# Neural networks for variable star classification

Vasily Belokurov, IoA, Cambridge

Supervised classification
Multi-Layer Perceptron (MLP)
*"Neural Networks for Pattern Recognition"* by C. Bishop

• Unsupervised learning and detection Self-Organizing Map (SOM) "Self-organizing maps" by T. Kohonen

Bayesian methods for learning algorithms
*"Information Theory, Inference and Learning Algorithms"* by D. MacKay

• **Training set** - a library of classified lightcurves. datasets from Hipparcos, OGLE etc Archive integration

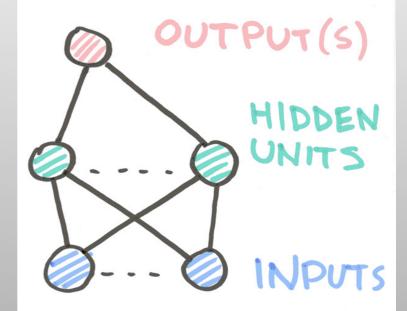
• Feature extraction - lightcurves need to be parameterized Tools for spectral analysis of sparsely sampled data

• P(C|x) - probability that the variability class is *C* given measured features *X* 

Find **P(Clx)** - nonlinear function in many dimensions Neural Networks can approximate this mapping

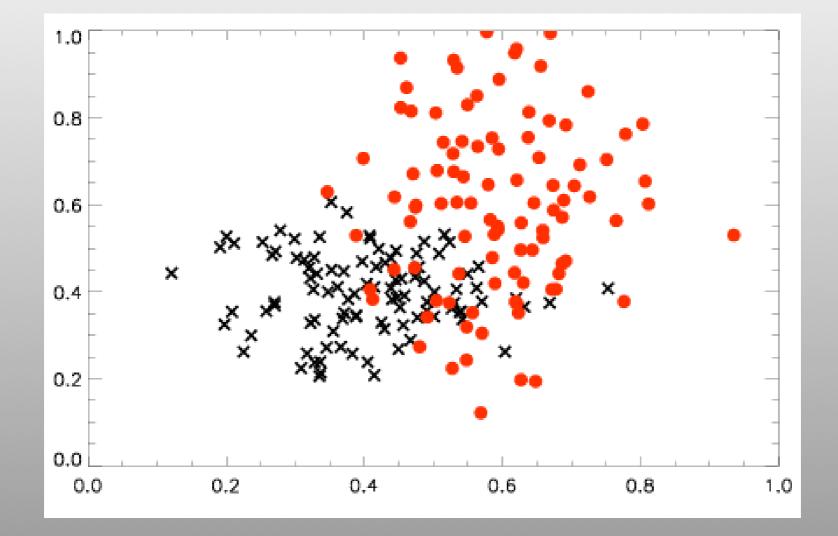
#### **Feed Forward Networks**

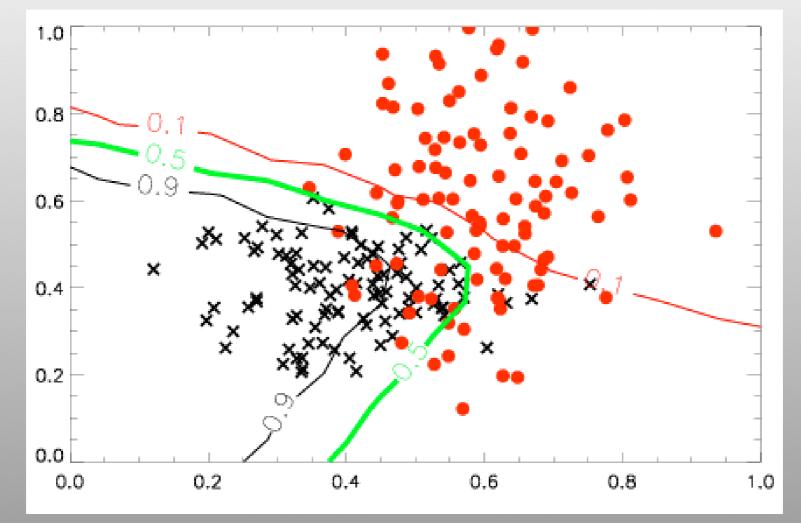
- Each connection has a *weight* assigned
- Network can have several *layers* of hidden (processing) units (neurons)
- Each neuron **receives** a *weighted sum* of values from the lower layer
- Each neuron then sends this linear sum transformed with a nonlinear *activation function*
- The output is calculated and can be compared to the *target*



