# NeuroSequences:

## A Method to Improve Neural Networks in Healthcare Predictive Modeling

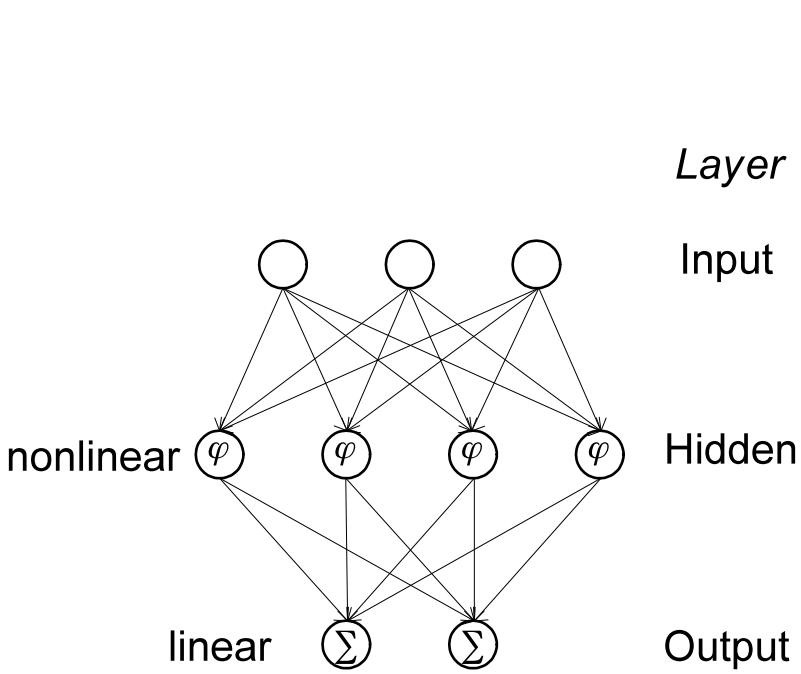
By Hung-Han Chen, Ph.D.

# Predictive modeling for Healthcare

- "a set of tools used to stratify a population according to its risk of nearly any outcome"[3]

- risk identification often requires careful balancing of sensitivity and specificity.

# Multi-Layer Perceptrons (MLP)




- Error Back Propagation through layers.
- Minimizing square of error by using Gradient descent.
- **Gradient descent** Method is an optimization algorithm (Hill-climbing technology)
- Intrinsic Nonlinearity, Computational Simplicity, and Resemblance to neurons of human brain.

# What can MLP do for you?

- They can be used as a machine learning method
  - **Pattern Recognition**
    - Assign new inputs to one of a number of classes
  - **Regression**
    - Predict the outputs as values of continuous variables
- They can also be used where traditional techniques do not work well

# BP Critics & remedies

Critics

- Trapped in local minima.
- Problems with scale-up.
- Not integrated with cost function.
- Long time to train.
- Over-fitting.
- Catastrophic unlearning.
- Mysticism.

Remedies

- Change structure
- Change learning algorithm

Or cooperate with others

- Linear equations
- Probability elements
- Decision Tree elements
- Fuzzy elements
- Clustering
- Bayesian

# Solution from Neuroscience

Neocortical Memory [6]:

- The neocortex stores *sequences* of patterns
- The neocortex recalls patterns *auto-associatively*
- The neocortex stores patterns in an *invariant form*
- The neocortex stores patterns in a *hierarchy*

Besides MLP, If we can add

- auto-association with *unsupervised learning*
- invariant form with *clustering analysis*
- *more layers* to the hierarchical structure

Answer:

SOM + MLPs

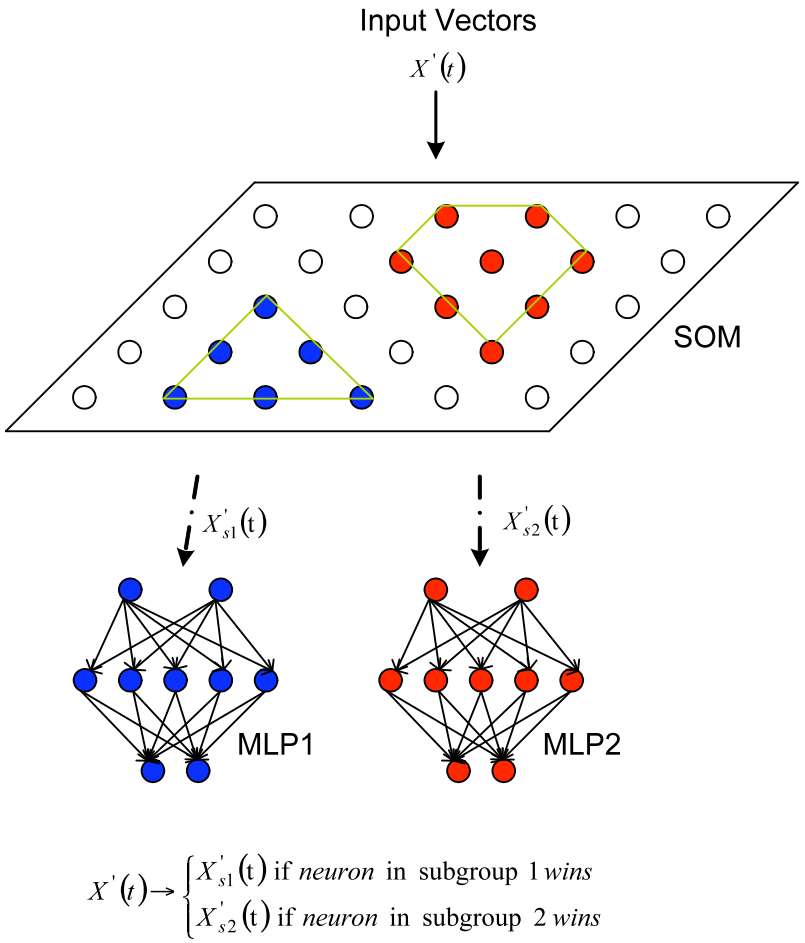
# Self-Organized Map (SOM)

- one of artificial neural models for the brain
- experimentally found "ordered maps" in the cortex layers
- Unsupervised learning and a relative to Clustering Analysis
- preserve the topological properties of the input space
- **Data Exploration and Visualization**

# Concept of Neurosequences

*Divide and Conquer*

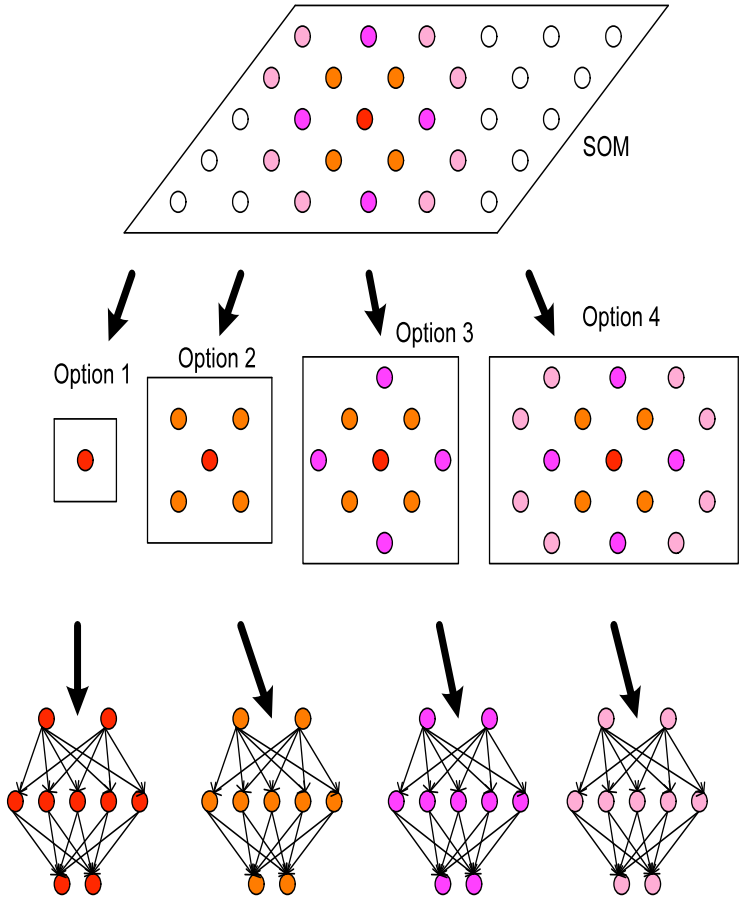
- SOM subgroups
- reduce the complexity of the problem
- relationships between the SOM neurons is very useful for data exploration since the *topographic error* is minimized [14]



