-	A supervised learning method in which an output error signal is feed back through the network, altering connection weights so as to minimize the error.
Categorization		A process in which ideas and objects are recognized, differentiated and understood where it implies that objects are grouped into categories, usually for some specific purpose. Ideally, a category illuminates a relationship between the subjects and objects of knowledge.
Connection	-	A link between nodes used to pass data from one node to the other. Each connection has an adjustable value call the weights.
Feature extraction	-	A special form of dimensionality reduction and is in the area of image processing also connected with shape recognition.
Generalization	-	A neural network ability to respond correctly to data no used to train it.
Input layer	-	A layer of nodes that forms a passive conduit for data entering a neural network.
Labeled data	-	Input pattern tagged with a target result, which provides the correct answer needed by supervised algorithms for training.
Neural Network	-	An implementation of a teaming algorithm derived from research about the brain. Often referred to as

		artificial neural network, it typically contains layers of so called artificial neurons composed of weights, connections and nodes.
Node	-	A single neuron-like element in a neural network. It typically has many inputs but only one output.
Output layer	-	The layer of nodes that produce neural network results.
Pattern recognition	-	Identification of shapes, forms, or configurations by automatic means.
Supervised learning	-	A learning process requiring a labeled training set.
Self-Organizing Maps (SOM)	-	A subtype of artificial neural networks and it is trained using unsupervised learning to produce low dimensional representation of the training samples while preserving the topological properties of the input space.
Target output	-	A correct results included with each input pattern in a training and testing set.
Testing	-	A process of measuring a neural network's performance, during which the network passes through an independent dataset to calculate a performance index, it does not change its weights.
Training	-	A process during which a neural network passes through a dataset repeatedly, changing the values of its weights to improve its performance.
Unsupervised learning	-	A learning process that does not require target result.
Weight	-	An adjustable value associated with a connection between nodes in a neural network.

CHAPTER 1

INTRODUCTION

We are living in a world full of data. Every day, people encounter a large amount of information and store or represent it as data, for further analysis and management. One of the vital means in dealing with these data is to classify or group them into a set of categories or clusters. Actually, as one of the most primitive activities of human beings, classification plays an important and indispensable role in the long history of human development. In order to learn a new object or understand a new phenomenon, people always try to seek the features that can describe it, and further compare it with other known objects or phenomena, based on the similarity or dissimilarity, generalized as proximity, according to some certain standards or rules.

Artificial neural networks (ANNs) are simple computational tools for examining data and developing models that help to identify interesting patterns or structures in the data. The data used to develop these models is known as training data. Once neural network has been exposed to the training data, and has learnt the patterns that exist in the data, it can be applied to new data thereby achieving variety outcomes. Neural networks can be used to

- Learn to predict future events based on the patterns that have been observed in the historical training data.

- Learn to classify unseen data into pre-defined groups based on characteristics observed in the training data.
- Learn to cluster the training data into natural groups based on the similarity of characteristics in the training data.

Recent research activities in ANN also have shown that ANN have powerful classification (Dorothea Heiss, C. and Bajla I., 2005) and pattern recognition (Xin-Hua, S. and Hopke, P.K., 1996) capabilities. Inspired by biological system, ANN is able to learn from and generalized from experienced. ANN explore many competing hypotheses simultaneously using massively parallel network composed of non linear relatively computational elements interconnect by links with variable weights. It is this interconnected set of weights that contains the knowledge generated by the ANN (Adya, M. and Collopy, F., 1998).

ANNs can be divided into two learning categories: supervised and unsupervised (Smith, K. A., 2002). In unsupervised learning, a desired output result for each input vector is required when the network is trained. An ANN of the supervised learning type, such as the multi-layer perceptron (MLP), uses the target result to guide the formation of the neural parameters. It is thus possible to make the neural network learn the behavior of the process under study. In contrast with unsupervised learning, the training of the network is entirely data driven, and no target results for the input data vectors are provided. An ANN of unsupervised learning type, such as the self-organizing maps (SOM), can be used for clustering the input data and find features inherent to the problem.