#### Error back-propagation

- *Error function* quantifies the difference between output and target
- Derivatives (with respect to the weights) of the error function are calculated by propagating the 'errors' from higher layers





Stuttgart Neural Network Simulator

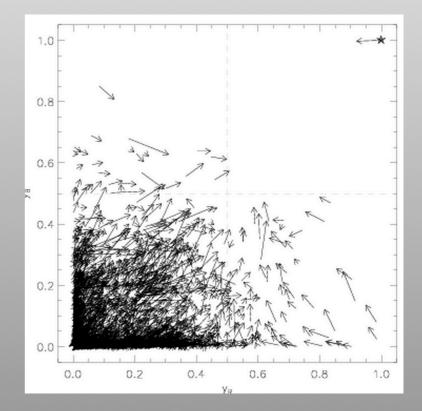
#### Advantages

- Distribituted processing of all available information
- Approximates complicated decision boundaries in many dimensions
- Novelty detection

#### Example

- MACHO LMC data
- 5 inputs (extracted from each lightcurve)

• Identifies 1 microlensing event in the tile of **2500** lightcurves *Belokurov, Evans & Le Du 2003* 

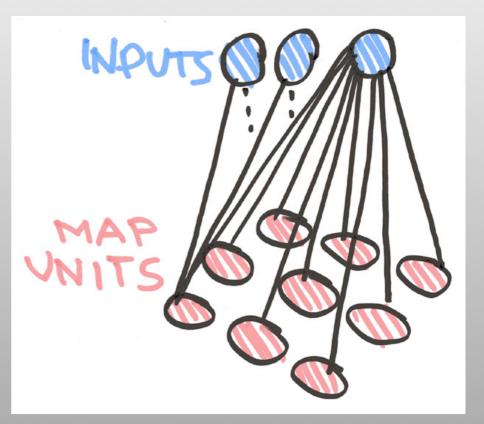


#### Desirables

- Visualize data in more than 3 dimensions
- Preserve the topology

#### SOM

• The weights of each map node define a *reference vector* of the same dimensionality as the input



• The node with the smallest Euclidean distance between its reference vector and the input defines *response* of the map to the input

