INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE AND EXPERT SYSTEMS (IJAE)

VOLUME 3, ISSUE 2, 2012

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 2180-124X

International Journal of Artificial Intelligence and Expert Systems (IJAE) is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJAE Journal is a part of CSC Publishers
Computer Science Journals
http://www.cscjournals.org
EDITORIAL PREFACE

The International Journal of Artificial Intelligence and Expert Systems (IJAE) is an effective medium for interchange of high quality theoretical and applied research in Artificial Intelligence and Expert Systems domain from theoretical research to application development. This is the second issue of volume three of IJAE. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJAE emphasizes on efficient and effective Artificial Intelligence, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of Expert Systems. IJAE comprehensively cover the system, processing and application aspects of Artificial Intelligence. Some of the important topics are AI for Service Engineering and Automated Reasoning, Evolutionary and Swarm Algorithms and Expert System Development Stages, Fuzzy Sets and logic and Knowledge-Based Systems, Problem solving Methods Self-Healing and Autonomous Systems etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 3, 2012, IJAE appears in more focused issues. Besides normal publications, IJAE intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

IJAE give an opportunity to scientists, researchers, and vendors from different disciplines of Artificial Intelligence to share the ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development. This journal is helpful for the researchers and R&D engineers, scientists all those persons who are involve in Artificial Intelligence and Expert Systems in any shape.

Highly professional scholars give their efforts, valuable time, expertise and motivation to IJAE as Editorial board members. All submissions are evaluated by the International Editorial Board. The International Editorial Board ensures that significant developments in image processing from around the world are reflected in the IJAE publications.

IJAE editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJAE. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJAE provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Artificial Intelligence and Expert Systems (IJAE)
EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)
Dr. Bekir Karlik
Mevlana University (Turkey)

ASSOCIATE EDITORS (AEiCs)

Assistant Professor. Tossapon Boongoen
Royal Thai Air Force Academy
Thailand

Assistant Professor. Ihsan Omur Bucak
Mevlana University
Turkey

Professor Ahmed Bouridane
Northumbria University
United Kingdom

EDITORIAL BOARD MEMBERS (EBMs)

Professor Yevgeniy Bodyanskiy
Kharkiv National University of Radio Electronics
Ukraine

Assistant Professor. Bilal Alatas
Firat University
Turkey

Associate Professor Abdullah Hamed Al-Badi
Sultan Qaboos University
Oman

Dr. Salman A. Khan
King Fahd University of Petroleum & Minerals
Saudi Arabia

Assistant Professor Israel Gonzalez-Carrasco
Universidad Carlos III de Madrid
Spain

Dr. Alex James
Indian Institute of Information Technology and Management- Kerala
India

Assistant Professor Dr Zubair Baig
King Fahd University
Saudi Arabia
Associate Professor Syed Saad Azhar Ali  
Iqra University  
Pakistan

Assistant Professor Israel Gonzalez-Carrasco  
Universidad Carlos III de Madrid  
Spain

Professor Sadiq M. Sait  
King Fahd University  
Saudi Arabia

Professor Hisham Al-Rawi  
University of Bahrain  
Bahrain

Dr. Syed Zafar Shazli  
Northeastern University  
United States of America

Associate Professor Kamran Arshad  
University of Greenwich  
United Kingdom

Associate Professor Ashraf Anwar  
University of Atlanta  
United States of America
<table>
<thead>
<tr>
<th>Pages</th>
<th>Title</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 - 38</td>
<td>Incorporating Kalman Filter in the Optimization of Quantum Neural Network Parameters</td>
<td>Hayder Mahdi Abdulridha</td>
</tr>
</tbody>
</table>
Incorporating Kalman Filter in the Optimization of Quantum Neural Network Parameters

Hayder Mahdi Alkhafaji
drengaider@uobabylon.edu.iq
Electrical Engineering Dept.
Babylon University
Babylon, Iraq

Abstract

Kalman filter has been used for the estimation of instantaneous states of linear dynamic systems. It is a good tool for inferring of missing information from noisy measurement. The quantum neural network is another approach to the merging of fuzzy logic with the neural network and that by the investment of quantum mechanics theory in building the structure of neural network. The gradient descent algorithm has been used widely in training the neural network, but the problem of local minima is one of the disadvantages of this algorithm. This paper presents an algorithm to train the quantum neural network by using the extended kalman filter.

