University of West Florida

# A Report on Recent Computer Science Conferences

**Eman El-Sheikh**

**Department of Computer Science**

**University of West Florida**

**March 6, 2009**

# WorldComp'08 Overview

- 2008 World Congress in Computer Science, Computer Engineering, and Applied Computing
  - July 14 – 17, 2008, Las Vegas, NV
  - "Many of the 25 joint conferences are the premier conferences for presentation of advances in their respective fields."
  - Over 2000 researchers from 82 different countries
  - International Conference on Artificial Intelligence (ICAI'08)
  - International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS'08)

# WorldComp'08 Conferences

- International Conference on Bioinformatics and Computational Biology (BIOCOMP'08)
- International Conference on Computer Design (CDES'08)
- International Conference on Computer Graphics and Virtual Reality (CGVR'08)
- International Conference on Communications in Computing (CIC'08)
- International Conference on Scientific Computing (CSC'08)
- International Conference on Data Mining (DMIN'08)
- International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE'08)
- International Conference on Engineering of Reconfigurable Systems and Algorithms (ERSA'08)
- International Conference on Embedded Systems and Applications (ESA'08)
- International Conference on Foundations of Computer Science (FCS'08)
- International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS'08)
- International Conference on Grid Computing and Applications (GCA'08)
- International Conference on Genetic and Evolutionary Methods (GEM'08)
- International Conference on Artificial Intelligence (ICAI'08)
- International Conference on Internet Computing (ICOMP'08)
- International Conference on Wireless Networks (ICWN'08)
- International Conference on Information and Knowledge Engineering (IKE'08)
- International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'08)
- International Conference on Information Theory and Statistical Learning (ITSL'08)
- International Conference on Machine Learning; Models, Technologies and Applications (MLMTA'08)
- International Conference on Modeling, Simulation and Visualization Methods (MSV'08)
- International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'08)
- International Conference on Security and Management (SAM'08)
- International Conference on Software Engineering Research and Practice (SERP'08)
- International Conference on Semantic Web and Web Services (SWWS'08)

# FECS'08 Overview

- Conference included recent research results and trends in CS and CE education:
  - Collaborative learning
  - Evaluation strategies
  - Game programming
  - Learning tools and emulators
  - Online learning
  - Teaching strategies
  - Virtual worlds
  - Web-based educational tools

# ICAI'08 Overview

- Conference included recent research results and trends in various AI topics:
  - Agent technologies
  - Genetic algorithms
  - Fuzzy logic and applications
  - Information retrieval
  - Knowledge-based systems
  - Natural language processing
  - Neural networks and applications
  - Robotics

# ICAI'08: Neural Networks

- Conference included several useful results in the area of Neural Networks (NNs).

- NNs have been successfully used for a variety of applications, ranging from stock price prediction to automatic target detection and recognition.

