Markus Stumptner
Dan Corbett
Mike Brooks (Eds.)

LNAI 2256

AI 2001: Advances in Artificial Intelligence

14th Australian Joint Conference on Artificial Intelligence
Adelaide, Australia, December 2001
Proceedings

Springer

Markus Stumptner Dan Corbett
Mike Brooks (Eds.)

AI 2001: Advances in Artificial Intelligence

14th Australian Joint Conference on Artificial Intelligence
Adelaide, Australia, December 10-14, 2001
Proceedings

Springer

Volume Editors

Markus Stumptner
Dan Corbett
University of South Australia, School of Computer and Information Science
Mawson Lakes, SA, Australia 5095
E-mail: {stumptner,corbett}@cs.unisa.edu.au

Mike Brooks
University of Adelaide, Department of Computer Science
Adelaide, SA, Australia 5001
E-mail: mjb@cs.adelaide.edu.au

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Advances in artificial intelligence : proceedings / AI 2001, 14th Australian Joint Conference on Artificial Intelligence, Adelaide, Australia, December 10 - 14, 2001. Markus Stumptner ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2001
(Lecture notes in computer science ; 2256 : Lecture notes in artificial intelligence)
ISBN 3-540-42960-3

CR Subject Classification (1998): I.2, F.1, F.4.1

ISBN 3-540-42960-3 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

Printed in Germany

Typesetting: Camera-ready by author
Printed on acid-free paper SPIN: 10846026 06/3142 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 2256

Subseries of Lecture Notes in Computer Science
Edited by J. G. Carbonell and J. Siekmann

Lecture Notes in Computer Science
Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Tokyo

Preface

This volume contains the proceedings of AI 2001, the 14th Australian Joint Conference on Artificial Intelligence. The aim of this conference series is to support the Australian artificial intelligence community with a forum for discussion and presentation. As before, this conference not only brought together a great deal of Australian AI research, but also attracted widespread international interest.

The conference this year saw an impressive array of about 110 submitted papers from no fewer than 16 countries. Full-length versions of all submitted papers were refereed by the international program committee. As a result, these proceedings contain 55 papers not just from Australia, but also Canada, France, Germany, The Netherlands, Japan, Korea, New Zealand, and the UK and USA.

The conference also comprised a tutorial program and several workshops, and featured five invited speakers on theoretical, philosophical, and applied topics: Didier Dubois of the Université Paul Sabatier, James Hendler of the University of Maryland, Liz Sonenberg of the University of Melbourne, Peter Struss of OCC'M Software and TU Munich, and Alex Zelinsky of the Australian National University.

We extend our thanks to the members of the program committee who processed a large review workload under tight time constraints. We especially thank our host, Professor Robin King, Pro Vice Chancellor of the Division of Information Technology, Engineering, and the Environment at UniSA, for providing infrastructure and financial support. We are also grateful to the US Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development, and the Commonwealth Defence Science and Technology Organisation for their financial support. Finally we would like to thank all those who contributed to the conference organization, without their help the conference could not have taken place.

December 2001

Dan Corbett
Markus Stumptner

AI 2001 Organization

Conference Chair: Dan Corbett, University of South Australia

Program Committee Co-chairs:
Mike Brooks, University of Adelaide
David Powers, Flinders University
Markus Stumptner, University of South Australia

Tutorial Chair: Alex Farkas, Tenix

Workshop Chair: Sam Mosel, Tenix

Applied AI Symposium Chair: Brian Hanisch, i2Net

Organizing Committee:
Committee Chair: Alex Farkas, Tenix
Carolyn Bellamy, University of South Australia
Richi Nayak, Queensland University of Technology
Dale Perin, University of South Australia
Sponsorship Chair: Helen Mitchard, Defence Science and Technology Organisation, Adelaide
Web Site: Anton van den Hengel, University of Adelaide