Keywords: Quantum Neural, Extended Kalman Filter, Training

1. INTRODUCTION

Since the innovation of first simple artificial neuron, the neural networks gain the interest of researchers. Many topologies of neural networks have been proposed as a trial to find the best architecture and make it more powerful in classification, recognition, function approximation, control, and other applications [1],[2],[3],[4]. In trying to enhance the ability of neural network, many approaches have been used in conjunction with neural networks, such as fuzzy and genetic [6],[7].

Quantum mechanics is one of the attractive approaches, which inspire the researchers great ideas and applications in various fields like communication, control, and others [5],[6],[7]. In 1941, Stevens and others [8] present a work paper used the quantum theory in the discrimination of loudness and pitch, by using rectilinear functions instead of classical integral functions, it can elude the unpredictable points in the classical functions. Purushothman and Karayiannis [9] proposed a neural network with multilevel squash function in the hidden layer nodes to imitate the fuzzy logic and overcome the problems of combining the neural network with fuzzy. The problems resulted from either explicitly training Fuzzy-Neural networks (FNNs) with fuzzy membership values estimated a priori, or by training FNNs with the available crisp membership information and interpreting their response as being fuzzy in itself.


Due to the problem of local minima in using gradient descent method in training neural networks, the researchers in this field try to find new approaches in training. Brady and others [12] prove that gradient descent on a surface defined by a sum of squared errors can fail to separate families of vectors.
Kalman filter is one of the alternatives in training neural networks that can be used to process the missing data [13],[14]. In a previous work many researchers invest the Kalman filter in training neural networks [15],[16],[17].

In this study, the quantum neural network parameters have been optimized using the extended Kalman filter. The results show the power of Kalman filter to speed-up the finding process of network parameters values in few iterations.

2. QUANTUM NEURAL NETWORK
Many topologies have been proposed by modeling the quantum neural network inspiring the mathematical background of quantum mechanics theory [18],[19],[20]. In this paper, a model proposed by Gopathy and Nicholaos [9] is adopted to be used as a classification network. The idea behind this topology of this network is to build multilevel neurons in the hidden layer to imitate the fuzzy sets. The nonlinear classifier divides the input data space into classes which are recognized by collapsing-in over regions of certainty or spreading-out over regions of uncertainty in the feature space. Every hidden unit is represented by a multilevel function to formulate the graded partitions instead of the linear partitions.

The network consists of three layers as shown in figure 1. The first layer receives the input vectors $x_i$, where $i$ stands for the index of the input vectors $x$. The input vectors should be

$$x = \begin{bmatrix}
    x_{i1} & \ldots & x_{iI} & \ldots & x_{iM} \\
    \vdots & \ddots & \ddots & \ddots & \vdots \\
    \vdots & & \ddots & \ddots & \vdots \\
    x_{ni} & \ldots & \ldots & x_{niM}
\end{bmatrix}$$

(1)

Where $n_i$ and $M$ are input vector lengths and number of input vectors, respectively.

![Architecture of quantum neural network](image)

**FIGURE 1:** Architecture of quantum neural network.

Every neuron in the hidden layer receives weighted sum of the input vector and can be evaluated as
The output of every hidden layer neuron is passed to a graded compound function which consists of the summation of a number of shifted sigmoid functions. The shift of each sigmoid function specifies the jump to the next level of the quantum based function. The following function represents the output of hidden neurons:

\[ h_b_j = \frac{1}{n} \sum_{r} f_h(h_a_j - \gamma_j^r) \]  

Where

\( n_s = \text{no. of quantum levels} \)
\( f_h = \text{squash function of hidden neurons} \)
\( \gamma = \text{quantum level shifts} \)
\( r = \text{index of quantum level shifts} \)

