UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Active Learning Algorithms for Multilayer Feedforward Neural Networks

By

Adebola Adebisi Adejumo

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Master of Science

in the Faculty of Science

AT

University of Pretoria

PRETORIA

2 December 1999

© University of Pretoria

To God Almighty who hath been my Tower of Refuge and Strenght

Abstract

Backpropagation (BP) has played a vital role in the resurgence of interest in artificial neural networks (ANNs). Eversince, a lot of research effort concentrated on finding ways to improve its performance. Active learning has emerged as an efficient alternative to improve the performance of multilayer feedforward neural networks. The learner is given active control over the information to include in the training set, and in doing so, the generalization accuracy is improved and the computational cost and complexity of the network are reduced compared to training on a fixed set of data.

While many research effort has been invested in designing new learning approaches, an elaborate comparison of active learning approaches is still lacking. The objective of this research study is to compare and critisize active learning approaches and also to propose a new selective learning algorithm.

This thesis presents a comparison of four selected active learning algorithms. The thesis concentrates on one type of application, namely function and time series approximation.

Opsomming

Terugwaartspropagering neurale netwerke het 'n belangrike rol gespeel in die oplewing van die belangstelling in kunsmatige neurale netwerke. Verskeie navorsingsstudies konsentreer op die verbetering van die prestasie van neurale netwerke. Aktiewe leer het getoon om 'n effektiewe alternatief te wees om die prestasie van multi-vlak vorentoe-voer neurale netwerke te verbeter. Die leerder word aktiewe beheer gegee oor die inligting wat in die leerversameling ingesluit word. Sodoende word veralgemening verbeter, en die berekeningskoste en -kompleksiteit van die netwerk verlaag in vergeleke met afrigting op 'n vaste leerversameling.

Terwyl vele navorsing gedoen is in die ontwikkeling van nuwe leerstrategieë, is daar 'n tekort aan 'n uitgebreide vergelykende studie van aktiewe leer. Die doelwit van hierdie studie is om aktiewe leer strategieë te vergelyk en te kritiseer. 'n Nuwe selektiewe leer algoritme word ook aangebied.

Hierdie tesis bied 'n vergelyking van vier aktiewe leer algoritmes aan. Die tesis konsentreer op die benadering van funksies en tydreekse.

Acknowledgements

I came, I saw and I conquered

I thank God Almighty for my stay and experience in South Africa, and most especially for His wisdom and guidance when I plunged my self into the unknown ocean of Neural Networks.

I deeply express my thanks to the following people:

- My supervisor, A.P Engelbrecht, for his time and advice.
- My sister and her hubby - Dr and Dr Ogunleye for making my postgraduate study a reality.
- My Uncle and his family - Uncle Akin, for his support and encouragement.
- My family at home especially my mother- For your prayers and encouragement.
- My Daddy - Late Prof J.A Adejumo, you inspired me, taught me the values of education, and that with power of knowledge, all men are equal.
- My Husband - Boye Aderogba, for waiting for me and being there when I needed you, thanks.

Contents

1 INTRODUCTION **1**

1.1 What is a Neural Network? 1

1.1.1 Characteristics Of A Neural Network? 3

1.2 Why Neural Networks? 4

1.2.1 Features of Neural Networks 6

1.3 Background to Neural Networks 7

1.4 Objective and Justification 10

1.5 Outline 11

2 MULTILAYER NEURAL NETWORK LEARNING **13**

2.1 Introduction 13

2.2 Biological Neural Networks 14

2.3 A Taxonomy Of Training 17

2.3.1 Topology of Neural Networks 19

2.3.2 Activation Functions 20

2.3.3 Neural Network Types 22

2.3.4 Optimization Algorithms 26

2.3.5 Why Neural Networks? 28

2.4 Gradient Descent Optimization 30

2.4.1 Introduction 30

2.4.2 Gradient descent training algorithm 32

2.4.3 Additional Features To The Training Algorithm 36

2.5 Learning Difficulties With Gradient Descent Optimization 37

2.5.1 Solutions to these learning difficulties 39

2.6 CONCLUSION 44

3 ACTIVE LEARNING **45**

3.1 Introduction 46

3.2 Concept of Active Learning 47

3.2.1 Overview on Active Learning 48

3.3 General Algorithm for Active Learning 51

3.4 A Comparative Study of Four Selected Active Learning Algorithms 53

3.4.1 A New Selective Learning Algorithm 54

3.4.2 Sensitivity Analysis Incremental Learning Algorithm 56

3.4.3 Dynamic Pattern Selection 59

3.4.4 Accelerated Learning by Active Example Selection 61

3.5 Experimental Results 62

3.5.1 Experimental Procedure 64

3.5.2 Results 70

3.6 Conclusion 83

4 CONCLUSION **86**

4.1 Future of Active Learning in Neural Networks 88

Bibliography **89**

A Symbols and Notations **98**

B Definitions **100**

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

List of Tables

3.1 Summary of the functions and time series used 63

3.2 Comparison results over 2000 epochs for times series problems 72

3.3 Comparison results over 2000 epochs for problems F1 and times series with noise and outliers 73

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

List of Figures

2.1 A simplified representation of a biological neuron 15

2.2 An artificial neuron 16

2.3 A multilayer neural network with a hidden layer 20

2.4 Activation functions 22

2.5 Recurrent Neural Networks (RNNs) 23

2.6 A functional link neural network (FLNN) 24

2.7 Product Unit Neural Networks 25

2.8 An illustration of local and global minimum 38

3.1 Function and Time series problems to be approximated 65

3.2 Time series problems to be approximated 66

3.3 Average generalization error vs epoch 74

3.4 Average generalization error vs epoch 75

3.5 Average generalization factor vs pattern presentations 78

3.6 Average generalization factor vs pattern presentations 79

3.7 Average number of patterns used per epoch 80

3.8 Average number of patterns used per epoch 81

3.9 Average computational cost per epoch 82

3.10 Average computational cost per epoch 83

3.11 Percentage simulations that converged 84

3.12 Percentage simulations that converged 85