Basically, classification systems are either supervised or unsupervised, depending on whether they assign new inputs to one of a finite number of discrete supervised classes or unsupervised categories. Hence, the context of this study is limited to the evaluation of SOM algorithm performance in classification task.

1.1 Problem Background

Kohonen SOM networks have been successfully applied as a classification tool to various problem domains. The self-organizing map (SOM) network is a special type of neural network that can learn from complex, multi-dimensional data and transform them into visually decipherable clusters. The theory of the SOM network is motivated by the observation of the brain operation. Various human sensory impressions are neurologically mapped into the brain such that spatial or other relations among stimuli correspond to spatial relations among the neurons are organized into a two-dimensional map. The main function of SOM networks is to map the input data from an n-dimensional space to a lower dimensional (usually one or two-dimensional) plot while maintaining the original topological relations. The physical location of points on the map shows the relative similarity between the points in the multi-dimensional space.

Self organizing maps (SOMS) are a form of competitive neural network (Kohonen, T., 1998), which transforms highly dimensional data onto a two dimensional grid, while keeping the data topology by mapping similar data items to the same cell on the grid (or to neighboring cells), using some form of distance measure usually Euclidean distance.

In other neural network models, all neurons adjust their weights in response to a training presentation while in competitive learning only one or few neurons are allowed to adjust their weights. Therefore, this ability made Kohonen networks to become more resource efficient compared to other networks. Moreover, the unsupervised training of Kohonen network does not require target output for training. The network is able to learn the pattern of data itself without knowing all the output. The nodes in the network converge to form clusters to represent groups of entities with similar properties. The number and composition of clusters can be visually determined based on the output distribution generated by the training process.

Besides unsupervised training, SOM is able to train in supervised manner (Lee T. E., 2005). This method is normally applied if the target outputs have been known in priori. The flexibility and ability of SOM has gained interest of the author to further research and apply the technique in variety of tasks such as classification.

ANN implementations that based on competition method often use some means of calculating distance between input vectors and weights (Gopalan, A. and Titus, A. H., 2003). Clearly, an important part of this process is the comparison of the input vector elements and weight vector elements. Mathematically, this comparison is achieved through the computation of a distance between vectors; vectors with the smallest distance are most similar. The goal is to minimize the distance between the stored weight vectors and the input vectors. Term distance is also used to convey the idea of dissimilarity. Naturally, this distance should only be applicable to real-valued patterns (Lourenço, F *et al.*, 2004). None of the distance measure, including Euclidean appropriately handle non-continuous input attributes.

Although the term similarity is often used, dissimilarity corresponds to the notion of distance, small distance means small dissimilarity, and large similarity (Veltkamp, R. C., 2001). So when comparing patterns, it is very useful if they are represented in a space that has a metric. The success of unsupervised algorithms, such as the SOM and clustering methods, depends crucially on the metric, the measure of the distance between the objects of interest. The metrics, on the other hand, depends on which kinds of variables selection and feature extraction (Kaski, S. *et al.*, 2001)

The choice of metric for neural network that implements competitive learning rule such as SOM is directly connected to the representation of data and it crucially influences the efficiency, accuracy and generalization ability of the results.

They are various types of distance measures which all define different kinds of metric space. Each method has its own properties and generally gives different perspectives of the data turning the matter of choice not trivial (Meyer, 2002). The most commonly used methods for calculating distance in SOM learning is Euclidean distance measure that considers each observation dimension with the same significance whatever the observation distribution inside classes (Fessant, F *et al.* 2001).

Among all distance measures, some have very similar behaviors in similarity queries. others may behave quite differently (Qian, G. *et al.*, 2003). For example, Bray Curtis distance and Canberra distance have favorable advantage where both measures perform their own standardization. Usually the method is chosen based on which distance measure that gives the 'best' results in terms of some error function or ability to classify/cluster certain data points. Changing the distance measure can have a major effect on the overall performance of a classification system.