The weights matrix \( v \) represents the weights between input layer and hidden layer, while matrix \( w \) contains the weights between hidden layer and output layer, as follows

\[
v_{ij} = \begin{bmatrix}
  v_{11} & \ldots & v_{1,j} & \ldots & v_{1,\text{nh}} \\
  \vdots & \ddots & \vdots & & \vdots \\
  \vdots & & \ddots & \ddots & \vdots \\
  v_{ni,1} & \ldots & \ldots & v_{ni,\text{nh}} 
\end{bmatrix}
\]

(4)

\[
w_{kj} = \begin{bmatrix}
  w_{11} & \ldots & w_{1,j} & \ldots & v_{1,\text{nh+1}} \\
  \vdots & \ddots & \vdots & & \vdots \\
  \vdots & & \ddots & \ddots & \vdots \\
  w_{\text{no},1} & \ldots & \ldots & w_{\text{no},\text{nh+1}} 
\end{bmatrix}
\]

(5)

where

\( \text{nh} = \text{hidden layer size} \)
\( \text{no} = \text{output layer size} \)

The jump positions matrix for the quantum hidden units can be represented as
The neural network output depends on finding the output of every node of the output layer by evaluating the result of squash function which receives its input from the hidden layer multiplied by the weights between hidden layer and output layer. To do so, the following two equations reveal that.

\[
\gamma_{j,r} = \begin{bmatrix}
\gamma_{11} & \cdots & \gamma_{1,r} & \cdots & \gamma_{1,ns} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\gamma_{nh,1} & \cdots & \gamma_{nh,r} & \cdots & \gamma_{nh,ns}
\end{bmatrix}
\]

(6)

The neural network output depends on finding the output of every node of the output layer by evaluating the result of squash function which receives its input from the hidden layer multiplied by the weights between hidden layer and output layer. To do so, the following two equations reveal that.

\[
hc = [hb^T \cdot w^T]^T
\]

(7)

\[
y_{hat}^l = f_o(h_c^k, l)
\]

(8)

Where

- \( f_o \) = squash function of output layer
- \( k \) = index of output layer neurons
- \( l \) = index of input vectors

3. QUANTUM NEURAL PARAMETERS OPTIMIZATION BASED ON KALMAN FILTER

Because of the nonlinear nature of neural networks, it is evident that other tools used with it should be capable of dealing with such paradigms. The unscented Kalman filter, which is the modified nonlinear version of Kalman filter, can be used in conjunction with neural network to predict the parameters of the network. The process and output equations for the classifier system formulated by the unscented Kalman filter are:

\[
\theta_{k+1} = f(\theta_k) + \omega_k
\]

(9)

\[
y_k = h(\theta_k) + \nu_k
\]

Where

- \( \theta_k \) = system state vector
- \( \omega_k \) = process noise
- \( \nu_k \) = measurement noise
- \( y_k \) = system output

The process noise is assumed to have zero mean and \( Q_k \) covariance, while the measurement noise has zero mean and \( R_k \) covariance. The state vector of the system can be represented as
\[ \theta = [w_{1,1} \ldots w_{1,nh+1} w_{2,1} \ldots w_{no,nh+1} v_{1,1} \ldots v_{ni,1} v_{1,2} \ldots v_{ni,nh} \ldots \\
\gamma_{1,1} \gamma_{1,2} \ldots \gamma_{1,ns} \ldots \gamma_{nh,1} \ldots \gamma_{nh,ns} ] \]

(10)

The nonlinear process and measurement system equations can be expanded around the state estimate \( \tilde{\theta}_k \) by Taylor series as follows:

\[
\begin{align*}
\quad \quad f(\theta_k) &= f(\tilde{\theta}_k) + F_k \times (\theta_k - \tilde{\theta}_k) + \text{higer order terms} \\
\quad \quad h(\theta_k) &= h(\tilde{\theta}_k) + H_k^T \times (\theta_k - \tilde{\theta}_k) + \text{higer order terms}
\end{align*}
\]

(11)

Where

\[
\begin{align*}
F_k &= \frac{\partial f(\theta)}{\partial \theta} \bigg|_{\theta=\tilde{\theta}_k} \\
H_k^T &= \frac{\partial h(\theta)}{\partial \theta} \bigg|_{\theta=\tilde{\theta}_k}
\end{align*}
\]