Program Committee

Grigoris Antoniou, Brisbane
Robert Dale, Sydney
Peter Deer, Adelaide
Susan George, Adelaide
Greg Gibbon, Newcastle
Ray Jarvis, Melbourne
Michael Maher, Brisbane
Mehmet Orgun, Sydney
Jeff Pelletier, Edmonton
Eric Tsui, Sydney
Peter Wallis, Melbourne
Mary-Anne Williams
Chengqi Zhang, Melbourne
Paul Compton, Sydney
John Debenham, Sydney
Peter Eklund, Brisbane
John Gero, Sydney
Achim Hoffmann, Sydney
Cara MacNish, Perth
Phan Minh Dung, Bangkok
Maurice Pagnucco, Sydney
Geoff Sutcliffe, Miami
Anton van den Hengel, Adelaide
Geoff Webb, Melbourne
Vilas Wuwongse, Bangkok

Additional Reviewers

Ayman Ammoura
Luigi Barone
John Bastian
Damien Brain
Vadin Bulitko
Samir Chopra
Jirapun Daengdej
Marc Denecker
Renee Elio
Matthias Fuchs
Bernd Groh
Thomas Havelka
Robert Holte
Ryszard Kozera
Rex Kwok
Yuefeng Li
Dekang Lin
Xudong Luo
Martin Luerssen
Wolfgang Mayer
Chris McDonald
Andrew Myers
Pavlos Peppas
Yongyuth Permpoontanalarp
Daniel Pooley
Krishna Rao
Jochen Renz
Chiaki Sakama
Stephan Schulz
Tran Cao Son
Michael Thielscher
Bruce Thomas
Quoc Bao Vo
Dominik Wieland
Franz Wotawa
Li-Yan Yuan
Dongmo Zhang
Hong Zhang
Zili Zhang
Zhi-Hua Zhou

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 2070: L. Monostori, J. Váncza, M. Ali (Eds.), Engineering of Intelligent Systems. Proceedings, 2001. XVIII, 951 pages. 2001.

Vol. 2080: D.W. Aha, I. Watson (Eds.), Case-Based Reasoning Research and Development. Proceedings, 2001. XII, 758 pages. 2001.

Vol. 2083: R. Goré, A. Leitsch, T. Nipkow (Eds.), Automated Reasoning. Proceedings, 2001. XV, 708 pages. 2001.

Vol. 2086: M. Luck, V. Mařík, O. Stěpánková, R. Trappl (Eds.), Multi-Agent Systems and Applications. Proceedings, 2001. X, 437 pages. 2001.

Vol. 2087: G. Kern-Isberner, Conditionals in Nonmonotonic Reasoning and Belief Revision. X, 190 pages. 2001.

Vol. 2099: P. de Groote, G. Morrill, C. Retoré (Eds.), Logical Aspects of Computational Linguistics. Proceedings, 2001. VIII, 311 pages. 2001.

Vol. 2100: R. Küsters, Non-Standard Inferences in Description Logics. X, 250 pages. 2001.

Vol. 2101: S. Quaglini, P. Barahona, S. Andreassen (Eds.), Artificial Intelligence in Medicine. Proceedings, 2001. XIV, 469 pages. 2001.

Vol. 2103: M. Hannebauer, J. Wendler, E. Pagello (Eds.), Balancing Reactivity and Social Deliberation in Multi-Agent Systems. VIII, 237 pages. 2001.

Vol. 2109: M. Bauer, P.J. Gmytrasiewicz, J. Vassileva (Eds.), User Modeling 2001. Proceedings, 2001. XIII, 318 pages. 2001.

Vol. 2111: D. Helmbold, B. Williamson (Eds.), Computational Learning Theory. Proceedings, 2001. IX, 631 pages. 2001.

Vol. 2112: R. Kowalczyk, S.W. Loke, N.E. Reed, G. Williams (Eds.), Advances in Artificial Intelligence. Proceedings, 2000. IX, 363 pages. 2001.

