An Innovative Predictive Modeling tool: NeuroSequences

Graphion LLC.

http://chenstory.com

NeuroSequences

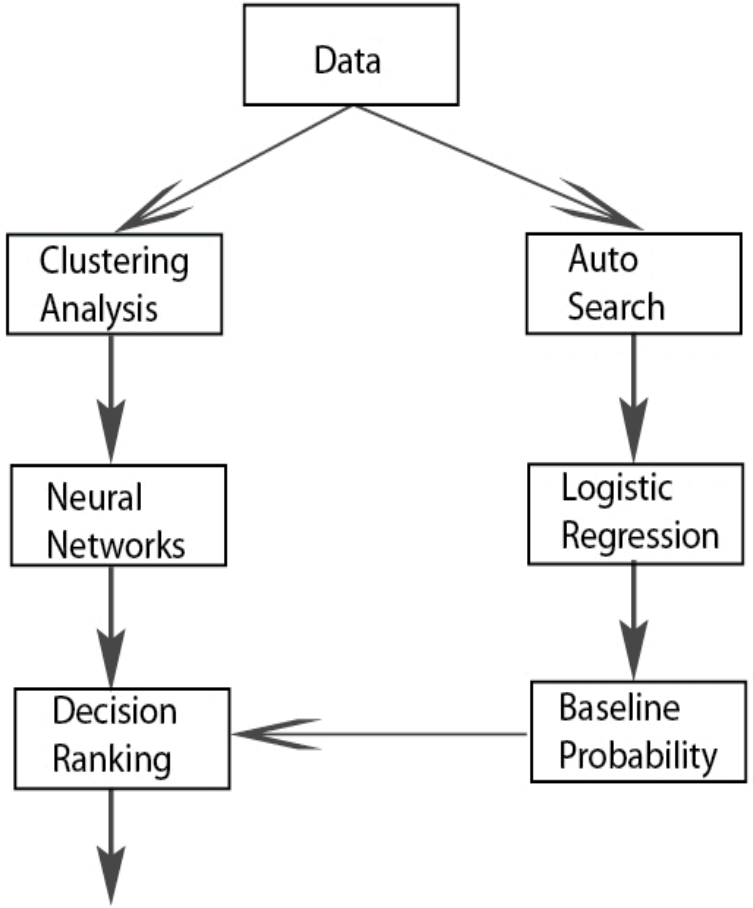
- Predictive Modeling
- Baseline Probability
- Auto Search Logistic Regression
- Clustering Analysis
- Escaping from local minima
- Monitoring MLP's free parameters
- Fast training with Lo-Shu data sampling