(12)

The higher order can be neglected to get

\[
\begin{align*}
\quad \quad \theta_k &= F_k \tilde{\theta}_k + \omega_k + \psi_k \\
\quad \quad y_k &= H_k^T \tilde{\theta}_k + \nu_k + \phi_k
\end{align*}
\]

(13)

where

\[
\begin{align*}
\psi_k &= f(\tilde{\theta}_k) - F_k \tilde{\theta}_k \\
\phi_k &= h(\tilde{\theta}_k) - H_k^T \tilde{\theta}_k
\end{align*}
\]

(14)

To estimate the network parameters value by using Kalman filter, it is necessary to formulate an objective function which stands as a condition for reaching the optimal state. The mean square error can be used, where the error represents the difference between the estimated output and the desired output, as follows:
\[ E = \frac{1}{2} \sum (h(\hat{\theta}) - y)^2 \] (15)

The recursion of the following equations, depending on an error limit for stopping the iteration, can result in an optimal state estimate of the network parameters.

\[
K_k = P_k H_k (R + H_k^T P_k H_k)^{-1}
\]

\[
\hat{\theta}_k = f(\hat{\theta}_{k-1}) + K_k [y_k - h(\hat{\theta}_{k-1})]
\]

\[
P_{k+1} = F_k (P_k - K_k H_k^T P_k) F_k^T + Q
\] (16)

Where

\( K_k \) = Kalman gain
\( P_k \) = estimation-error covariance
\( R \) = measurement noise covariance
\( Q \) = process noise covariance
\( H_k \) = partial derivative of the network output with respect to network parameters.

The partial derivative matrix is obtained as below:

\[
H = [H_1 \ H_2 \ H_3]^T
\] (17)

Where \( H_1, H_2, \) and \( H_3 \) are evaluated as

\[
H_1 = \begin{bmatrix}
\frac{\partial f_o^{1,1}}{\partial w_{1,1}} & \cdots & \frac{\partial f_o^{1,M}}{\partial w_{1,1}} & \cdots & \frac{\partial f_o^{n_0,M}}{\partial w_{1,1}}
\end{bmatrix}
\]

\[
H_2 = \begin{bmatrix}
\frac{\partial f_o^{1,1}}{\partial w_{1,nh}} & \cdots & \cdots & \cdots & \frac{\partial f_o^{n_0,M}}{\partial w_{1,nh}}
\end{bmatrix}
\]

\[
H_3 = \begin{bmatrix}
\frac{\partial f_o^{1,1}}{\partial w_{n_0,nh+1}} & \cdots & \cdots & \cdots & \frac{\partial f_o^{n_0,M}}{\partial w_{n_0,nh+1}}
\end{bmatrix}
\]
4. SIMULATION RESULTS
The aim of this paper is to find the best values for the selected parameters of quantum neural network. The network will be tested in a classification problem, where the classified data will be the known iris data set. The data set contains three categories of 50 patterns for each category. Each pattern consists of four features. The dataset will be divided into two categories. The first one is used for training and the second is used for testing.

The Quantum neural network composes of three layers: input, hidden and output layers. The input layer receives the feature vectors with four nodes in length. The hidden one contains the quantum hidden nodes, where every node is composed of a multi level squash function. The output layer gives the classification result.

Figure 2 shows the quantum function of the hidden layer neurons for the case where six neurons are used in this layer. It can be seen that every neuron function has different shape from others and this because of the moving of each single component of the compound function to right or left according the updating rule of the training algorithm. The composite quantum function will be resulted from the addition of these single components, where each component has its own mean as a consequence to the shifting operation during training phase.

To evaluate the quantum neural network by testing the parameters of the network, it should be taken in regard the performance of the network through introducing different set of data from that
used in training process. An illustration of the performance of the quantum neural network is shown in Figure 3. Twenty five vectors of testing data are introduced to the network with different number of hidden neurons to get the highest correct classification results.

From figure 3, it can be seen that the classification ability is degraded when the hidden layer nodes exceed certain number, so, the number of hidden layer nodes must be suitable and it can be selected to be the sum of input and output layer lengths. In our classification problem the hidden layer length could be seven nodes.
FIGURE 2: The configuration of quantum neuron squash function for every node of the hidden layer, in case when nh=6.

