

**IMPROVED VOTING TECHNIQUE FOR ENSEMBLE OF MLP SYSTEM APPLIED
ON VARIOUS CLASSIFICATION DATA**



**INSTITUT PENGURUSAN PENYELIDIKAN
UNIVERSITI TEKNOLOGI MARA
40450 SHAH ALAM, SELANGOR
MALAYSIA**

BY:

**SAODAH OMAR
IZA SAZANITA ISA
ASS. PROF. DR JUNITA MOHD SALEH**

MAY 2010

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS	ix
ABSTRAK	x
ABSTRACT	xi
1 INTRODUCTION	1
1.0 Introduction	1
1.1 Neural Network for Classification	1
1.2 Problem Statement	2
1.3 Objectives and Scope	2
1.4 Outline of Project	3
2 LITERATURE REVIEW	5
2.0 Introduction	5
2.1 Biological Neural Networks versus Artificial Neural Network	5
2.2 Artificial Neural Network Characteristics	8
2.3 Artificial Neural Network	9
2.3.1 Feed Forward ANN	10
2.3.2 Training Process	11
2.3.2.1 Supervised Training	11
2.3.2.2 Unsupervised Training	12
2.3.2.3 Reinforcement Training	12
2.4 Multilayer Perceptron	13
2.4.1 Activation Function	15
2.4.2 MLP Learning Algorithm	17
2.4.2.1 Levenberg Marquardt	20

2.4.2.2	Resilient Back Propagation (RP)	20
2.5	Neural Network Application	22
2.6	Application of MLP for Classification Problem	23
2.7	Application of Voting Technique	24
3	METHODOLOGY	25
3.0	Introduction	25
3.1	Proposed System	25
3.2	Development of Individual MLP Classifier	26
3.3	Development of Multiple MLP System	28
3.4	Majority voting	30
3.4.1	Implementation of Confidence Level	30
3.5	Case Study: Triangular Waveform Classification	30
3.5.1	Data Preparation	31
3.5.2	Development of Multiple MLP	34
3.5.3	Implementation of Voting Technique	35
3.6	Case Study: Breast Cancer Detection	38
3.6.1	Data Preparation	39
3.6.2	Development of Multiple MLP System	41
3.6.3	Implementation of Voting Technique	41
3.7	Case Study: Transportation Classification	43
3.7.1	Data Preparation	43
3.7.2	Development of Multiple MLP System	44
3.7.3	Implementation of Voting Technique	45
3.8	Performance Assessment	45
3.9	Summary	46
4	RESULTS AND ANALYSIS	47
4.0	Introduction	47
4.1	Case Study: Triangular Waveform Classification	47
4.1.1	Levenberg Marquardt (LM)	47
4.1.2	Resilient Back Propagation (RP)	48
4.1.3	Performance on Testing Phase of Two Algorithm	50
4.1.4	Majority Voting: Triangular Waveform Classification	51

4.1.5	Comparison with Previous Study	51
4.2	Case Study: Breast Cancer Detection	52
4.2.1	Levenberg Marquardt	52
4.2.2	Resilient Back Propagation	55
4.2.3	Majority Vote for Breast Cancer Detection	57
4.2.4	Proposed System Versus Previous Studies	58
4.3	Case Study: Transportation Classification	58
4.3.1	Levenberg Marquardt	59
4.3.2	Resilient Back Propagation	59
4.3.3	Majority Vote for Transportation Recognition	60
4.4	Summary	61
5	CONCLUSION	62
5.0	Conclusion	62
5.1	Recommendation for Future Development	63
APPENDIX A		- 1 -

ABSTRACT

An Artificial Neural Network (ANN) system has been extensively applied to numerous data classification problems such as cloud classification, business applications (sales forecasting), and medical domain for clinical diagnosis. The most well-known ANN architecture is the Multilayer Perceptron (MLP) network which is widely used for solving problems related to data classifications. However, the conventional ANN theory selects the best MLP (after training) for classification based on one which has the least number of hidden neurons, and gives the highest percentage of correct classification when if there are other MLPs (with more number of hidden neurons) which gives the same highest percentage of correct classification. The concept may not be correct since the other MLPs may perform better when presented with new datasets. Therefore, this project intends to investigate the capability of multiple MLP system with majority voting technique. It is a system which consists of all the best-performed MLPs and a single final output from these MLPs is selected by the voting system. The work employs MATLAB Neural Network Toolbox and Borland C++ programming language as the tools to develop the proposed system. The MLP networks are trained using two types of learning algorithm, which are the Levenberg Marquardt and the Resilient Back Propagation algorithms. The performance of the multiple MLP networks are calculated based on the percentage of correct classification. Data from two case studies; triangular waveform classification and breast cancer detection, have been used to test the performance of the developed system. The results show that the multiple MLP system with voting technique had the capability to improve the classification correctness.