Predictive Modeling

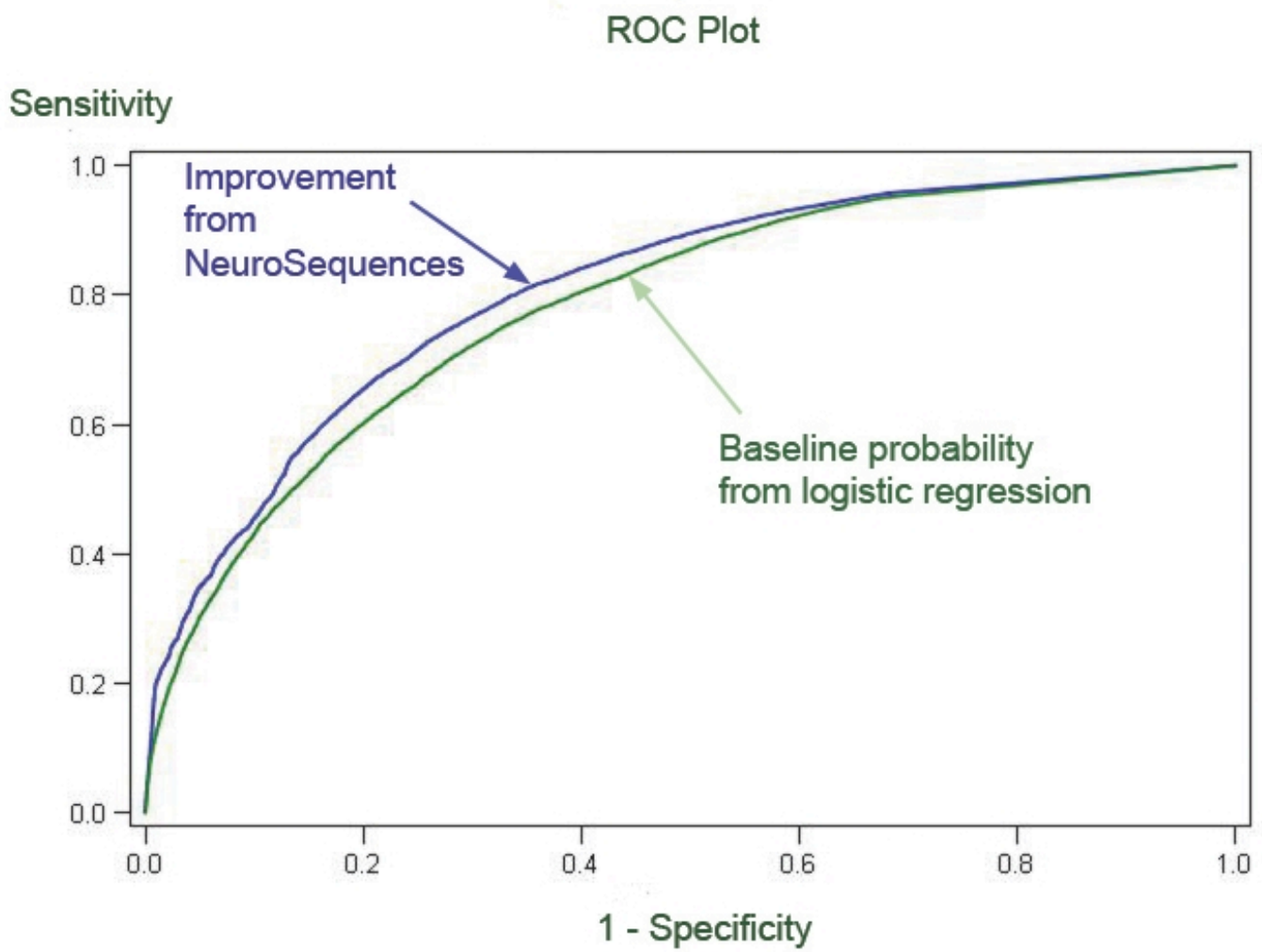
- Decision
- Ranking
- Estimation

Baseline Probability

- Auto Search
- Logistic Regression
- Probability Estimation



ROC



MLP

4 major concerns

- Big Scale/Rare Events
- Local Minima
- Generalization
  (Over Training)
- Slow Speed

