ROLE OF COMPUTATIONAL INTELLIGENCE IN TARGETED FORECASTING

Mohamed A. El-Sharkawi

Mohamed A. El-Sharkawi

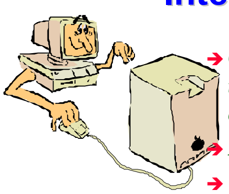
Department of EE, Box 352500
University of Washington
Seattle, WA 98195-2500
USA
elsharkawi@ee.washington.edu

NN Application to Electric Load Forecasting

Mohamed A. El-Sharkawi

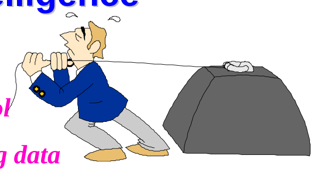
Mohamed A. El-Sharkawi
Department of EE, Box 352500
University of Washington
Seattle, WA 98195-2500
USA
e-mail: elsharkawi@ee.washington.edu
URL: http://cialab.ee.washington.edu

Why Computational Intelligence?



- ***Complex nonlinear mapping through a set of input/output examples.***
- ***No structured model.***
- ***Variables can be easily include or excluded.***
- ***Superior noise rejection capability.***
- ***Fast execution.***

Challenges to Computational Intelligence

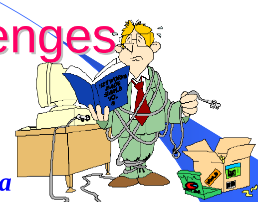


- ***Architecture***
- ***Learning protocol***
- ***Range of training data***
- ***Data spanning the operational space***
- ***Data statistical properties.***
- ***Correlated Features.***

Challenges to Computational Intelligence

- ***Sensitivity to variation in topology, operation, and control characteristics.***
- ***Memorization.***
- ***Saturation.***
- ***Adaptation to system dynamics.***

Fundamental Challenges

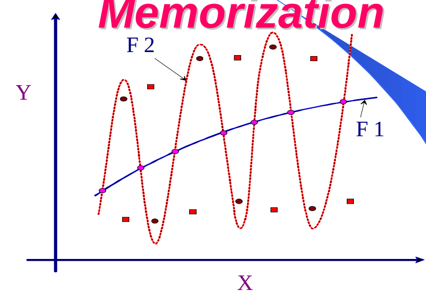


- ***NN architecture***
- ***Learning protocol***
- ***Range of training data***
- ***Distribution of data in operational space***
- ***Statistical properties of the data.***
- ***Formulation of features***

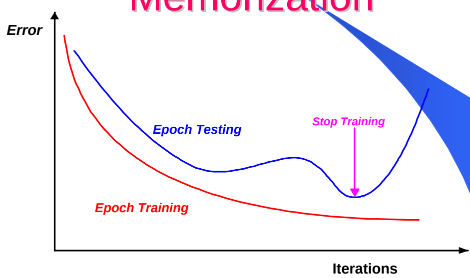
Fundamental Challenges

- ***Memorization***
- ***Saturation***
- ***Sensitivity to variations in topology, operation, and control characteristics.***
- ***Adaptation to system dynamics***
- ***Border Tracking***