One way of comparing distance measures is to study their retrieval performance on a particular application (Qian, G. et al., 2002). Choosing a particular distance measures also concern on the impact of computational overhead on system performance (Qian, G. *et al.*, 2003). Understanding the relationship among distance measures is helpful in choosing a proper one for a particular application.

Fessant, F *et al.* (2001) compare the performance of supervised self-organizing maps designed with different distance measures: Euclidean distance and Mahalanobis distance on data classification application. Concerning on classification problems, Mahalanobis distance turns out to be more effective concerning classification problem with 92.8% classification accuracy compared to Euclidean distance with 94.1%. This is because of the large range of data components variations. In fact, the giving up of Euclidean distance is advisable when the variances of input vectors components are highly different.

In Cure, J. D. and Hill, J. J. (1981) paper, proposed a scheme where both the Euclidean distance measure and a simpler unweighted city-block distance are utilized together for improving the classification speed of clustering algorithm which used a Euclidean distance metric. The proposed scheme described, allow the algorithm to decide whether the classification of each pattern vectors is to be achieved by the computationally slow Euclidean distance or the faster city-block distance.

Keeratipranon, N. and Maire, F. (2005) also highlighted the differences between three natural similarity measures for bearing vectors. The researches has demonstrated the clear superiority of the Mahalanobis distance for localization based on bearings problems with reaches the best classification accuracy of 99.32% compared to Euclidean distance achieves 92.17% classification accuracy and Naive Bayes distance with 97.14%.

Huang, Y. *et al*. (1998), evaluate the performance of the self-organizing maps (SOMs) with different distance measures; Euclidean distance and Bhattacharyya distance in retrieving similar images when a full or a partial query image is presented to the SOM. The results show that the Bhattacharyya distance is superior to Euclidean distance with 98% of retrieval rates compared to Euclidean distance which is 95%. The standard Euclidean distance not yield the best results in retrieving partial images based on their histograms due to long time needed to compute color histograms compared to Bhattacharyya distance.

The SOM as conceived should live the input patterns space, i.e. the codebook patterns should lie in the space of the input patterns. The original SOM algorithm was defined for real valued patterns. However, when using binary input patterns and as consequence of computation, the codebook patterns will assume non-binary (i.e., real) values. To keep applying the binary similarity measures we have to envisage some way to convert the real valued pattern to a binary one in order to compute the best matching unit (BMU). However, for binary data the usual Euclidean distance can be replaced by binary similarity measures that take into account possible asymmetries and therefore provide a different point of view for looking at the data.

Fernando, L. *et al.*, (2002) in their study, they had proposed two SOM architecture that approaches to the BMU problem, the “hard” logic and the dot product. When using “hard” logic and the dot product approach, BMU can be computed using other types binary similarity measures instead of Euclidean distance especially when dealing with binary patterns. In the context of SOM it is clear that the range of variation allowed by Euclidean distance cannot be matched by binary-based measures. This means that at this time it is not realistic to use binary-based similarity measures to produce “fine-resolution” clustering, although most of the measures used revealed the ability to distinguish major clusters. In their work, they have showed that binary-based similarity measures might provide a different insight into data, effectively revealing interesting patterns and relations in the data.

Based on the previous research, has gained interest for the accomplishment of this study in order to evaluate and compare the performance of SOM using different distance measures in classification tasks. From these past researches also shows that, there is no highly difference in the performance of SOM in terms of its classification accuracy when different distance measures is employ to this algorithm. For this reason, a new learning methodology is developed to be implemented in SOM algorithm to see whether it able to enhanced the performance of SOM in classification tasks. This new proposed method is known as multilevel SOM learning.