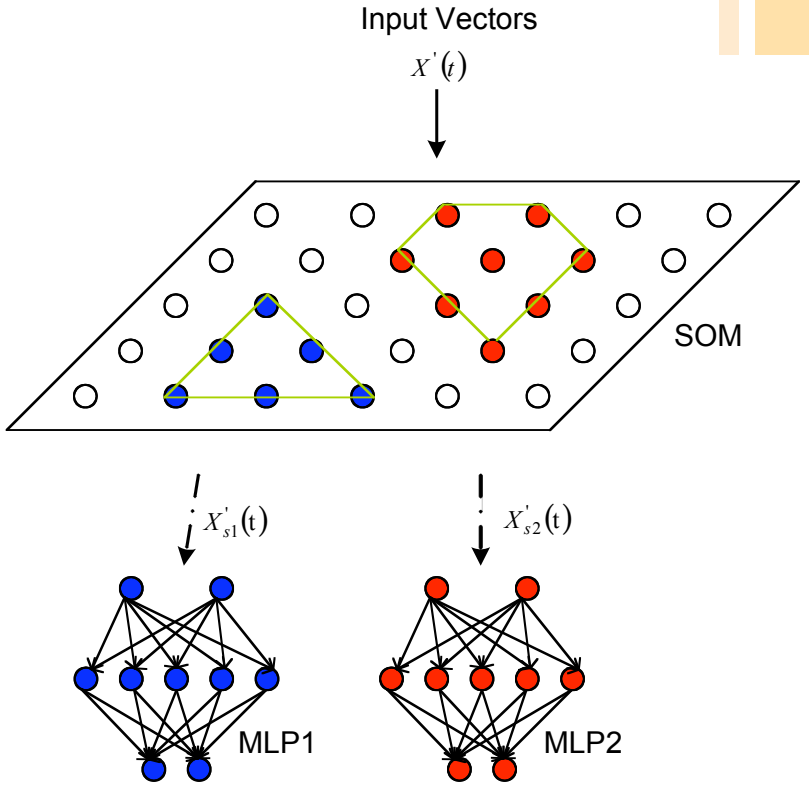
Solutions

- NeuroSequences
  (FFC 2008, DC, USA)
- Retreat & Turn Search
  (ISNN 2008, Beijing, China)
- Monitoring Parameters
  (AIQED 2009, Cambridge, UK)
- Lo-Shu Data Sampling
  (AIQED 2009, Cambridge, UK)

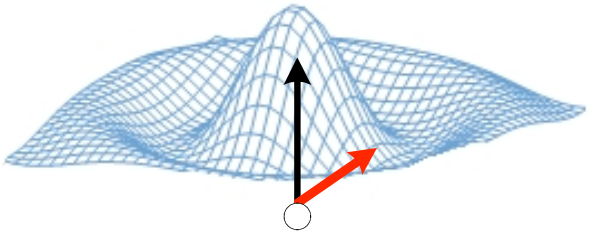
Clustering Analysis

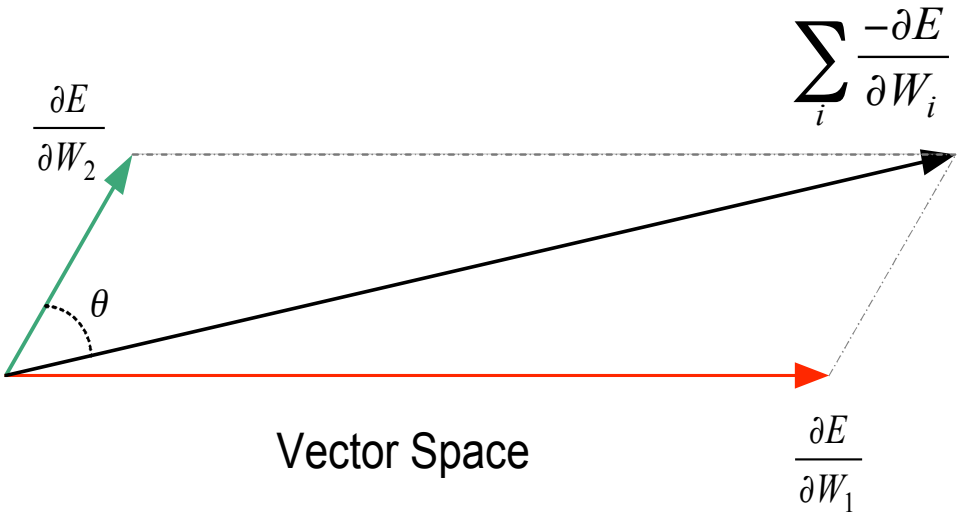
Divide and Conquer

- SOM subgroups
- reduce the complexity of the problem
- relationships between the SOM neurons is very useful for data exploration since the *topographic error* is minimized.