Vol. 2116: V. Akman, P. Bouquet, R. Thomason, R.A. Young (Eds.), Modeling and Using Context. Proceedings, 2001. XII, 472 pages. 2001.

Vol. 2117: M. Beynon, C.L. Nehaniv, K. Dautenhahn (Eds.), Cognitive Technology: Instruments of Mind. Proceedings, 2001. XV, 522 pages. 2001.

Vol. 2120: H.S. Delugach, G. Stumme (Eds.), Conceptual Structures: Broadening the Base. Proceedings, 2001. X, 377 pages. 2001.

Vol. 2123: P. Perner (Ed.), Machine Learning and Data Mining in Pattern Recognition. Proceedings, 2001. XI, 363 pages. 2001.

Vol. 2132: S.-T. Yuan, M. Yokoo (Eds.), Intelligent Agents: Specification, Modeling, and Application. Proceedings, 2001. X, 237 pages. 2001.

Vol. 2143: S. Benferhat, P. Besnard (Eds.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty. Proceedings, 2001. XIV, 818 pages. 2001.

Vol. 2148: A. Nareyek (Ed.), Local Search for Planning and Scheduling. Proceedings, 2000. IX, 171 pages. 2001.

Vol. 2155: H. Bunt, R.-J. Beun (Eds.), Cooperative Multimodal Communication. Proceedings, 1998. VIII, 251 pages. 2001.

Vol. 2157: C. Rouveirol, M. Sebag (Eds.), Inductive Logic Programming. Proceedings, 2001. X, 261 pages. 2001.

Vol. 2159: J. Kelemen, P. Sosík (Eds.), Advances in Artificial Life. Proceedings, 2001. XIX, 724 pages. 2001.

Vol. 2166: V. Matoušek, P. Mautner, R. Mouček, K. Taušer (Eds.), Text, Speech and Dialogue. Proceedings, 2001. XIII, 542 pages. 2001.

Vol. 2167: L. De Raedt, P. Flach (Eds.), Machine Learning: ECML 2001. Proceedings, 2001. XVII, 618 pages. 2001.

Vol. 2168: L. De Raedt, A. Siebes (Eds.), Principles of Data Mining and Knowledge Discovery. Proceedings, 2001. XVII, 510 pages. 2001.

Vol. 2173: T. Eiter, W. Faber, M. Truszczynski (Eds.), Logic Programming and Nonmonotonic Reasoning. Proceedings, 2001. XI, 444 pages. 2001.

Vol. 2174: F. Baader, G. Brewka, T. Eiter (Eds.), KI 2001: Advances in Artificial Intelligence. Proceedings, 2001. XIII, 471 pages. 2001.

Vol. 2175: F. Esposito (Ed.), AI*IA 2001: Advances in Artificial Intelligence. Proceedings, 2001. XII, 396 pages. 2001.

Vol. 2182: M. Klusch, F. Zambonelli (Eds.), Cooperative Information Agents V. Proceedings, 2001. XII, 288 pages. 2001.

Vol. 2190: A. de Antonio, R. Aylett, D. Ballin (Eds.), Intelligent Virtual Agents. Proceedings, 2001. VIII, 245 pages. 2001.

Vol. 2198: N. Zhong, Y. Yao, J. Liu, S. Ohsuga (Eds.), Web Intelligence: Research and Development. Proceedings, 2001. XVI, 615 pages. 2001.

Vol. 2225: N. Abe, R. Khardon, T. Zeugmann (Eds.), Algorithmic Learning Theory. Proceedings, 2001. XI, 379 pages. 2001.

Vol. 2226: K.P. Jantke, A. Shinohara (Eds.), Discovery Science. Proceedings, 2001. XII, 494 pages. 2001.

Vol. 2250: R. Nieuwenhuis, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. Proceedings, 2001. XV, 738 pages. 2001.