By using this approach, original SOM algorithm is divided to two learning level, where each level will implement different distance measures during the learning process instead of one measures as in original SOM algorithm. Hence, in this study, five types of distance such as Euclidean distance, Manhattan distance, Bray Curtis distance, Canberra distance and Chebyshev distance was evaluated and their performance in SOM learning process was investigated. The new proposed multilevel learning method is then analyzed to find out whether it can improve the performance of SOM in pattern classification task. The performance of new proposed SOM-based classification system is evaluated in terms of classification accuracy and computation times.

1.2 Problem Statements

The choice of metric for neural network that implements competitive learning rule such as SOM is directly connected to the representation of data and it crucially influences the efficiency, accuracy and generalization ability of the results. Euclidean distance is commonly used metric in SOM application. Besides Euclidean distance, there are different types of distance measures; which all define different kinds of metric space. From previous studies shows that, by employing different distance measures in SOM has affect the performance of this network in classification context. So, this study attempt to evaluate the performance of SOM using different distance measures in several real world classification problems. The hypothesis of the study can be stated as:

"*Could the selection of distance measures used to train the SOM can affect the performance of SOM?*"

Based on past researches (Huang, Y. *et al.*, 1998), (Fessant, F *et al.*, 2001) and (Keeratipranon, N. and Maire, F., 2005), shows that although the used of different distance measures has affect the SOM performance in classification tasks but the results obtained is nearly equivalent and not quite promising. For this reason, an enhancement learning methodology for SOM algorithm is proposed that is known as multilevel SOM learning in order to find out whether it can give better improvement on SOM classification results. The hypothesis for this study can be stated as:

"Could multilevel learning approach used in Kohonen Self-Organizing Maps (SOM) neural network enhanced the accuracy of classification result?"

1.3 Project Aim

The aim of this study is to apply multilevel learning approach in Self-Organizing Map (SOM) algorithm. This approach is evaluated and analyzed to determine weather it can improve SOM learning performance in terms of its capability to produce the accurate classification result in less computation times. Further more, different types of real-valued dataset are used to represent the classification problem that is going to be solved using SOM algorithm designed with multilevel learning approach.

1.4 Objectives Of The Project

The objectives of the study are outlined as below:

1. To propose multilevel learning methodology in SOM algorithm known as multilevel SOM.

2. To design and develop standard SOM and multilevel SOM model which uses various distance measures.

3. To evaluate and compare the learning and classification performance of standard SOM models and multilevel SOM models.

1.5 Project Scopes

Below defined the scope of the study, which involved several areas:

1. Five types of distance measures are employed in SOM algorithm.
   The distance measures are Euclidean Distance, Manhattan Distance, Bray Curtis distance, Canberra distance and Chebyshev distance.

2. These algorithms are tested using real-valued data set.
   Four set of universal data being used are Iris, Wine, Glass, Diabetes and Pendigits.

3. The programs are built on a Windows environment using Microsoft Visual C++ 6.0 programming language.

1.6 Significance Of The Project

The study investigates the capabilities of multilevel learning method used in Self-Organizing Maps (SOM) to perform in pattern classification tasks. The performance of standard SOM and multilevel SOM trained using various distance measures such as Euclidean distance, Manhattan distance, Bray Curtis distance, Canberra distance and Chebyshev distance are evaluated and compared. The performance of SOM which employ multilevel learning approach is evaluate to examine whether this new proposed method is able to give better performance than the standard SOM in terms of classification accuracy and computation time.

1.7 Organization Of The Report

This report consists of four chapters. Chapter 1 presents the introduction of the study. The remainders of this report are structured as follows. Chapter 2 covers the literature review of this project, which is divided into 4 parts. The first part, recall the basic concepts of SOM network, mainly focused on the architecture and training particular processes in Kohonen Self-Organizing Maps (SOM). Next, four types of distance measures will be described and their performance differences also discussed. A review on relevant and related literature on classification using SOM algorithm will be presented. Chapter 3 provides the methodologies in terms of data and classification techniques used in this study. Chapter 4 presents the experimental result of this project, where the results shows the performance of standard SOM and multilevel SOM trained using Euclidean distance, Manhattan distance, Bray Curtis distance, Canberra distance and Chebyshev distance when it tested using the dataset of real world classification problems, which is taken from universal data. Finally, suggestion of future research direction and the conclusion for this study are given in Chapter 5.