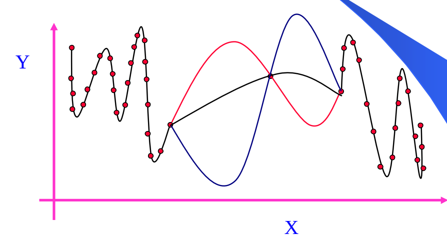
Learning versus Memorization



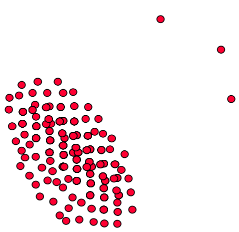
Learning versus Memorization



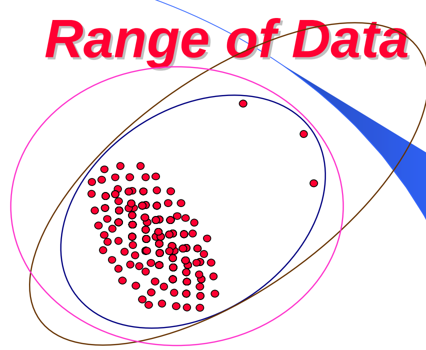
Data Distribution



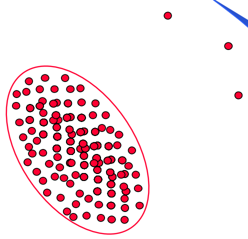
Range of Data



Range of Data



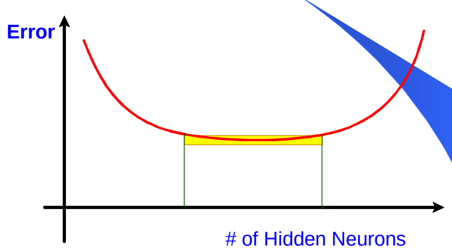
Range of Data



Neural Net size

- *The number of hidden neurons, must be matched to the complexity of the classification boundary.*
- *Cross validation among neural networks can provide good approximation of the net size.*
- *The best NN structure is where the error of the NN is relatively unchanged.*

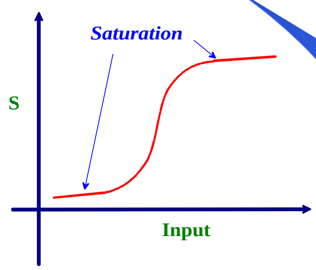
Neural Net size



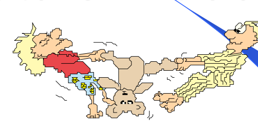
Network Saturation

- *Nonlinear functions reaches its limits*
- *A wide change in the input produce minimal change in the output*
- *With large number of saturated neurons, the NN can be paralyzed.*
- *If saturated, neurons must be randomly perturbed.*

Network Saturation



Feature Extraction

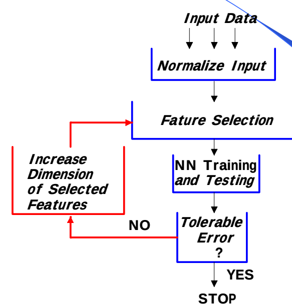


- *To train a NN with reasonable accuracy, a sufficiently large data set spanning the operating space is required.*
- *When correlated data is hard to obtain.*

Why Feature Extraction?

- *Eliminates curse of dimensionality.*
- *Enhances class separability.*
- *Reduces pattern dimension*
- *Maintains classification accuracy.*

Feature Extraction



Electric Load Forecasting

- *Efficient power system operation.*
- *Scheduling of spinning reserves.*
- *Economical energy interchange.*
- *Operational security.*

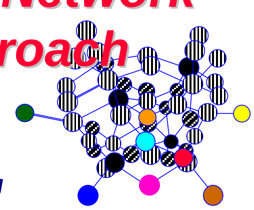
Conventional Solutions

- *Regression*
- *Time series*
- *Expert system*
- *Human Forecast*

Challenges

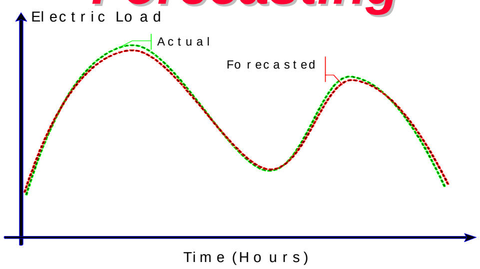
- *Relevancy of variables*
- *Geographical dependency*
- *Accuracy of weather forecasting.*
- *Size of training data set.*
- *Adaptive forecasting*
- *Accuracy of extrapolation (cold snap, heat wave, pickup loads)*
- *Thermal inertia*
- *Peak forecasting*

Neural Network Approach



- *Nonlinear mapping*
- *No pre-assumed functional relationships*
- *Emphases can be placed on certain times*
- *Ease of implementation*

NN Load Forecasting



Targeted Forecasting

- *Similar Day Forecasting*
- *Peak Forecasting*
- *Adaptive Forecasting*

Similar Day Forecasting:

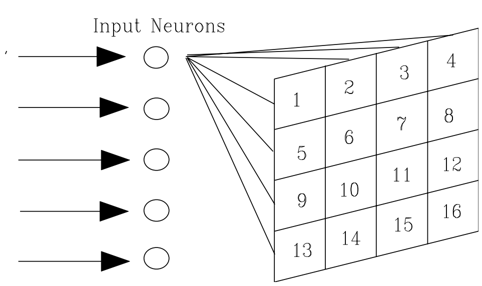
KOHONEN NEURAL NETWORK Classifier

- **It is a clustering technique**
- **Output status is not needed for training patterns.**
- **Only a selected sample of the training pattern needs to be assessed through simulations.**
- **Once the network is adequately trained, neurons that respond to certain status are self organized in clusters.**

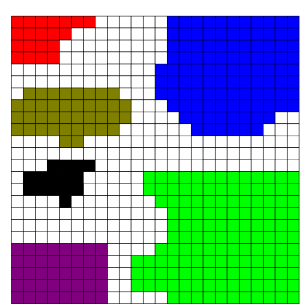
KOHONEN NEURAL NETWORK Classifier

- **In the testing stage, the pattern status is determined by correlating the test pattern with a cluster of a known status.**
- **Kohonen network is usually a two-dimensional array of neurons.**
- **Each input is connected to every neuron in the array.**
- **The distance between two neurons is expressed in terms of the "neighborhood order."**

KOHONEN NEURAL NETWORK Classifier



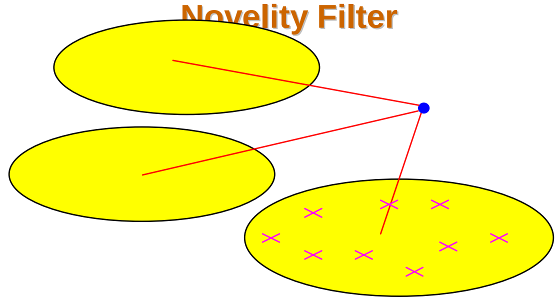
KOHONEN MAP



Similar Day Forecasting: Novelity Filter

- Similar days are bounded in a cluster in the forecasting space
- Bounding is an optimization problem
- The number of bounded clusters depends on the number and uniqueness of the special days
- Novelity filtering technique can identify the "closeness" of the tested day to each cluster

Similar Day Forecasting: Novelity Filter



Fuzzified Novelity Filter

- Fuzzy Systems can be used to identify the load of the test date based on a collective data from all clusters
- An Adaptive mechanism can be established to modify and generate new clusters

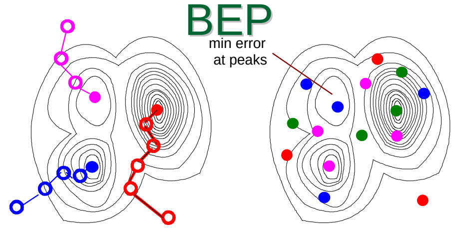
Peak Forecasting

- To enhance the forecasting accuracy during load peaks
- NN designed to forecast 24 hour can be modified to allow a weighted back-error-propagation
- NN error can be structured to forecast peaks
- Genetic Algorithm (GA) can identify selective minima

Peak Forecasting: GA

- *Single Search Technique (SST), back-error-propagation, may trap the NN is undesirable minimum*
- *Convergence of SST to a minima for peak loads depends on the selected initial condition of NN weights*

Genetic Algorithm vs BEP



Single Search

Genetic Search

ADAPTIVE FORECASTING

- For dynamically varying systems with/without large data sets (Load forecasting, security, etc.)
- Weights are automatically adjusted based on new data
- Effect of old and invalid patterns (data) are eventually and automatically deleted (forgotten)
- Perturbation in the NN weights are restricted to chosen boundaries
- Global optimality can be obtained
- Adaptive training does not drift
- Data can be weighted based on its importance

ADAPTIVE FORECASTING

$$e(i) = [\, t(i) - o(i)\,]^2$$
$$E(N) = \Sigma_N \, [t(N) - o(N)]^2$$
$$E(N+1) = E(N) + e(N+1)$$
$$W(N+1) = W(N) + \Delta W(N+1)$$

Adaptive condition

$$D\, W(N+1) < \phi$$

Drifting condition

If $x(N+1) = x(i)$, $i = 1, 2, \ldots, \text{or } N$

then

$$\Delta\, W(N+1) = 0$$