Vol. 2256: M. Stumptner, D. Corbett, M. Brooks (Eds.), AI 2001: Advances in Artificial Intelligence. Proceedings, 2001. XII, 666 pages. 2001.

Lecture Notes in Computer Science

Vol. 2201: G.D. Abowd, B. Brumitt, S. Shafer (Eds.), Ubicomp 2001: Ubiquitous Computing. Proceedings, 2001. XIII, 372 pages. 2001.

Vol. 2202: A. Restivo, S. Ronchi Della Rocca, L. Roversi (Eds.), Theoretical Computer Science. Proceedings, 2001. XI, 440 pages. 2001.

Vol. 2204: A. Brandstädt, V.B. Le (Eds.), Graph-Theoretic Concepts in Computer Science. Proceedings, 2001. X, 329 pages. 2001.

Vol. 2205: D.R. Montello (Ed.), Spatial Information Theory. Proceedings, 2001. XIV, 503 pages. 2001.

Vol. 2206: B. Reusch (Ed.), Computational Intelligence. Proceedings, 2001. XVII, 1003 pages. 2001.

Vol. 2207: I.W. Marshall, S. Nettles, N. Wakamiya (Eds.), Active Networks. Proceedings, 2001. IX, 165 pages. 2001.

Vol. 2208: W.J. Niessen, M.A. Viergever (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001. Proceedings, 2001. XXXV, 1446 pages. 2001.

Vol. 2209: W. Jonker (Ed.), Databases in Telecommunications II. Proceedings, 2001. VII, 179 pages. 2001.

Vol. 2210: Y. Liu, K. Tanaka, M. Iwata, T. Higuchi, M. Yasunaga (Eds.), Evolvable Systems: From Biology to Hardware. Proceedings, 2001. XI, 341 pages. 2001.

Vol. 2211: T.A. Henzinger, C.M. Kirsch (Eds.), Embedded Software. Proceedings, 2001. IX, 504 pages. 2001.

Vol. 2212: W. Lee, L. Mé, A. Wespi (Eds.), Recent Advances in Intrusion Detection. Proceedings, 2001. X, 205 pages. 2001.

Vol. 2213: M.J. van Sinderen, L.J.M. Nieuwenhuis (Eds.), Protocols for Multimedia Systems. Proceedings, 2001. XII, 239 pages. 2001.

Vol. 2214: O. Boldt, H. Jürgensen (Eds.), Automata Implementation. Proceedings, 1999. VIII, 183 pages. 2001.

Vol. 2215: N. Kobayashi, B.C. Pierce (Eds.), Theoretical Aspects of Computer Software. Proceedings, 2001. XV, 561 pages. 2001.

Vol. 2216: E.S. Al-Shaer, G. Pacifici (Eds.), Management of Multimedia on the Internet. Proceedings, 2001. XIV, 373 pages. 2001.

Vol. 2217: T. Gomi (Ed.), Evolutionary Robotics. Proceedings, 2001. XI, 139 pages. 2001.

Vol. 2218: R. Guerraoui (Ed.), Middleware 2001. Proceedings, 2001. XIII, 395 pages. 2001.

Vol. 2220: C. Johnson (Ed.), Interactive Systems. Proceedings, 2001. XII, 219 pages. 2001.

Vol. 2221: D.G. Feitelson, L. Rudolph (Eds.), Job Scheduling Strategies for Parallel Processing. Proceedings, 2001. VII, 207 pages. 2001.

Vol. 2224: H.S. Kunii, S. Jajodia, A. Sølvberg (Eds.), Conceptual Modeling – ER 2001. Proceedings, 2001. XIX, 614 pages. 2001.

Vol. 2225: N. Abe, R. Khardon, T. Zeugmann (Eds.), Algorithmic Learning Theory. Proceedings, 2001. XI, 379 pages. 2001. (Subseries LNAI).

