School of Information Technology

Neural Network Classification Based On Quantification of Uncertainty

Pawalai Kraipeerapun

This thesis is presented for the Degree of

Doctor of Philosophy of

Murdoch University

March, 2009

Declaration

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

Pawalai Kraipeerapun

March, 2009

Acknowledgements

I would like to take this opportunity to acknowledge and thank the following people and organizations who helped me to complete this thesis.

I would like to thank my supervisor, Associate Professor Dr. Lance Chun Che Fung for his guidance and encouragement throughout the period of my Ph.D. research. I would also like to thank my associate supervisor, Associate Professor Dr. Kevin Kok Wai Wong for his suggestion. I wish to express my gratefulness to Dr. Warick Brown for his helpful advice and his efforts to explain things to me. I also wish to thank Dr. Arta Dilo for her support in order to start my Ph.D. thesis. Many thanks to my Ph.D. colleagues and Thai students who have given me nice friendships. Thanks to all of them.

I am very grateful to Royal Thai Government for the financial support during the period of my study. I also wish to thank Ramkhamhaeng University for giving me the opportunity to study in Perth. I am deeply grateful to the Office of Education Affairs, Royal Thai Embassy in Canberra for all advices during my study. Finally, I would like to thank the School of Information Technology, Murdoch University, for providing me all the necessary facilities for my research.

Abstract

This thesis deals with feedforward backpropagation neural networks and interval neutrosophic sets for the binary and multiclass classification problems. Neural networks are used to predict "true" and "false" output values. These results together with the uncertainty of type error and vagueness occurred in the prediction are then represented in the form of interval neutrosophic sets. Each element in an interval neutrosophic set consists of three membership values: truth, indeterminacy, and false. These three membership values are then used in the classification process. For binary classification, a pair of neural networks is first applied in order to predict the degrees of truth and false membership values. Subsequently, bagging technique is applied to an ensemble of pairs of neural networks in order to improve the performance. For multiclass classification, two basic multiclass classification methods are proposed. A pair of neural networks with multiple outputs and multiple pairs of binary neural network are experimented. A number of aggregation techniques are proposed in this thesis. The difference between each pair of the truth and false membership values determines the vagueness value. Error occurred in the prediction are estimated using an interpolation technique. Both vagueness and error then form the indeterminacy membership. Two and three dimensional visualization of the three membership values are also presented. Ten data sets obtained from UCI machine learning repository are experimented with the proposed approaches. The approaches are also applied to two real world problems: mineral prospectivity prediction and lithofacies classification.

List of Publications Related to This Thesis

Journal

1. P. Kraipeerapun and C. C. Fung, Binary Classification using Ensemble Neural Networks and Interval Neutrosophic Sets, ACCEPTED for publication in the Journal of *Neurocomputing*, Elsevier.

2. P. Kraipeerapun, C. C. Fung and K. W. Wong, Uncertainty Assessment using Neural Networks and Interval Neutrosophic Sets for Multiclass Classification Problems. *WSEAS Transactions on Computers*, Issue 3, Vol. 6, March 2007, pp. 463–470.

3. P. Kraipeerapun, C. C. Fung and K. W. Wong, Lithofacies classification from Well Log Data using Neural Networks, Interval Neutrosophic Sets and Quantification of Uncertainty. *International Journal of Applied Mathematics and Computer Sciences*, Vol. 3, No. 1, 2006, pp. 28–32.

LNCS/LNAI

1. P. Kraipeerapun, C. C. Fung, W. Brown, K. W. Wong and T.D. Gedeon, Uncertainty in Mineral Prospectivity Prediction, *Lecture Notes in Computer Science*,

Springer Verlag, LNCS 4233, 2006, pp. 841–849.

2. P. Kraipeerapun, C. C. Fung and W. Brown, Assessment of Uncertainty in Mineral Prospectivity Prediction Using Interval Neutrosophic Set, *Lecture Notes in Artificial Intelligence*, Springer Verlag, LNAI 3802, 2005, pp. 1074–1079.

Conference Proceedings

1. P. Kraipeerapun and C. C. Fung, Comparing Performance of Interval Neutrosophic Sets and Neural Networks with Support Vector Machines for Binary Classification Problems. In *Proceedings of the Second IEEE International Conference on Digital Ecosystems and Technologies*, Phitsanulok, Thailand, 26-29 February 2008, pp. 34–37.

2. P. Kraipeerapun and C. C. Fung, Uncertainty Visualization in Mineral Prospectivity Prediction. In *Proceedings of the Eighth Postgraduate Electrical Engineering and Computing Symposium (PEECS 2007)*, Perth, Australia, November 2007, pp. 127–129.

3. P. Kraipeerapun, C. C. Fung and K. W. Wong, Ensemble Neural Networks Using Interval Neutrosophic Sets and Bagging. In *Proceedings of the 3rd International Conference on Natural Computation (ICNC'07)*, Haikou, China, 24-27 August 2007, pp. 386–390.

4. P. Kraipeerapun, C. C. Fung and K. W. Wong, Quantification of Vagueness in Multiclass Classification Based On Multiple Binary Neural Networks. In *Proceedings of the international conference on Machine Learning and Cybernetics (ICMLC07)*, Hong Kong, China, 19-22 August 2007, pp. 140–144.