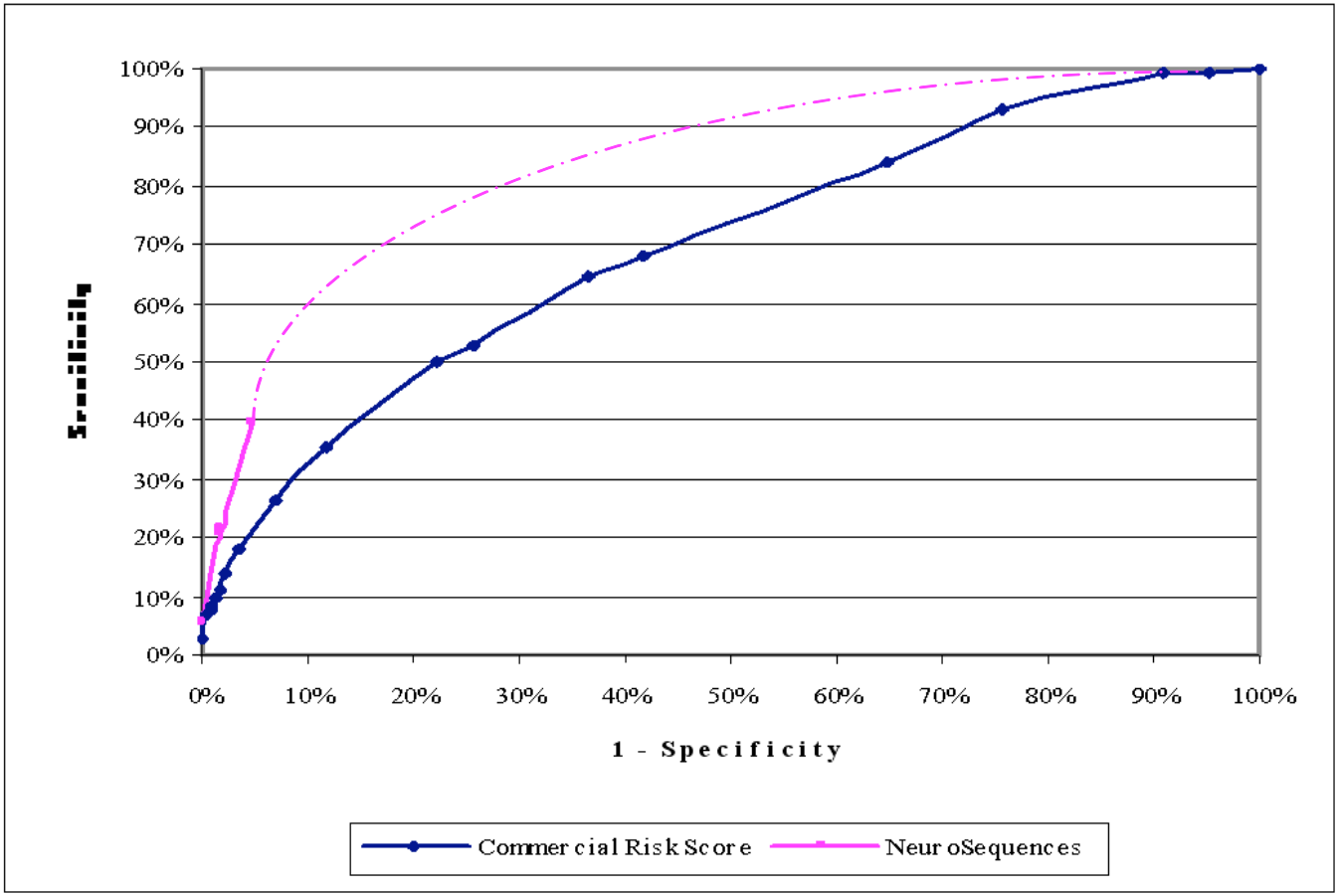
# Impacts of NeuroSequences

*Controlling Data Flow*

- SOM subgroups
- drive the solution to a specific target region
- RED color to identify the accuracy of a SOM neuron for 40% or above. ORANGE for 30%, PURPLE for 20%, PINK for 10%, and the rest is uncolored.



# ROC curve


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
1 - Specificity
Commercial Risk Score
NeuroSequences

# 3 month inpatient risk

Target:

2.4 Million Members, binary output, prevalence ~ 1.3%

Data Source:

Member Table: Enrollment data.