REFERENCES

Adya, M., Collopy, F. (1998). How Effective Neural Networks at Forecasting and Prediction? A Review and Evaluation. *Journal Of Forecasting,* vol. 17, pp. 481-495.

Alam, P., Booth, D., Lee, K., and Thordarson, T. (2000). The Use Of Fuzzy Clustering Algorithm And Self-Organizing Neural Network For Identifying Potentially Failing Banks: An Experimental Study. *Expert System With Applications,* vol. 18, pp. 185-199.

Amato, N. M. *et al.*, 2000). Choosing Good Distance Metrics and Local Planners for Probabilistic Roadmap Methods. *IEEE Transactions On Robotics And Automation,* vol. 16, no. 4.

Barthelemy, S. and Filippi, J. B. (2003). A Typology Of Very Small Companies Using Self-Organizing Maps. *IEEE.*

Blazejewski, A. and Coggins, R. (2004). Application Of Self-Organizing Maps To Clustering Of High-Frequency Financial Data. *(AWDMandWI2004), Dunedin, New Zealand. Conferences in Research and Practice in Information Technology,* vol. 32.

Cure, J. D. and Hill, J. (1981). A Method for Improving the Classification Speed of Clustering Algorithms Which Use a Euclidean Distance Metric. *Proceedings of the IEEE,* vol. *69,* no. 1.

Cottrell, M., Girard, B., and Rousset, P. (1998c). Forecasting Of Curves Using A Kohonen Classification. *Journal of Forecasting*, vol. 17, pp. 429–439.

Deboeck, G. and Kohonen, T. (1998). Visual Explorations in Finance with Self-Organizing Maps. London: Spinger-Verlag.

Fessant, F., Aknin, P., Oukhellou, L. and Midenet, S. (2001). Comparison Of Supervised Self-Organizing Maps Using Euclidean or Mahalanobis Distance in Classification Context. *6th IWANN2001, Granada*

Gopalan, A. and Titus, A. H. (2003). A New Wide Range Euclidean Distance Circuit for Neural Network Hardware Implementations. *IEEE Transactions On Neural Networks,* vol. 14, no. 5.

Hagenbuchner, M., Sperduti, A. and Tsoi, A. C. (2003). A Self-Organizing Map For Adaptive Processing Of Structured Data. *IEEE Transactions On Neural Networks,* vol. 14, no. 3.

Heiss, D. C. and Bajla, I. (2005). Using Self-Organizing Maps For Object Classification In Epo Image Analysis. *Measurement Science Review,* vol. 5, sec. 2.

Herrero, J., Valencia, A. and Dopazo, J. (2000). A Hierarchical Unsupervised Growing Neural Network For Clustering Gene Expression Patterns.

Hinton, G. and Sejnowski, T. (1999). Unsupervised Learning: Foundations of Neural Computation. Cambridge, MA: The MIT Press.

Honkela T. (1998). Description of Kohonen's Self-Organizing Map.

Honkela T. (1999). Learning to Understand -General Aspects of Using Self-Organizing Maps in Natural Language Processing

Huang, Y., Suganthan, P. N., Krishnan, S. M. and Xiang, C. (2001). Evaluation of the Distance Measures for Partial Retrieval Using Self-Organizing Map. *ICANN 2001,LNCS 2130*, pp 1042-1047.

Hung C. and Wermter S. (2002), A Dynamic Adaptive Self-Organising Hybrid Model for Text Clustering.

Ichiki, H., Hagiwara M. and Nakagawa M. (1993). Kohonen Feature Maps as a Supervised Learning Machine. *Proceedings of the IJCNN-IEEE 1993 Conference, S. Francisco CA,* pp. 1944-1948.