Vol. 2226: K.P. Jantke, A. Shinohara (Eds.), Discovery Science. Proceedings, 2001. XII, 494 pages. 2001. (Subseries LNAI).

Vol. 2227: S. Boztaş, I.E. Shparlinski (Eds.), Applied Algebra, Algebraic Algorithms and Error-Correcting Codes. Proceedings, 2001. XII, 398 pages. 2001.

Vol. 2229: S. Qing, T. Okamoto, J. Zhou (Eds.), Information and Communications Security. Proceedings, 2001. XIV, 504 pages. 2001.

Vol. 2230: T. Katila, I.E. Magnin, P. Clarysse, J. Montagnat, J. Nenonen (Eds.), Functional Imaging and Modeling of the Heart. Proceedings, 2001. XI, 158 pages. 2001.

Vol. 2232: L. Fiege, G. Mühl, U. Wilhelm (Eds.), Electronic Commerce. Proceedings, 2001. X, 233 pages. 2001.

Vol. 2233: J. Crowcroft, M. Hofmann (Eds.), Networked Group Communication. Proceedings, 2001. X, 205 pages. 2001.

Vol. 2234: L. Pacholski, P. Ružička (Eds.), SOFSEM 2001: Theory and Practice of Informatics. Proceedings, 2001. XI, 347 pages. 2001.

Vol. 2237: P. Codognet (Ed.), Logic Programming. Proceedings, 2001. XI, 365 pages. 2001.

Vol. 2239: T. Walsh (Ed.), Principles and Practice of Constraint Programming – CP 2001. Proceedings, 2001. XIV, 788 pages. 2001.

Vol. 2240: G.P. Picco (Ed.), Mobile Agents. Proceedings, 2001. XIII, 277 pages. 2001.

Vol. 2241: M. Jünger, D. Naddef (Eds.), Computational Combinatorial Optimization. IX, 305 pages. 2001.

Vol. 2242: C.A. Lee (Ed.), Grid Computing – GRID 2001. Proceedings, 2001. XII, 185 pages. 2001.

Vol. 2250: R. Nieuwenhuis, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. Proceedings, 2001. XV, 738 pages. 2001. (Subseries LNAI).

Vol. 2248: C. Boyd (Ed.), Advances in Cryptology – ASIACRYPT 2001. Proceedings, 2001. XI, 603 pages. 2001.

Vol. 2256: M. Stumptner, D. Corbett, M. Brooks (Eds.), AI 2001: Advances in Artificial Intelligence. Proceedings, 2001. XII, 666 pages. 2001. (Subseries LNAI).

Table of Contents

Papers

Invited Contribution

A Memetic Pareto Evolutionary Approach to Artificial Neural Networks

H.A. Abbass

University of New South Wales, School of Computer Science, ADFA Campus,
Canberra, ACT 2600, Australia, h.abbass@adfa.edu.au.

Abstract. *Evolutionary Artificial Neural Networks* (EANN) have been a focus of research in the areas of *Evolutionary Algorithms* (EA) and *Artificial Neural Networks* (ANN) for the last decade. In this paper, we present an EANN approach based on pareto multi-objective optimization and differential evolution augmented with local search. We call the approach *Memetic Pareto Artificial Neural Networks* (MPANN). We show empirically that MPANN is capable to overcome the slow training of traditional EANN with equivalent or better generalization.

Keywords: neural networks, genetic algorithms

1 Introduction

Evolutionary Artificial Neural Networks (EANNs) have been a key research area for the last decade. On the one hand, methods and techniques have been developed to find better approaches for evolving *Artificial Neural Networks* and more precisely - for the sake of our paper – Multi–layer feed–forward *Artificial Neural Networks* (ANNs). On the other hand, finding a good ANNs' architecture has been an issue as well in the field of ANNs. Methods for network growing (such as Cascade Correlation [4]) and for network pruning (such as Optimal Brain Damage [14]) have been used to overcome the long process for determining a good network architecture. However, all these methods still suffer from their slow convergence and long training time. In addition, they are based on gradient–based techniques and therefore can easily stuck in a local minimum. EANNs provide a better platform for optimizing both the network performance and architecture simultaneously. Unfortunately, all of the research undertaken in the EANN literature ignores the fact that there is always a trade–off between the architecture and the generalization ability of the network. A network with more hidden units may perform better on the training set, but may not generalize well on the test set. This trade–off is a well known problem in *Optimization* known as the Multi-objective Optimization Problem (MOP).