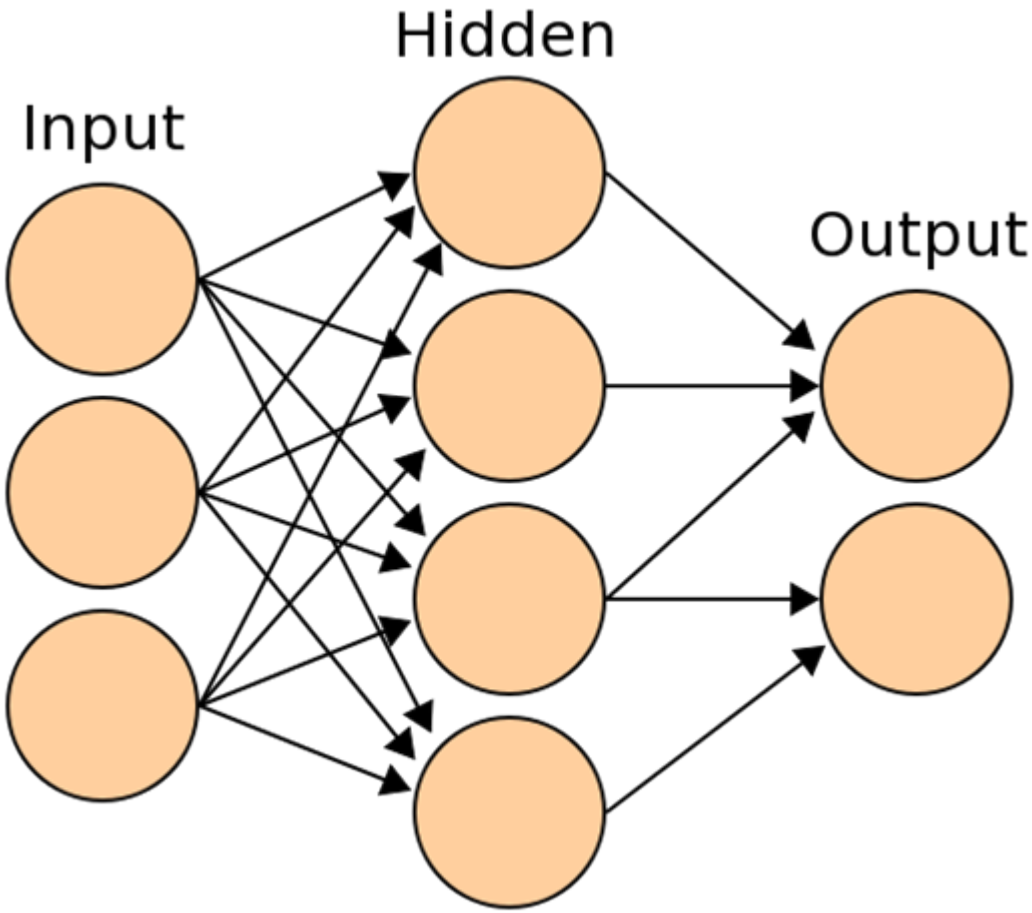
- Especially useful as pattern recognizers/classifiers due to their:
  - Robustness
  - Fault tolerance
  - Universal function approximation capability

# Background Information

- A multi-layer feed-forward NN with enough nodes and hidden layers can approximate any bounded function and its derivative to any arbitrary accuracy given a “large enough” training set.

- A network with *N* inputs, one hidden layer of *H* units, and a total of *W* weights requires O(*W*/*e*) training patterns to yield error < *e* on the test set.

# Neural Network Model



# Is Your Neural Network Learning or Memorizing?

**Feng-Jen Yang**
**Computer Science Department**
**Prairie View A&M University**

- Ideally, learning should be meaningful and comprehensive but rote learning may happen in real life as well as in simulated biological intelligence such as artificial neural networks.
- The author performed a series of experiments on five different NNs to learn the same quadratic equation:

$$y(x) = 2x^2 + 3x + 1, \text{ where } 1 \leq x \leq 30.$$

- Analyzed three statistical measures to indicate whether a NN is learning or memorizing:
  - Learning curve, mean and variance of error ratios
- All NNs had the same input layer, output layer, learning rate, etc, but different number of neurons in the hidden layer.
- Hypothesis: Do more neurons lead to better learning?

# Is Your Neural Network Learning or Memorizing (cont'd)?

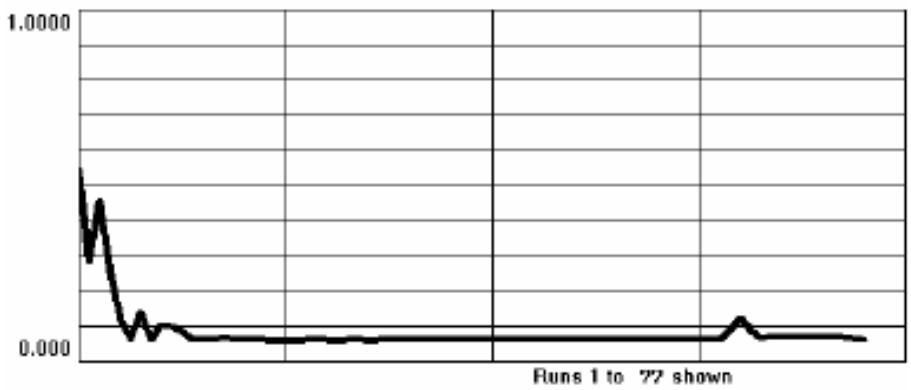
- Having more than enough neurons can turn a neural network from learning into memorizing and hurt the overall learning process.