### Learning

1. Find the *winner* 

2. Update the winner to minimize the distance

3. Allow nodes in the vicinity of the winner to be updated as well (defined by the *neighborhood kernel*)

4. Kernel monotonically decreases with time

### **Quality of learning**

- Quantization error
- Distortion measure

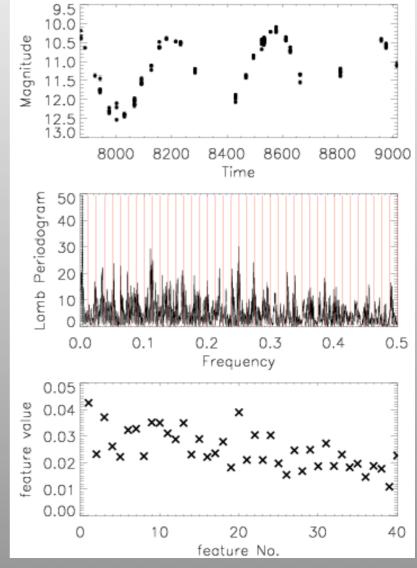
#### Calibration

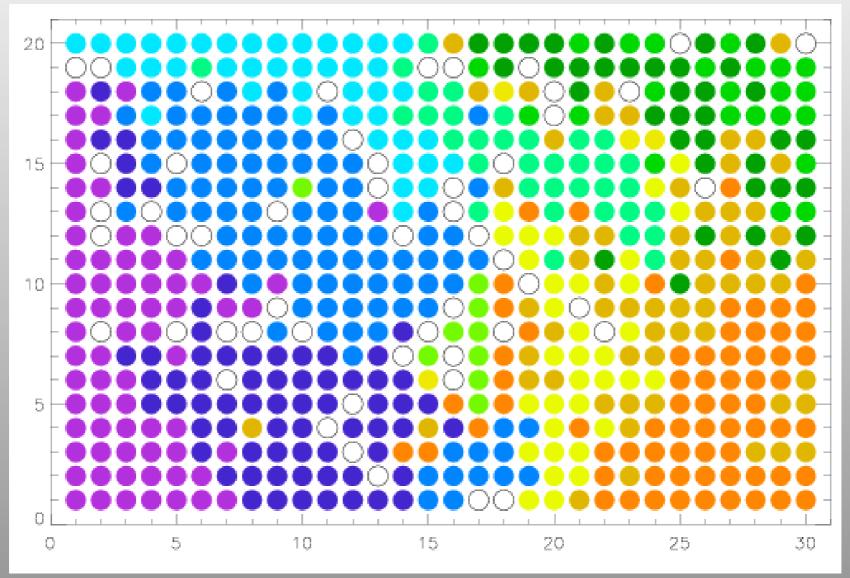
Set reference points with manually analyzed data

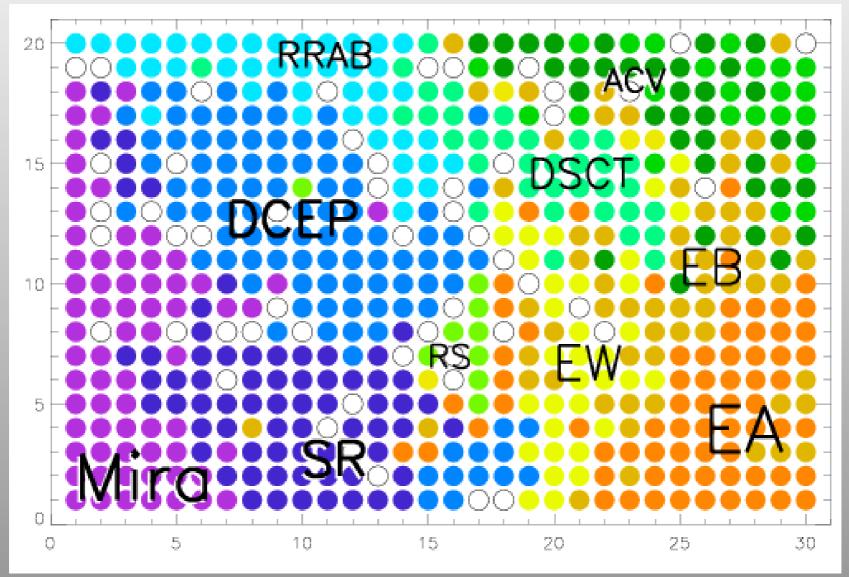
#### Example - Hipparcos Variables 52 input features in total

- Lomb Periodogram integrated in 40 bins and normalized
- 5,15,25,35,45,55,65,75,85,95 percentiles with median magnitude subtracted
- Ratio of magnitude above/below median
- V-I
- Color evolution?...

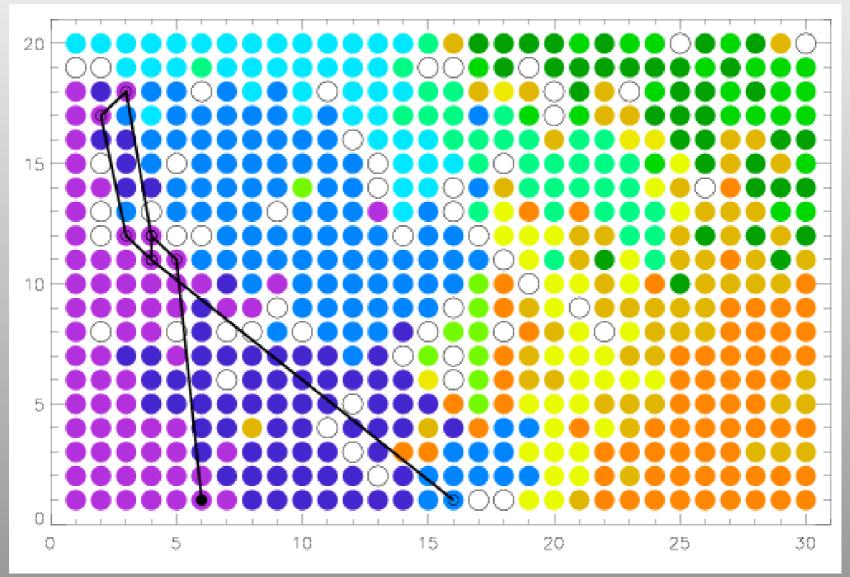
**Dataset** - more than **2500** 'solved' Hipparcos variables **Software** - SOM\_PAK (free)