With the trade–off between the network architecture – taken in this paper to be the number of hidden units – and the generalization error, the EANN problem is in effect a MOP. It is, therefore, natural to raise the question of why not applying a multi–objective approach to EANN.

The objective of this paper is to present a *Memetic* (*ie. evolutionary algorithms* (EAs) augmented with local search [18]) Pareto Artificial Neural Networks (MPANN). The rest of the paper is organized as follows: In Section 2, background materials are covered followed by an explanation of the methods in Section 3. Results are discussed in Section 4 and conclusions are drawn in Section 5.

2 Background Materials

In this section, we introduce necessary background materials for Multi-objective Optimization, ANNs, *Differential Evolution* (DEs), Evolutionary Multi-objective, and EANN.

2.1 Multi-objective Optimization

Consider a *Multi-Objective Optimization Problem* (MOP) model as presented below:-

$$\begin{gathered}\text{Optimize } F(\boldsymbol{x}) \\ \text{subject to: } \Omega = \{\boldsymbol{x} \in R^n | G(\boldsymbol{x}) \leq 0\}\end{gathered}$$

Where $\boldsymbol{x}$ is a vector of decision variables $(x_1, \ldots, x_n)$ and $F(\boldsymbol{x})$ is a vector of objective functions $(f_1(\boldsymbol{x}), \ldots, f_K(\boldsymbol{x}))$. Here $f_1(\boldsymbol{x}), \ldots, f_K(\boldsymbol{x})$, are functions on R^n and Ω is a nonempty set in R^n. The vector $G(\boldsymbol{x})$ represents a set of constraints.

In MOPs, the aim is to find the optimal solution $\boldsymbol{x}^* \in \Omega$ which optimize $F(\boldsymbol{x})$. Each objective function, $f_i(\boldsymbol{x})$, is either maximization or minimization. Without any loss of generality, we assume that all objectives are to be minimized for clarity purposes. We may note that any maximization objective can be transformed to a minimization one by multiplying the former by -1.

To define the concept of non-dominated solutions in MOPs, we need to define two operators, $\ncong$ and $\precsim$ and then assume two vectors, $\boldsymbol{x}$ and $\boldsymbol{y}$. We define the first operator as $\boldsymbol{x} \ncong \boldsymbol{y}$ iff $\exists\, x_i \in \boldsymbol{x}$ and $y_i \in \boldsymbol{y}$ such that $x_i \neq y_i$. And, $\boldsymbol{x} \precsim \boldsymbol{y}$ iff $\forall\, x_i \in \boldsymbol{x}$ and $y_i \in \boldsymbol{y}, x_i \leq y_i$, and $\boldsymbol{x} \ncong \boldsymbol{y}$. The operators $\ncong$ and $\precsim$ can be seen as the "not equal to" and "less than or equal to" operators respectively, between two vectors. We can now define the concepts of local and global optimality in MOPs.

Definition 1: Neighborhood or open ball The open ball (*ie.* a neighborhood centered on $\boldsymbol{x}^*$ and defined by the Euclidean distance) $B_\delta(\boldsymbol{x}^*) = \{\boldsymbol{x} \in R^n | \; ||\boldsymbol{x} - \boldsymbol{x}^*|| < \delta\}$.