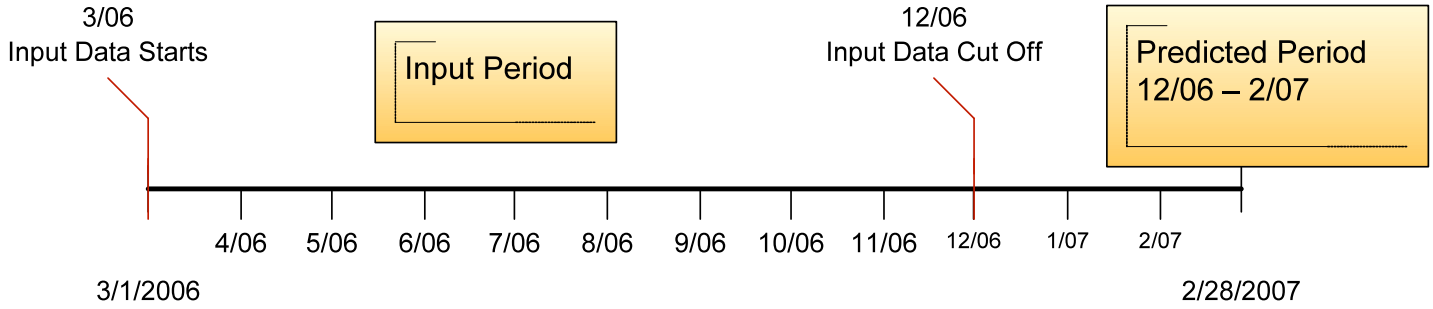
Service Table: Medical Claims, Pharmacy claims.

Model's Inputs:

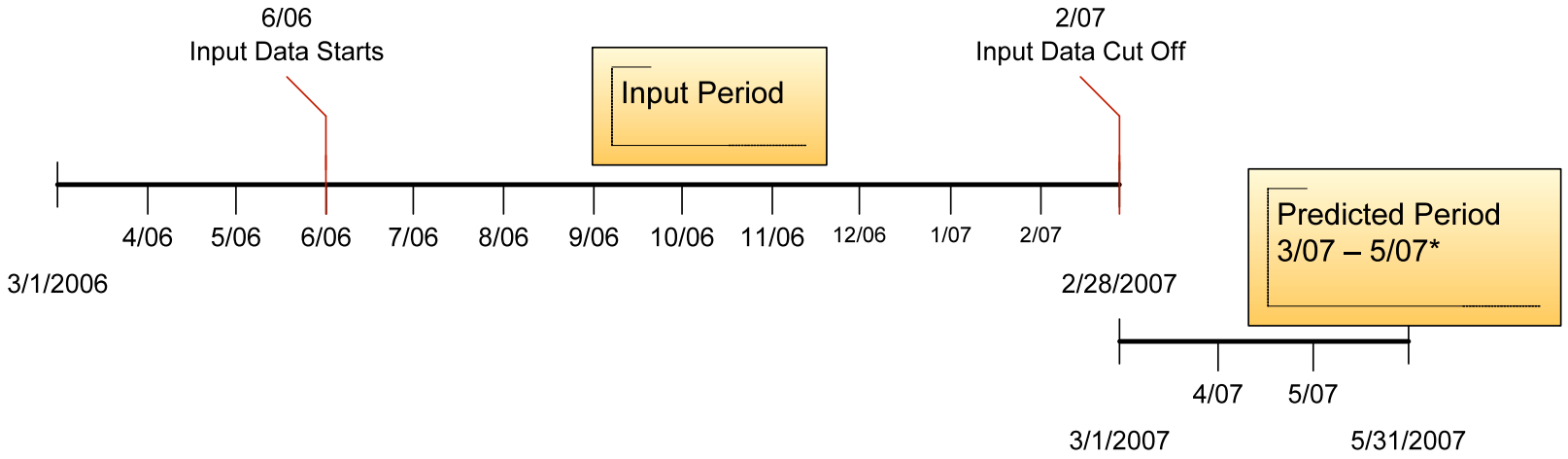
Age, Gender, ICD9, CPT4, NDC, CCS Groups etc.