- Overall comparison:

| Number of Hidden neurons | 10 | 15 | 100 | 200 | 300 |
|---|---|---|---|---|---|
| Learning Curve | Smooth | uneven | uneven | rough | very rough |
| Mean of Error Rate | Small | slightly high | high | very high | very high |
| Variance of Error Rate | Small | slightly high | high | very high | very high |
| Learn Behavior | meaningful | less meaningful | memorizing | highly memorizing | extremely memorizing |

- Sample test results & learning curve NN# 4 (200 hidden neurons):

| X | $y_d(x)$ | y(x) | e(x) |
|---|---|---|---|
| 4 | 45 | 12.16 | 0.73 |
| 8 | 153 | 33.78 | 0.78 |
| 12 | 323 | 136.87 | 0.58 |
| 16 | 561 | 452.11 | 0.19 |
| 20 | 861 | 958.33 | 0.11 |
| 24 | 1225 | 1391.30 | 0.14 |
| 28 | 1653 | 1634.30 | 0.01 |
| The mean of e(x): 0.36 | | | |
| The variance of e(x): 0.32 | | | |
| The minimum of e(x): 0.11 | | | |
| The maximum of e(x): 0.78 | | | |



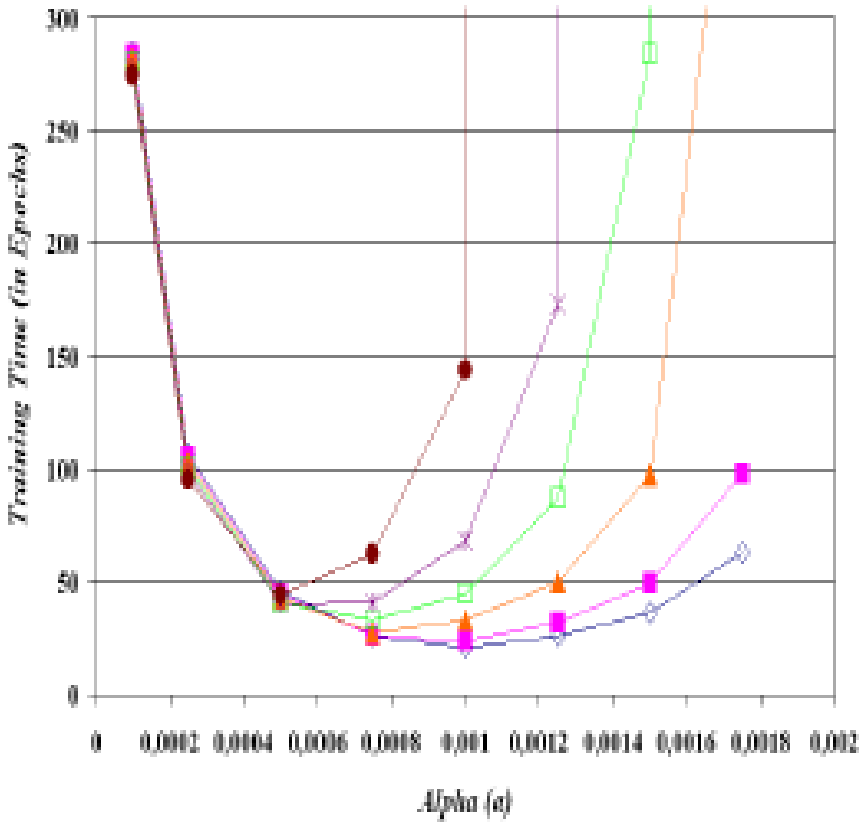
# A Practical Method for Finding an Efficient Static Learning Rate for an ANN

**Celso Camilo[1] and Keiji Yamanaka[2]**
**[1]Federal University of Grande Dourados, Brazil**
**[2]Federal University of Uberlandia, Brazil**

- Learning algorithms often have difficulties converging quickly:
  - Need to adjust learning rate carefully
- The authors propose a method for finding an efficient static learning rate (α) for single-layer nets based on:
  - Input vector dimensions
  - Input values