Iordanova, I., Rialle, V. and Vila, A. (1992) Use of Unsupervised Neural Networks for Classification Tasks in Electromyography.

James, S. K. and Jacek M. Z. (2004). Topography-Enhanced BMU Search in Self-Organizing Maps. *ISNN 2004, LNCS 3174,* pp. 695–700, 2004.

Ji, C. Y. (2000). Land-Use Classification Of Remotely Sensed Data Using Kohonen Selforganizing Feature Map Neural Network. *Photogrammetric Engineering and RemoteSensing*, 66, pp. 1451-1460.

Kaski, S., Sinkkonen, J., and Peltonen, J. (2001). Bankruptcy Analysis with Self-Organizing Maps in Learning Metrics. *IEEE Transactions on Neural Networks*, vol. 12, pp. 936-947.

Keeratipranon, N. and Maire, F. (2005). Bearing Similarity Measures for Self-Organizing Feature Maps. Springer Verlag Berlin Heidelberg.

Keith-Magee, R. (2001). Learning and Development In Kohonen Style Self-Organizing Maps. *Thesis for Degree Of Doctor of Philosophy. Curtin University of Technology*

Ketelare, D. B., Moshou D., Coucke P. and Baerdemaeker J. D. (1994). A Hierarchical Self-Organizing Map For Classification Problems

Kiang, M. Y. (2001). Extending The Kohonen Self-Organizing Map Networks For Clustering Analysis. *Computational Statistics and Data Analysis*, vol. 38, pp. 161–180.

Kohonen, T. (1982). Self-organized Formation Of Topologically Correct Feature Maps. *Biological Cybernatics*, vol. 43, pp. 59-69.

Kohonen T., Hynninen J., Kangas J., and Laaksonen J. (1996). SOM K: The Self-Organizing Map Program Package (Report A31), *Helsinki University of Technology, Laboratory of Computer And Information Science.*

Kohonen, T. (1998). The Self-Organizing Maps. Springler, Berlin, Heidelberg.

Kohonen T, Kaski S., Lagus K., Salojärvi J., Honkela J., Paatero V. and Saarela A. (2000), Self Organization Of A Massive Document Collection, *IEEE Transactions on Neural Networks,* vol. 11, no. 3, pp. 574-585.

Koua, E. L. (2003). Using Self-Organizing Maps For Information Visualization And Knowledge Discovery In Complex Geospatial Datasets. *Proceedings of the 21st International Cartographic Conference (ICC) Durban, South Africa, 10-16 August 2003*

Lamirel, J.C., Al Shenabi, S., Hoffman M. and Francois, C. (2003). Intelligent Patent Analysis Through The Use Of A Neural Network: Experiment Of Multi-Viewpoint Analysis With The Multisom Model, *Proceedings of the ACL-2003 Workshop on Patent Corpus Processing, pp. 7-23.*

Lampinen, J. (1992). On Clustering Properties Of Hierarchical Self-Organizing maps, *in Aleksander, I. and Taylor, J. eds, 'Artificial Neural Networks, 2', Vol. II, North-Holland, Amsterdam, Netherlands*, pp. 1219–1222.

Lee, K., Booth D. and Alam P. (2005). A Comparison Of Supervised And Unsupervised Neural Networks In Predicting Bankruptcy Of Korean Firms. *Expert Systems With Applications*, vol. 29, pp. 1-16.

Lee, T. E. (2005). Stock Market Prediction Using Kohonen Network. *Ijazah Sarjana Muda Sains Komputer, Universiti Teknologi Malaysia.*

Lourenço, F., Lobo, V., and Bação, F. (2004). Binary-Based Similarity Measures For Categorical Data And Their Application In Self-Organizing Maps.

Mangiameli, C., Chen, S. K. and West, D. (1996). A Comparison of SOM Neural Network and Hierarchichal Clustering Methods. *European Journal of Operation Research*, vol. 93, no. 2.