# Analysis Periods

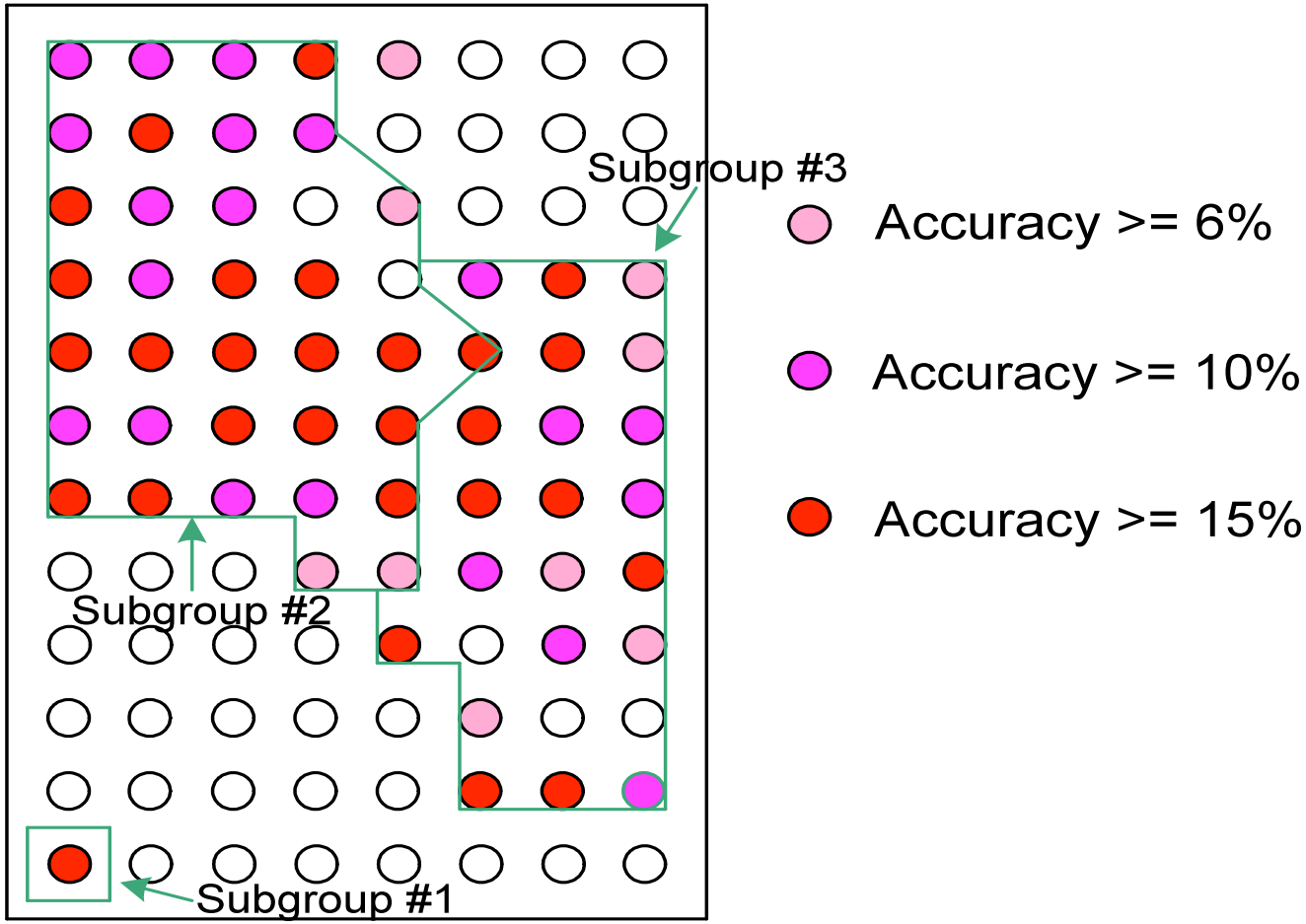
*NeuroSequences Training*




*NeuroSequences Validation*



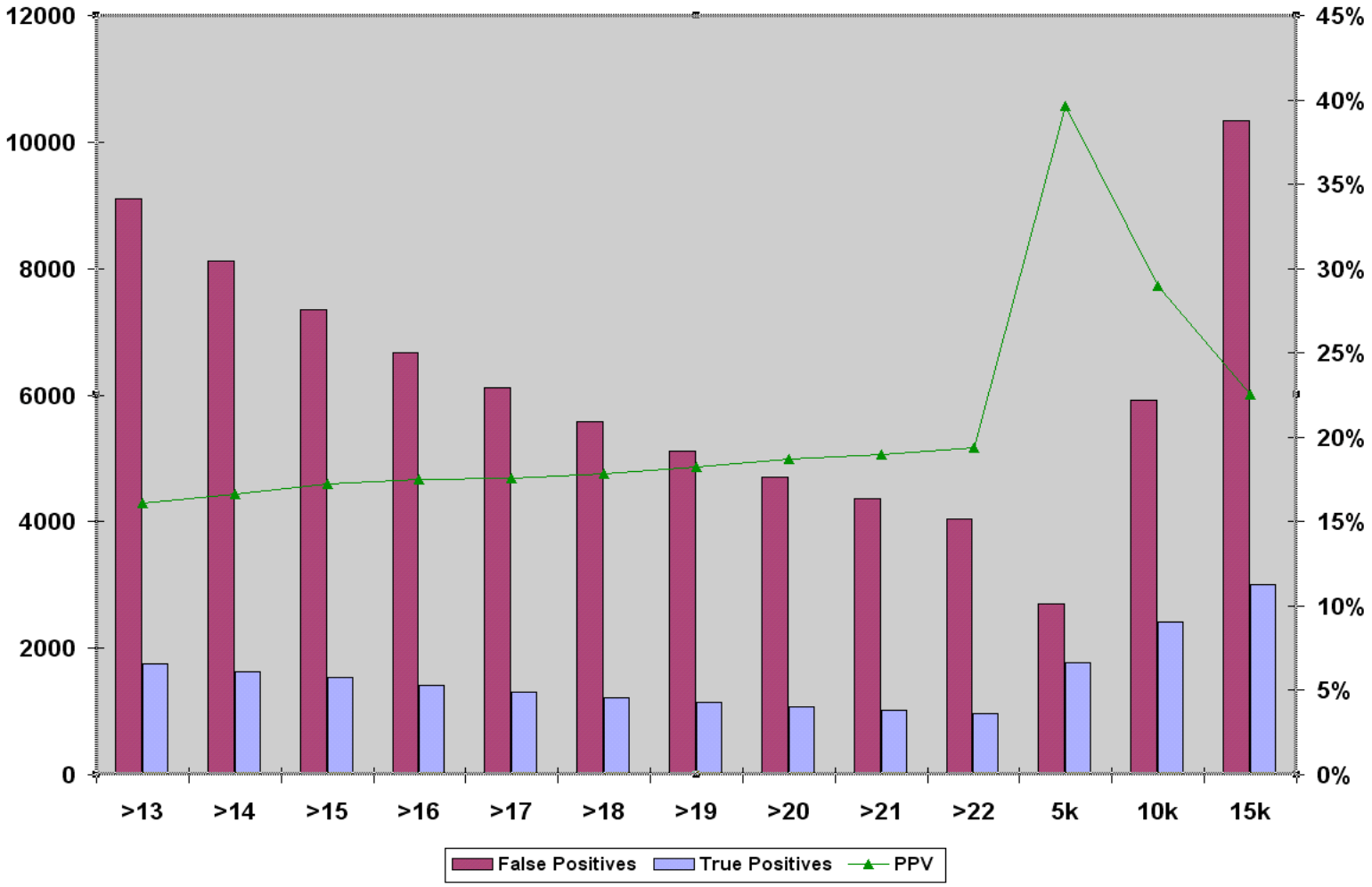
# SOM subgroups



# Result Comparison

| **Commercial Software** | True Positives | False Positives | Total | Sensitivity | PPV |
|---|---|---|---|---|---|
| >13 | 1748 | 9099 | 10847 | 5.31% | 16.12% |
| >14 | 1619 | 8124 | 9743 | 4.92% | 16.62% |
| >15 | 1531 | 7346 | 8877 | 4.65% | 17.25% |
| >16 | 1416 | 6679 | 8095 | 4.30% | 17.49% |
| >17 | 1302 | 6121 | 7423 | 3.96% | 17.54% |
| >18 | 1213 | 5582 | 6795 | 3.69% | 17.85% |
| >19 | 1143 | 5116 | 6259 | 3.47% | 18.26% |
| >20 | 1081 | 4695 | 5776 | 3.29% | 18.72% |
| >21 | 1019 | 4356 | 5375 | 3.10% | 18.96% |
| >22 | 973 | 4042 | 5015 | 2.96% | 19.40% |
| ***NeuroSequences*** | | | | | |
| *5k Model* | 1778 | 2708 | 4486 | 5.40% | 39.63% |
| *10k Model* | 2412 | 5913 | 8325 | 7.33% | 28.97% |
| *15k Model* | 3004 | 10336 | 13340 | 9.13% | 22.52% |

# Comparison Chart



# Conclusion

- NeuroSequences utilize two neural technologies to build predictive models.
- "Divide and Conquer" and "Control Data Flow" help Neurosequences to perform better.
- Remain computational simplicity and resemblance to neural system.