Definition 2: Local efficient (non-inferior/ pareto-optimal) solution A vector $\boldsymbol{x}^* \in \Omega$ is said to be a local efficient solution of MOP iff $\nexists\, \boldsymbol{x} \in (B_\delta(\boldsymbol{x}^*) \cap \Omega)$ such that $F(\boldsymbol{x}) \precsim F(\boldsymbol{x}^*)$ for some positive δ.

Definition 3: Global efficient (non-inferior/ pareto-optimal) solution A vector $\boldsymbol{x}^* \in \Omega$ is said to be a global efficient solution of MOP iff $\nexists\, \boldsymbol{x} \in \Omega$ such that $F(\boldsymbol{x}) \precsim F(\boldsymbol{x}^*)$.

Definition 4: Local non-dominated solution A vector $\boldsymbol{y}^* \in F(\boldsymbol{x})$ is said to be local non-dominated solution of MOP iff its projection onto the decision space, $\boldsymbol{x}^*$, is a local efficient solution of MOP.

Definition 5: Global non-dominated solution A vector $\boldsymbol{y}^* \in F(\boldsymbol{x})$ is said to be global non-dominated solution of MOP iff its projection onto the decision space, $\boldsymbol{x}^*$, is a global efficient solution of MOP.

In this paper, the term "non-dominated solution" is used as a shortcut for the term "local non-dominated solution".

2.2 Artificial Neural Networks

We may define an ANN by a graph: $G(N, A, \psi)$, where N is a set of neurons (also called nodes), A denotes the connections (also called arcs or synapses) between the neurons, and ψ represents the learning rule whereby neurons are able to adjust the strengths of their interconnections. A neuron receives its inputs (also called activation) from an external source or from other neurons in the network. It then undertakes some processing on this input and sends the result as an output. The underlying function of a neuron is called the activation function. The activation, a, is calculated as a weighted sum of the inputs to the node in addition to a constant value called the bias. The bias can be easily augmented to the input set and considered as a constant input. From herein, the following notations will be used for a single hidden layer MLP:

- I and H are the number of input and hidden units respectively.
- $\mathbf{X}^p \in \mathbf{X} = (x_1^p, x_2^p, \ldots, x_I^p), p = 1, \ldots P$, is the pth pattern in the input feature space $\mathbf{X}$ of dimension I, and P is the total number of patterns.
- Without any loss of generality, $\mathbf{Y}_o^p \in \mathbf{Y_o}$ is the corresponding scalar of pattern $\mathbf{X}^p$ in the hypothesis space $\mathbf{Y_o}$.
- w_{ih} and w_{ho}, are the weights connecting input unit i, $i = 1 \ldots I$, to hidden unit h, $h = 1 \ldots H$, and hidden unit h to the output unit o (where o is assumed to be 1 in this paper) respectively.
- $\Theta_h(\mathbf{X}^p) = \sigma(a_h)$; $a_h = \sum_{i=0}^{I} w_{ih} x_i^p$, $h = 1 \ldots H$, is the h^{th} hidden unit's output corresponding to the input pattern $\mathbf{X}^p$, where a_h is the activation of hidden unit h, and $\sigma(.)$ is the activation function that is taken in this paper to be the logistic function $\sigma(z) = \frac{1}{1+e^{-Dz}}$, with D the function's sharpness or steepness and is taken to be 1 unless it is mentioned otherwise.
- $\hat{Y}_o^p = \sigma(a_o)$; $a_o = \sum_{h=0}^{H} w_{ho} \Theta_h(\mathbf{X}^p)$ is the network output and a_o is the activation of output unit o corresponding to the input pattern $\mathbf{X}^p$.

MLPs are in essence non-parametric regression methods which approximate underlying functionality in data by minimizing a risk function. The data are presented to the network and the risk function is approximated empirically R_{emp} by summing over all data instances as follows:

$$R_{emp}(\alpha) = \sum_{p=1}^{P} (Y_o^p - \hat{Y}_o^p)^2 \tag{1}$$