Introduction to Cascade Correlation Networks

Willy Rempel

February 2, 2017

Features of the Beamer Class

- Normal LaTeX class. Stuff!! [, 1]

Features of the Beamer Class

- Normal LaTeX class. Stuff!! [, 1]
- Easy overlays.

Features of the Beamer Class

- Normal LaTeX class. Stuff!! [, 1]
- Easy overlays.
- No external programs needed.

Train input weight - correlation

S is sum over all output units o of correlation with error

$$S = \sum_o |\sum_p (V_p - V)(E_{p,o} - E_o)|$$

o	network output at which the error is measured
ρ	the training pattern
σ	network output
V_p	candidate output for input pattern ρ
$E_{p,o}$	network output error for output o, pattern ρ
$\overline{V}$	average of candidate output over all patterns
$\overline{E_o}$	average of output errors over all patterns

0.1 trainingInputun

Figure: image

Train input weight - correlation

$$\frac{\delta S}{\delta w_i} = \sum_{p,o} \sigma_o (E_{p,o} - \overline{E_o}) f_p{}' I_{i,p}$$

o	network output at which the err
ρ	the training pattern
$\overline{V}$	average of candidate unit outpu
$\overline{E_o}$	average of output errors over al

Figure: image

- doesn't require backprop? https://en.wikipedia.org/wiki/Types_of_artificial_neural_networks#Cascading_neural_networks

book

In spite of the many CoNN algorithms surveyed in (Kwok & Yeung, 1997a), the most popular for regression problems is no doubt the Cascade Correlation algorithm (CasCor) and maybe the second most popular is the DNC. While the DNC algorithm constructs neural networks with a single hidden layer, the CasCor creates them with multiple hidden layers, where each hidden layer has one hidden neuron. The popularity of CasCor can be attested by the various ways this algorithm has inspired new variations and also has been used in the combined approaches between learning methods. p15.

constructive algo's paper

- describes overall field.
- describes evo techniques as a broad category different from CCAs

other

other slides

- input units code the problem being presented to the network
- output units code the network's response to the input problem

shultz slides

- 'when error stagnates a hidden unit is recruited'
- correlation slide shows normalizing denominator. Not in original paper.
- states same quickprop algo used for both correlation maximization and error minimization
- the added randomized output weights are opposite sign of neurons correlation with network error –true?
  - confirmed in `http://www.ra.cs.uni-tuebingen.de/SNNS/UserManual/node167.html` tutorial
- aside: he has 2nd tutorial on encoder option for CCNNs. Later.

things to say

- why did CCann fall out? what happened since inception?
- current uses
- for practical uses, would like pictures of the headings of papers to show what some researches have to say
- 2 main problems of backprop :
  - step size problem
  - moving target problem
- original paper acknowledges vanishing gradient problem
- original paper citation count 3716
- creates it's own topology starting with minimal network
  - input and output layers only, as usual connected by weights.
- opposite of DNNs, they start big and stay big. CCNN start small and grows with training.
  - in a class called Dynamic Node Creation, DNC . . . change this
- a 'multi layered perceptron'
- S Fahlman did RCC - recurrent CCNN same year '90

References I


S. E. Fahlman and C. Lebiere, The cascade-correlation learning architecture.


K. Khatter and J. Kaur, Global Journal of Engineering Science and Research Management.


S. K. Sharma and P. Chandra, CONSTRUCTIVE NEURAL NETWORKS: A REVIEW, International Journal of Engineering Science and Technology, vol. 1, no. 2, pp. 78477855.


T.-Y. Kwok and D.-Y. Yeung, Constructive algorithms for structure learning in feedforward neural networks for regression problems, IEEE Transactions on Neural Networks, vol. 8, no. 3, pp. 630645, 1997.


G. Balzs, Cascade-Correlation Neural Networks: A Survey.


Y. Guo, L. Bai, S. Lao, S. Wu, and M. S. Lew, A Comparison between Artificial Neural Network and Cascade-Correlation Neural Network in Concept Classification, in Advances in Multimedia Information Processing PCM 2014, 2014, pp. 248253.


B. K. Wong, T. A. Bodnovich, and V. S.-K. Lai, The Use of Cascade-Correlation Neural Networks in University Fund Raising, The Journal of the Operational Research Society, vol. 51, no. 8, pp. 913920, 2000.


A. B. Nassif, L. F. Capretz, and D. Ho, Software Effort Estimation in the Early Stages of the Software Life Cycle Using a Cascade Correlation Neural Network Model, in 2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2012, pp. 589594.


S. Saha, A. Konar, A. Saha, A. K. Sadhu, B. Banerjee, and A. K. Nagar, EEG based gesture mimicking by an artificial limb using cascade-correlation learning architecture, in 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 46804687.

References II

P. Blonda, G. Pasquariello, and J. Smid, Comparison of backpropagation, cascade-correlation and Kokonen algorithms for cloud retrieval, in Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan), 1993, vol. 2, pp. 12311234 vol.2.

N. Sharma and H. Om, Cascade correlation neural network model for classification of oral cancer, WSEAS Transactions on Biology and Biomedicine, vol. 11, pp. 4551, 2014.

B. Chandra and P. P. Varghese, Applications of Cascade Correlation Neural Networks for Cipher System Identification, World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, vol. 1, no. 2, pp. 369372, 2007.

N. A. Singh and T. Naryanan, Application of Cascaded Correlation Neural Network for Financial Performance Prediction and Analysis of BSNL, in Swarm, Evolutionary, and Memetic Computing, 2014, pp. 502513.