# Questions and Answers

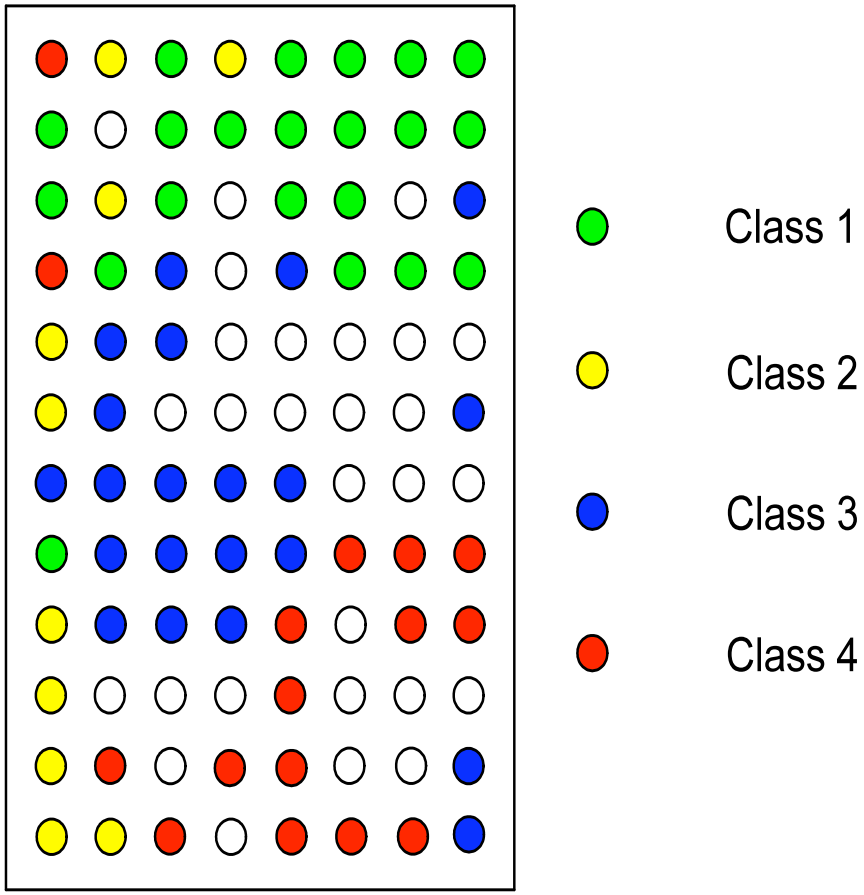
# Appendix

An image segment problem

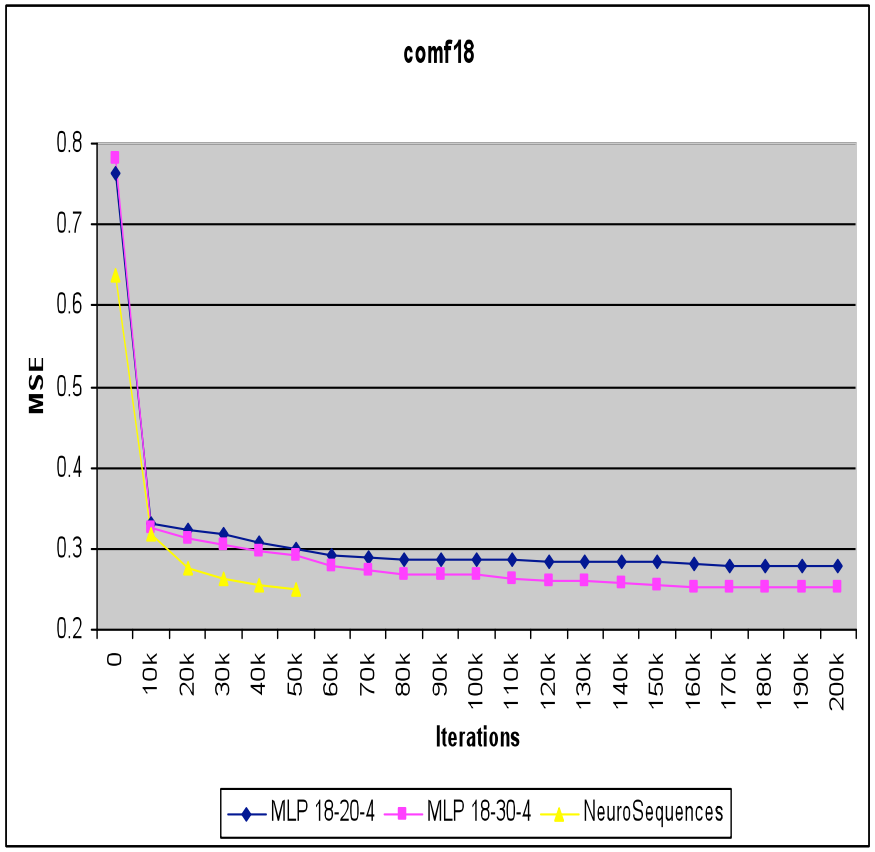
# USGS Land Use/Cover Categories

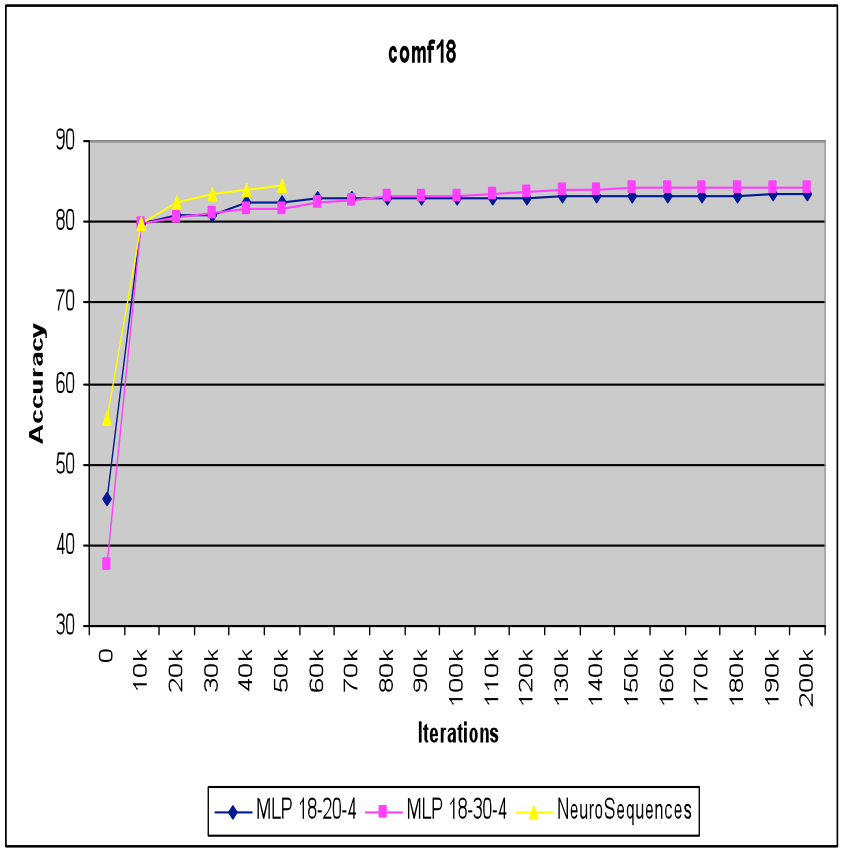
- Each segmented region is separately histogram equalized to 20 levels. Then the joint probability density of pairs of pixels separated by a given distance and a given direction is estimated. We use 0, 90, 180, 270 degrees for the directions and 1, 3, and 5 pixels for the separations. The density estimates are computed for each classification window. For each separation, the co-occurrences for for the four directions are folded together to form a triangular matrix. From each of the resulting three matrices, six features are computed: angular second moment, contrast, entropy, correlation, and the sums of the main diagonal and the first off diagonal. This results in 18 features for each classification window.
- 4 classes, even-distributed.

# SOM subgroups



# MLP vs NeuroSequences



# Subgroups in NeuroSequences