FIGURE 3: Average performance against no. of hidden neurons for neural network and quantum neural network. Dashed line represents neural network and continuous line for quantum neural network.

Another criterion can be used to evaluate the network which is the number of iterations taken by training algorithm to reach the optimal values of network parameters. This criterion which is against the number of hidden units is presented in figure 4. The curves of this figure reveal the ability of quantum neural network over the feed forward neural network in reach the optimum values of network parameters in less iterations for any number of hidden units.

FIGURE 4: average Iteration against no. of hidden neurons dashed for neural continuous for quantum neural.
5. CONCLUSIONS

The merit of this kind of networks, which is the quantum neural network, over other networks, is its ability to imitate the fuzzy logic by a simple way. Every specialist in artificial intelligence knows the importance of fuzzy logic as an efficient tool in reasoning and it is considered as a glass box because of its transparency in formulating and treating the problems in contrast with neural networks which is considered as a black box. Many topologies have been proposed by merging the neural networks and fuzzy logic to gather the adaptation ability of neural networks and fuzzy reasoning to get an efficient tool. The quantum neural network can be considered as another image to the former, but it has simpler structure.

Due to the problem of local minima, the researcher try to find other methods to overcome this problem. One of the good solutions is using extended Kalman filter in the optimization of neural network. In this paper, a combination of the benefits of quantum neural network and Kalman filter have been proposed and the results of classifying Iris data show the efficiency of the network.

6. REFERENCES


INSTRUCTIONS TO CONTRIBUTORS

The main aim of International Journal of Artificial Intelligence and Expert Systems (IJAE) is to provide a platform to AI & Expert Systems (ES) scientists and professionals to share their research and report new advances in the field of AI and ES. IJAE is a refereed journal producing well-written original research articles and studies, high quality papers as well as state-of-the-art surveys related to AI and ES. By establishing an effective channel of communication between theoretical researchers and practitioners, IJAE provides necessary support to practitioners in the design and development of intelligent and expert systems, and the difficulties faced by the practitioners in using the theoretical results provide feedback to the theoreticians to revalidate their models. IJAE thus meets the demand of both theoretical and applied researchers in artificial intelligence, soft computing and expert systems.

IJAE is a broad journal covering all branches of Artificial Intelligence and Expert Systems and its application in the topics including but not limited to technology & computing, fuzzy logic, expert systems, neural networks, reasoning and evolution, automatic control, mechatronics, robotics, web intelligence applications, heuristic and AI planning strategies and tools, computational theories of learning, intelligent system architectures.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJAE.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 3, 2012, IJAE appears in more focused issues. Besides normal publications, IJAE intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

LIST OF TOPICS
The realm of International Journal of Artificial Intelligence and Expert Systems (IJAE) extends, but not limited, to the following:

- AI for Web Intelligence Applications
- AI Parallel Processing Tools
- AI Tools for Computer Vision and Speech Understand
- Application in VLSI Algorithms and Mobile Communication
- Case-based reasoning
- Derivative-free Optimization Algorithms
- Evolutionary and Swarm Algorithms
- Expert Systems Components
- Fuzzy Sets and logic
- Hybridization of Intelligent Models/algorithms
- Inference
- Intelligent Planning
- AI in Bioinformatics
- AI Tools for CAD and VLSI Analysis/Design/Testing
- AI Tools for Multimedia
- Automated Reasoning
- Data and Web Mining
- Emotional Intelligence
- Expert System Development Stages
- Expert-System Development Lifecycle
- Heuristic and AI Planning Strategies and Tools
- Image Understanding
- Integrated/Hybrid AI Approaches
- Intelligent Search
CALL FOR PAPERS

Volume: 3 • Issue: 4 – August 2012


iii. Issue Publication: August 2012
CONTACT INFORMATION

Computer Science Journals Sdn Bhd
B-5-8 Plaza Mont Kiara, Mont Kiara
50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6207 1607
       006 03 2782 6991

Fax:   006 03 6207 1697

Email: cscpress@cscjournals.org