# A Practical Method for Finding an Efficient Static Learning Rate for an ANN (cont'd)

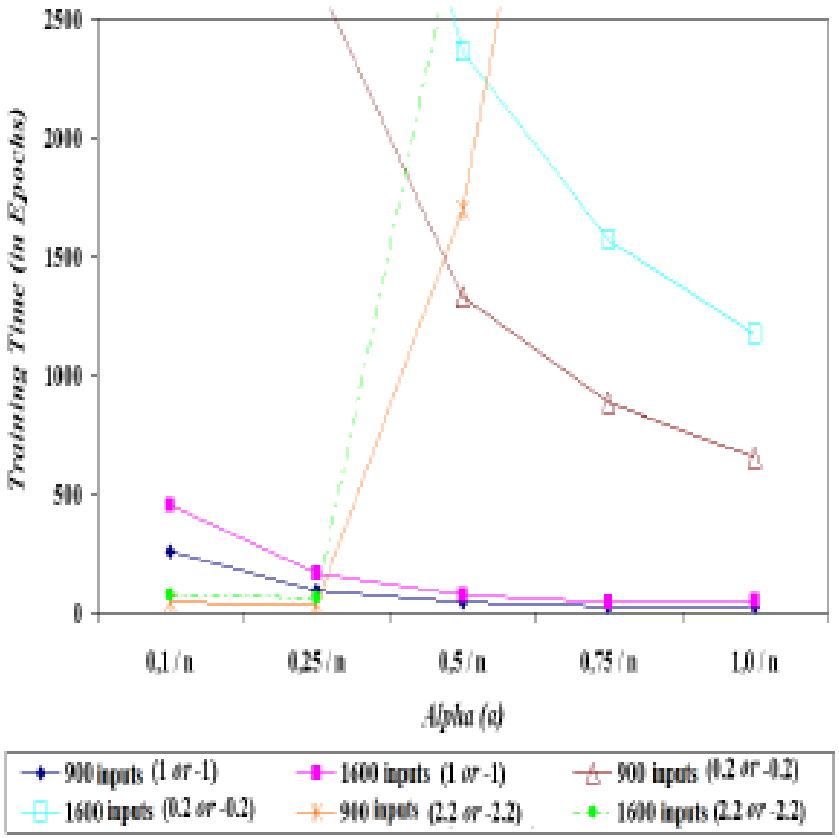
- When the number of inputs increases, the best α decreases:




| Code | Inputs | Outputs |
|---|---|---|
| 1 | 900 (Matrix 30 x 30) | 6 |
| 2 | 961 (Matrix 31 x 31) | 6 |
| 3 | 1089 (Matrix 33 x 33) | 6 |
| 4 | 1225 (Matrix 35 x 35) | 6 |
| 5 | 1369 (Matrix 37 x 37) | 6 |
| 6 | 1600 (Matrix 40 x 40) | 6 |

# A Practical Method for Finding an Efficient Static Learning Rate for an ANN (cont'd)

- When the input value <= 1, the best α tends to the upper limit (1/n), and when the value > 1, it tends to the lower limit (0.1/n):




| Code | input neurons | input values |
|---|---|---|
| 1 | 900 (Matrix 30 x 30) | (1 or -1) |
| 2 | 1600 (Matrix 40 x 40) | (1 or -1) |
| 3 | 900 (Matrix 30 x 30) | (0.2 or -0.2) |
| 4 | 1600 (Matrix 40 x 40) | (0.2 or -0.2) |
| 5 | 900 (Matrix 30 x 30) | (2.2 or -2.2) |
| 6 | 1600 (Matrix 40 x 40) | (2.2 or -2.2) |

# Use of Neural Networks to Detect Impaired Cheating on the Computerized CERAD Word List Memory Test

**Timothy Chan and Amar Raheja**
**Department of Computer Science**
**California State Polytechnic University**

- The computerized CERAD Word List Memory Test is used by the Long Term Care insurance industry to screen applicants for mild cognitive impairment, and for early detection of Alzheimer's Disease.
- The authors developed a method to detect cheating by cognitively impaired individuals using a two-stage approach:
  1) distinguish cheaters from non-cheaters
  2) distinguish impaired cheaters from unimpaired-cheaters
- They employed the use of NNs and compared results with a statistical method.