Escaping local minima - Retreat & Turn

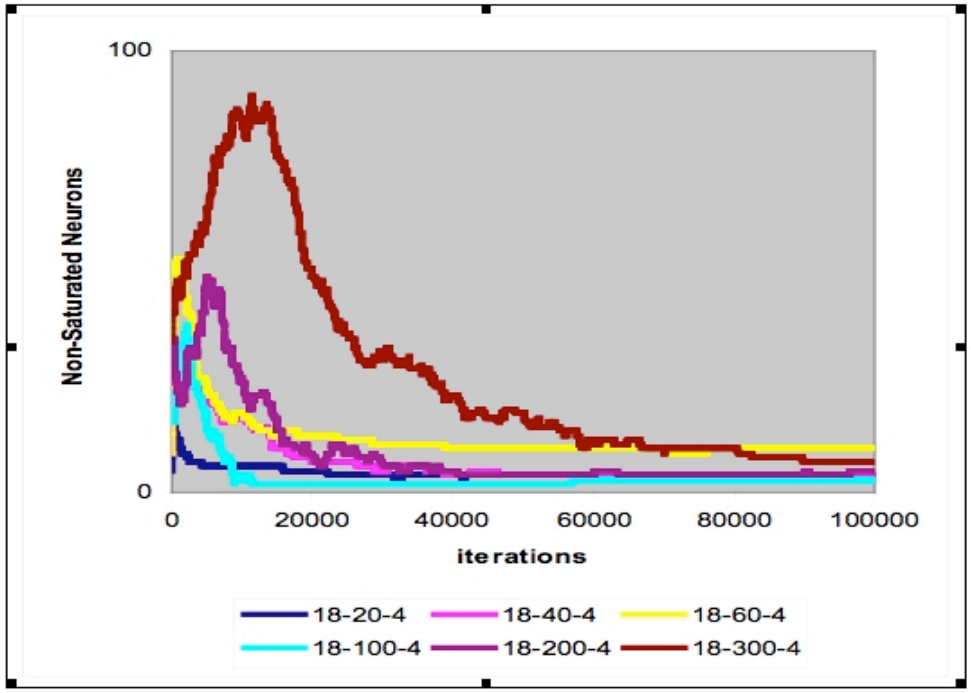


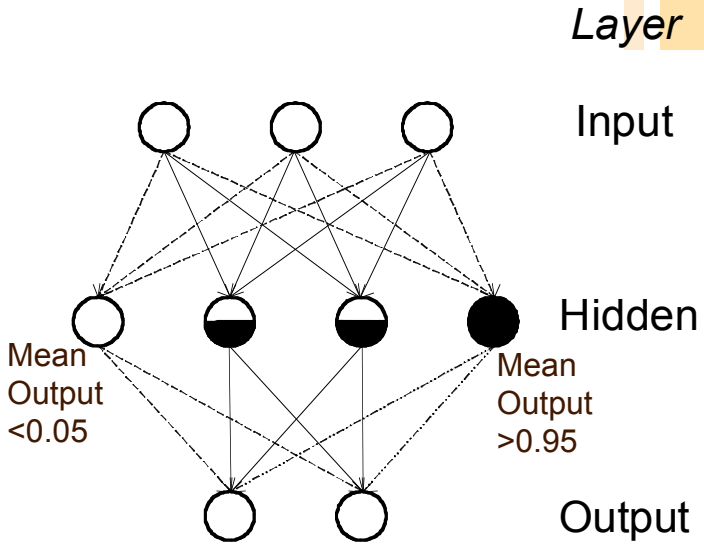


If we can't move the mountain, at least we can make the road turn.

When error increase, retreat and turn can search a better path to descend on the MLP's error surface.

Monitoring Free Parameters





The changes on the neurons, through time, can form the path to classify all the objects. Classifying these objects is embraced by the movement of neurons' outputs, between the state of inhibitory (or 0, Yin) and excitatory (or 1, Yang). This process can continue till a state of harmony has been reached (Laozi, Chapter 42, *Tao Te Ching*).