5. P. Kraipeerapun, C. C. Fung and K. W. Wong, Multiclass Classification using Neural Networks and Interval Neutrosophic Sets. In *Proceedings of The 5th WSEAS International Conference on Computational Intelligence, Man-Machine

Systems and Cybernetics (CIMMACS '06), Venice, Italy, 20-22 November 2006, pp. 123–128.

6. P. Kraipeerapun, C. C. Fung and K. W. Wong, Lithofacies Classification from Well Log Data using Interval Neutrosophic Sets. In *Proceedings of the Seventh Postgraduate Electrical Engineering and Computing Symposium (PEECS 2006)*, Perth, Australia, November 2006, pp. 178–180.

7. P. Kraipeerapun, K. W. Wong, C. C. Fung and W. Brown, Quantification of Uncertainty in Mineral Prospectivity Prediction Using Neural Network Ensembles and Interval Neutrosophic Sets, In *Proceedings of the 2006 IEEE World Congress on Computational Intelligence: A Joint Conference of the International Joint Conference on Neural Networks (IJCNN 2006)*, Vancouver, Canada, 16-21 July 2006, pp. 5341–5346.

8. P. Kraipeerapun, C. C. Fung, W. Brown and K. W. Wong, Neural Network Ensembles using Interval Neutrosophic Sets and Bagging for Mineral Prospectivity Prediction and Quantification of Uncertainty, In *Proceedings of the 2006 IEEE International Conferences on Cybernetics and Intelligent Systems*, Bangkok, Thailand, 7-9 June 2006, pp. 388–393.

9. P. Kraipeerapun, C. C. Fung, W. Brown and K. W. Wong, Mineral Prospectivity Prediction using Interval Neutrosophic Sets, In *Proceedings of IASTED International Conference on Artificial Intelligence and Applications*, Innsbruck, Austria, 13-16 February 2006, pp. 235–239.

10. P. Kraipeerapun, C. C. Fung, W. Brown and K. W. Wong, Quantification of uncertainty in the prediction of mineral prospectivity, In *Proceedings of the Sixth Postgraduate Electrical Engineering and Computing Symposium*, Perth, Australia, September 2005, pp. 163–165.

Contributions of This Thesis

In general, binary neural network classification is processed using a single neural network or an ensemble of several neural networks. In this thesis, a novel approach for binary neural network classification is proposed. A pair of neural networks and an ensemble of several pairs of neural networks have been considered. Each pair constitutes two opposite networks trained to predict the degree of truth and false membership values. Normally, the predicted outputs always contain uncertainty. Quantification of uncertainty of type vagueness and error in binary neural network classification is also proposed. These uncertainties are presented in the form of indeterminacy membership values. The three memberships: truth, indeterminacy, and false memberships form the interval neutrosophic sets. Therefore, the proposal is based on a combination of binary neural networks and interval neutrosophic sets. Results from this study have been published in journal paper 1 and in conference paper 4. Furthermore, the results obtained from the proposed approach are compared to the results obtained from other existing approaches. This comparison has been published in conference paper 2. The study of binary classification is described in Chapter 3. In order to realize the binary classification approach, the proposed methodology has been applied to a real world problem of mineral prospectivity prediction. The study of mineral data has been published in two Springer lecture notes as papers 10 and 14, as well as in five conference papers 3, 11, 12, 13, and 15. This study is described in Chapter 5.

The proposed technique of using a pair of neural networks and the quantification of vagueness and error occurred in the prediction has also been applied to solve the problem of multiclass neural network classification. A combination of the proposed multiclass neural networks and interval neutrosophic sets has also been proposed in this study. This finding has been published in journal paper 6 and in two conference papers 5 and 8. This study is described in Chapter 4. The proposed multiclass classification approach has also been applied to the real world problem of lithofacies classification. This study has been published in journal paper 7 and in conference paper 9. This study is described in Chapter 6.

Contents

List of Figures

List of Tables

Nomenclature

ANN	Artificial Neural Network
Bayesian DT	Bayesian Decision Tree
BPNN	Feedforward Backpropagation Neural Networks
CBD	Constraint Based Decomposition
DT	Sonic Travel Time
EM	Expectation Maximization
F	False membership function
FPSVM	Fuzzy Proximal Support Vector Machine
GEPSVM	Generalized Eigenvalue Proximal Support Vector Machine
GIS	Geographic Information System
GR	Gramma RAy
GRNN	Generalized Regression Neural Network
I	Indeterminacy membership function
ILD	Deep Induction Resistivity
INLS	Interval Neutrosophic Logic System
INS	Interval Neutrosophic Sets
ITI	Incremental Decision Tree Induction
K5	k-nearest neighbors with k=5
LMDT	Linear Machine Decision Tree
LS-SVM	Least Squares Support Vector Machine
LS^2-SVM	Least Squares version of the Least Squares Support Vector Machine
LVQ	Learning Vector Quantization
NEVP	Nevada Backpropagation
NN	Neural Network
OCI	Induction of Oblique Trees

PNN	Polynomial Neural Network
PrNN	Probabilistic Neural Network
PSVM	Proximal Support Vector Machine
QoS	Quality of Service
SA	Simulated Annealing
SOM	Kohonen's Self-Organizing Maps
SVM	Support Vector Machine
SWS	Semantic Web Services
T	Truth membership function
TWSVM	Twin Support Vector Machine