# Use of Neural Networks to Detect Impaired Cheating on the Computerized CERAD Word List Memory Test (cont'd)

- They found that the NN did a better job in stage 1 (2Stage1) by virtue of having a dominating average ROC (Receiver Operating Characteristic) curve.
- They also explored a one stage approach (1Stage).
- They compared the two stage and one stage NN and found that the former approach marginally outperformed the latter with a higher sensitivity yet similar positive predictive value (PPV) and specificity.

| | *Confusion matrix* | Predicted Negative | Predicted Positive |
|---|---|---|---|
| Actual | Negative | 5652.5±55.4 | 128.08±7.6 |
| | Positive | 10.3±2.2 | 4.71±1.4 |

TABLE II
CONFUSION MATRIX FOR *1Stage*. PPV IS 3.56 ±1.05%, SENSITIVITY IS 31.49 ± 8.05%, WHILE SPECIFICITY IS 97.78 ± 0.13%.

| | *Confusion matrix* | Predicted Negative | Predicted Positive |
|---|---|---|---|
| Actual | Negative | 5629.7±54.7 | 134.3±8.5 |
| | Positive | 7.48±1.73 | 5.26±1.53 |

TABLE III
CONFUSION MATRIX FOR TWO STEP METHOD: COMBINING *2Stage1* AND *2Stage2*. PPV IS 3.77±1.09%, SENSITIVITY IS 41.27 ± 9.66%, WHILE SPECIFICITY IS 97.67 ± 0.15%.

- Results still outperformed other statistical methods.

# A Comparison of Neural Network Architectures for Handwritten Digit Recognition

**Eman El-Sheikh, Bradley Swain, and Mohamed Khabou**

**University of West Florida**

- Project focus:
  - Development, use, and analysis of an artificial neural network architecture for recognizing handwritten digit data.
- Project Goal:
  - To study the effects of varying architectures, learning rates, number and sensitivity of neurons on the ability to distinguish the digits.
- Potential applications:
  - Mail sorting
  - Automated check reading
  - Data entry for hand-held devices

# Neural Network Architecture

- Architecture was implemented in Java, using a modular and parameterized design model.

- Network parameters include:
  - number of input neurons
  - number of hidden layers
  - number of neurons per hidden layer
  - number of output neurons
  - learning rate
  - number of training epochs

# Data Description

- A collection of handwritten digits collected by U.S. Postal Service from ZIP codes.
- Three digit features were chosen and compared for their ability to distinguish the digits:
  - Information measure
  - Orthogonality measure
  - Combination measure
- All three feature data sets consisted of data for 6000 digits, of which 2000 were used for training, and 3500 used for testing.

# Sample Data

# Results

| Hidden Layer(s) | Info Features | Time (mm:ss) | Orthogonal Features | Time (mm:ss) | Combination Features | Time (mm:ss) |
|---|---|---|---|---|---|---|
| 25/20 | 90.63% | 04:18 | 93.77% | 04:20 | 93.74% | 04:15 |
| 25/15 | 90.83% | 03:56 | 93.34% | 03:59 | 93.34% | 03:59 |
| 25 | 90.54% | 03:18 | 93.57% | 03:19 | 93.43% | 03:20 |
| 15 | 89.91% | 02:08 | 92.86% | 02:08 | 92.83% | 02:08 |

# Results

- Using simple back-propagation, feed-forward neural networks, we could correctly predict approx. 94% of 3500 handwritten digits within four minutes.
- An asymptotic bound was approached for the time/accuracy trade-off.
- An increase of only 0.2% accuracy cost another minute of processing time.

# Conclusions

- Architecture successfully classified test data using each of three feature sets.
- Orthogonal features produced best results, which influenced the combination features.
- Information feature set did not perform as well due to similarity in features.
- Single hidden layer network with 25 nodes performed nearly as well as two hidden layer network with 25 & 20 nodes, with better time efficiency.