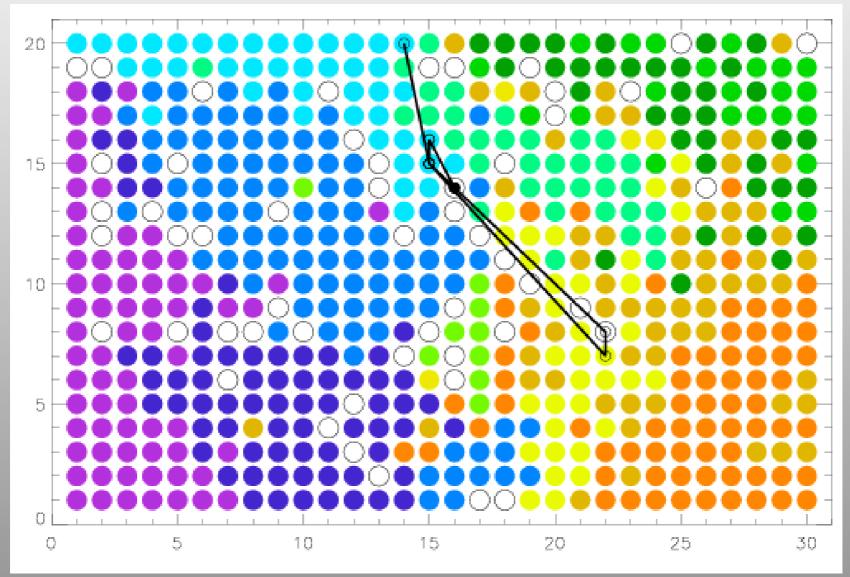


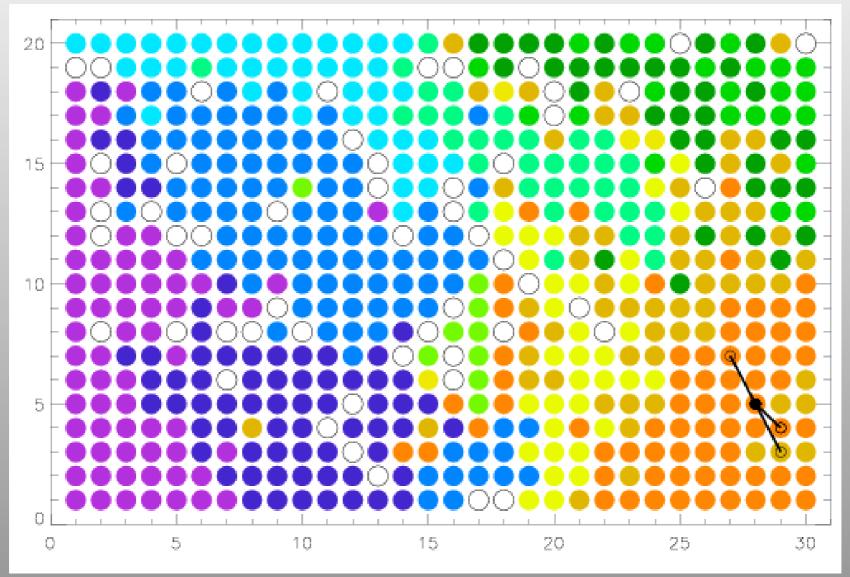




GAIA Science Alerts & Variable Star WGs







#### GAIA Science Alerts & Variable Star WGs

ACV ACV ACV SPB ACV SPB SPB SXARI BCEP ACV SXARI ACV ACV DCEP BRAB CEP(B) BRAB RRAB RRAB RRAB SXPHESXPHE FLL SRA DEEP RRAB DEEP RRC DSCT SPB GCAS RRAB RRAB RRAB RRAB RRAB DCEPS DSCT DSCT DSCT DSCT ELL: ELL SBA DCEP DCEPS DCEP CEP(B) DCEP DCEPS CEP RRAB DCEP RRC SXPHE EW DSCT DSCT DSCT DSCT DSCT DCEP DCEP CEP EW/KW EW EB EB DOEP DOEP DOEP ERGS EA DCEP DCEP DCEP FKCOM RS EA EB BBC EB. EB EA EB EB EW/KW EW EW/KE EA/DM SR DCEPS EB EW/KWEW/KW EB EA/DM EA/DM EB/KE EB/K DCEP EW/KW EB EW/KW EW SR SARV SR DCEP DCEP EW/KWEW/KWEW/KW EA/SD\_EA/SD\_EA/SD\_EA/SD SRB EA/DW EA/SD EA/DS EA/S CWA EW/KWEW/KWEW/KW EA

### Unsupervised learning with Self-Organizing Maps

#### GAIA Science Alerts & Variable Star WGs

**Organizing Maps** TAE HAAE HAAE HAAE HAAE HAAE HAAE HAAE RAAB RRAB RRAB RRAB RRAB RRAB RRAB STAR SXPHE EB ACV ACV ACV ACV SPB ACV SPB SPB ACV RRAB RRAB RRAB SXPHESXPHE CWA DCEP: DCEP RRAB CEP(B) RRAB DCEP RRAB DCEP CEP DCEP CEP(B) DCEP DCEP RRAB RRAB RRAB DSCT DSCT RRC DSCT SPB RRAB RRAB RRAB RRAB RRAB DCEPS DSCT DSCT DSCT SR CEP(B) DCEP DCEPS CEP DCEP DCEP DCEP DCEP RRAB RRAB RRAB DCEP RRC SXPHE EW DSCT DSCT DSCT DSCT **BB EW/KW** DCEP FW SBC DCEP DCEP DCEP DOEP DOEP DOEP EBGS DCEP DCEP DCEP DCEP DCEP DCEP FKCOM DCEP EB EB EW/KWEW/KW EB SARV EW DCEP DCEP DCEP EW/KWEW/KWEW/KWEB/DW DCEP DCEP DCEP DCEP EW/KWEW/K EA/SD EA/SD EA/SD EA/S SRB CWA EW/KWEW/KWEW/KW

Unsupervised learning with Self-