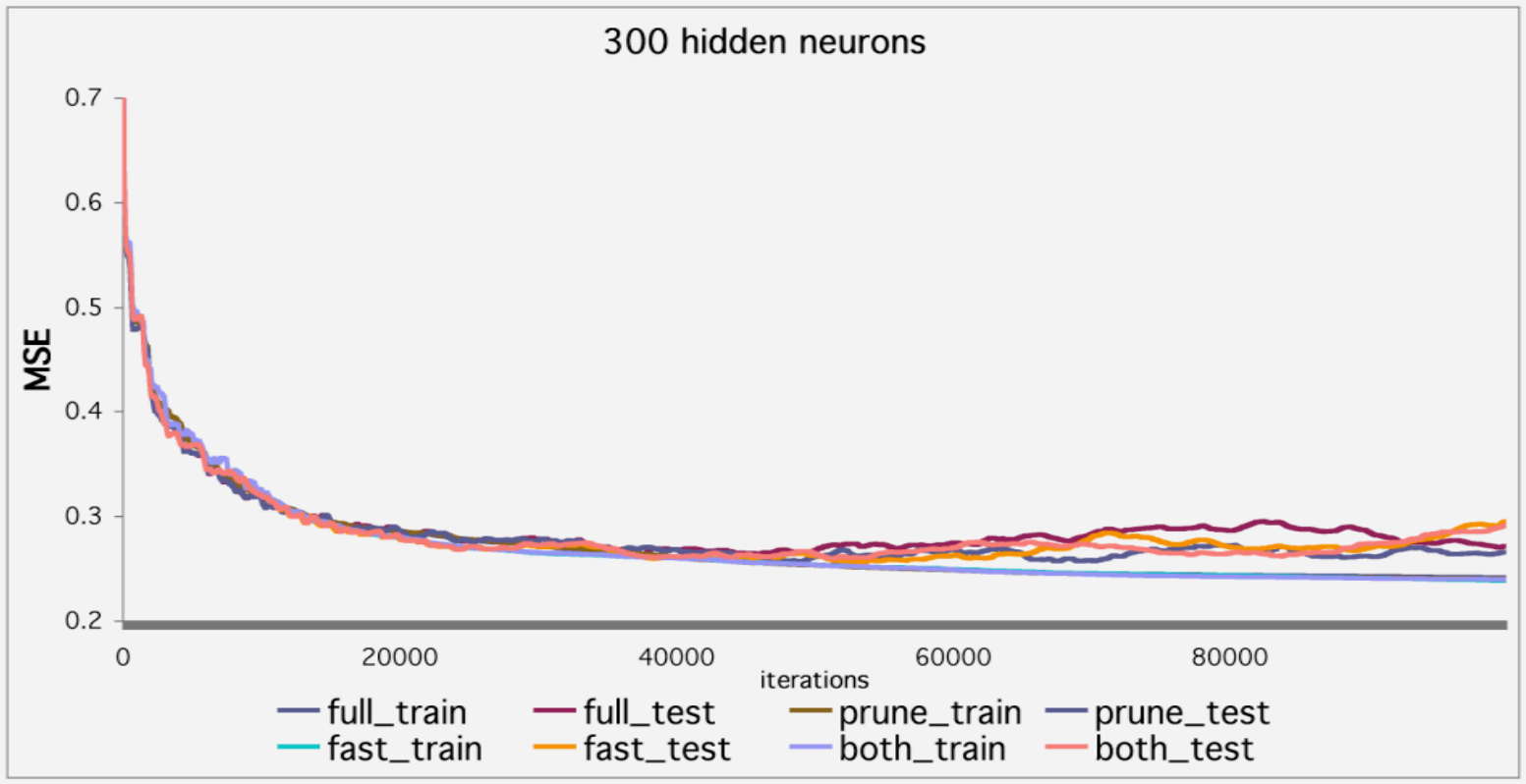
Lo-Shu Data Sampling

- **Potentially saving 2/3 of training time:**
  - **9 Subsets**: Systematically or randomly generated.
  - **Each shift**: Use only 3 subsets of data.
- **Lo-Shu data sampling can save time before convergence:**
  - Average time for at least one subset to reappear is less than two shifts away
  - Average time for at least two subsets to reappear is less than three shifts away

{4, 9, 2} (戴九)
{8, 1, 6} (履一)
{4, 3, 8} (左三)
{2, 7, 6} (右七)
{9, 5, 1}
{3, 9, 7} (二四為肩)
{2, 5, 8}
{4, 5, 6} (六八為足)
{3, 1, 7} (五在中間)

9 shifts based upon 9 groups of subseted data

Training & Validation Examples



USGS Land Use/Cover Categories: segmented images of four classes