Martin, L., Carsten, P., Guido, B., Lars, E., and Leif, S. (2000). Clustering ECG Complexes Using Hermite Functions and Self-Organizing Maps. *IEEE Transactions On Biomedical Engineering,* vol. 47, no. 7.

Masters, T. (1993). Practical Neural Network Recipes In C++. New York, NY: Academic Press, pp. 266-267, 328-341, 411.

Mu-Chun, S., Ta-Kang, L. and Hsiao-Te, C. (2002). Improving the Self-Organizing Feature Map Algorithm Using an Efficient Initialization Scheme *Tamkang Journal of Science and Engineering,* Vol. 5, No. 1, pp. 35-48

Nakkrasae. S., Sophatsathit, P., and Edward William, R. J. (2004). Fuzzy Subtractive Clustering Based Indexing Approach For Software Components Classification. International Journal of Computer and Information Science, Vol. 5, No. 1.

Negnevitsy, M. (2002). Artificial Intelligence: A Guide To Intelligent System. Pearson Education Limited, England.

Pakkanen, J. and Jukka, I., 2003. A Novel Self-Organizing Neural Network For Defect Image Classification.

Papadimitriou, S., Mavroudi S, Vladutu L.. Pavlides, G., and Bezerianos, A.

(2002). The Supervised Network Self-Organizing Map For Classification Of Large Data Sets. *Applied Intelligence. Kluwer Academic Publishers. Manufactured in The Netherlands.* vol. 16, pp. 185–203.

Qian, G., Surat, S. and Pramanik, S. (2002). Similarity Between Euclidean And Consine Angle Distance For Nearest Neighbor Queries. *IEEE ICIP*.

Rui Xu, (2005). Survey of Clustering Algorithms. *IEEE Transactions On Neural Networks*, Vol. 16, No. 3

Smith, K. A. (2002). Neural Networks: An Introduction. *Neural Network For Business.*

Tsao, E. C., Bezdek, J. C. and Pal, N. R. (1994). Fuzzy Kohonen Clustering Network. *Pattern Recognition,* Vol. 27, pp. 757-764

Tso, B. and Mather, P. M, (2001). Classification Methods for Remotely Sensed Data, New York: Taylor and Francis

Ultsch Guimarães (1993). Knowledge Extraction from Artificial Neural Networks and Applications. *Transputer-Anwender-Treffen, Aachen, Springer Verlag.*

Veltkamp, R. C. (2001). Shape Matching: Similarity Measures and Algorithms.

Vesanto, J. and Alhoniemi, E. (2000). Clustering of the Self-Organizing Map. *IEEE Transactions On Neural Network*, vol. 11, no. 3.

Vesanto, J., Himberg, J., Alhoniemi, E. and ParhanKangas. J. (2000). SOM Toolbox For Matlab.

Vesanto, J. and Sulkave, M. (2002). Distance Matrix Based Clustering of the Self-Organizing Map. *ICANN 2002, LNCS 2415,* pp. 951–956

Villman, T., Merenyi,, E. and Hammer, B. (2003). Neural Maps In Remote Sensing Analysis. *Neural Networks*, vol. 16, pp. 389-403.

Xin-Hua, S. and Hopke, P. K. (1996). Kohonen Neural Network As A Pattern Recognition Method Based On The Weight Interpretation. *Analytica Chimica Acta*, vol. 336, pp. 57-66.

Ying, H., Tian-Jin F., Jun-Kuo C., Xiang-Qiang D. and Ying-Hua Z. (2002). Research On Some Problems In Kohonen SOM Algorithm. *Proceedings Of the First Conference On Machine Learning and Cybernatics, Beijing.*

Zhe, L. and Ronald, E. J. (2004). The Nature And Classification Of Unlabelled Neurons In The Use Of Kohonen's Self-Organizing Map For Supervised Classification.

Zorin, A. (2003). Stock Price Prediction: Kohonen Versus Backpropogation. *Proceeding of International Conference on Modelling and Simulation of Business System – MOSIBUS' 2003, Vilnus, Lithuania.*