SPB SPB SXARI ACV ACV SXARI ACV RRAB RRAB RRAB SXPHESXPHE ELL ELL DCEP RRAB CEP(B) RRAB ELU RRAB DCEP CEP DCEP CEP(B) DCEP DCEP RRAB RRAB RRAB DSCT DSCT RRC DSCT SPB GCAS GCAS RRAB RRAB RRAB RRAB RRAB DCEPS DSCT DSCT DSCT DCEP DCEPS DCEP DCEP DCEP CWB RRAB RRAB RRAB DSCT DCEP DCEP DCEP DCEP RRAB RRAB BRAB DOEP BRO SXPHE EV DOEP DOEP DOEP CEP DCEP RR EW/KW DOEP DCEP DCEP DCEP DCEPSDCEPS ACYG DOEP DOEP DOEP DOEP DOEP DOEP DOEP ERGS DCEP DCEP DCEP DCEP DCEP FKCOM EB DCEPS RRC RRAB DCEP EB EW/KW **ISR** EW DCEP DCEP DCEP EW/KWEW/KWEW/KWEB/DW SR DCEP DCEP DCEP DCEP DCEP EW/KWEW/KWEW/KW EA/SD EA/SD EA/SD EA/S SB SR ISB ISR ĭSRÌ DCEP DCEP GWA GWA EW/KWEW/KWEW/KW EA EA/AR EA/D/R EA/D EA EA/DW EA/SD EA/DS E GAIA Science Alerts & Variable Star WGs

50000 variables in the POINT-AGAPE dataset

#### **Strategy 1**

Train the map with the whole dataset available Calibrate with known (archived) variables

#### **Strategy 2**

Train the map only with known variables

For each new lightcurve find the location on the map

#### **Science Alerts**

Novelty detection - the black spots and walls - regions with low reference vector density

Numbers on the x and y-axis do **NOT** carry any (usual) meaning. However, SOMs can be used for <u>feature extraction</u> and then the (x,y) positions on the map can be analysed with MLPs.