# ECM3412/ECMM409
# Nature Inspired Computation

## Lecture 16

Neural Networks 4: Variations on Neural Networks and Unsupervised Learning

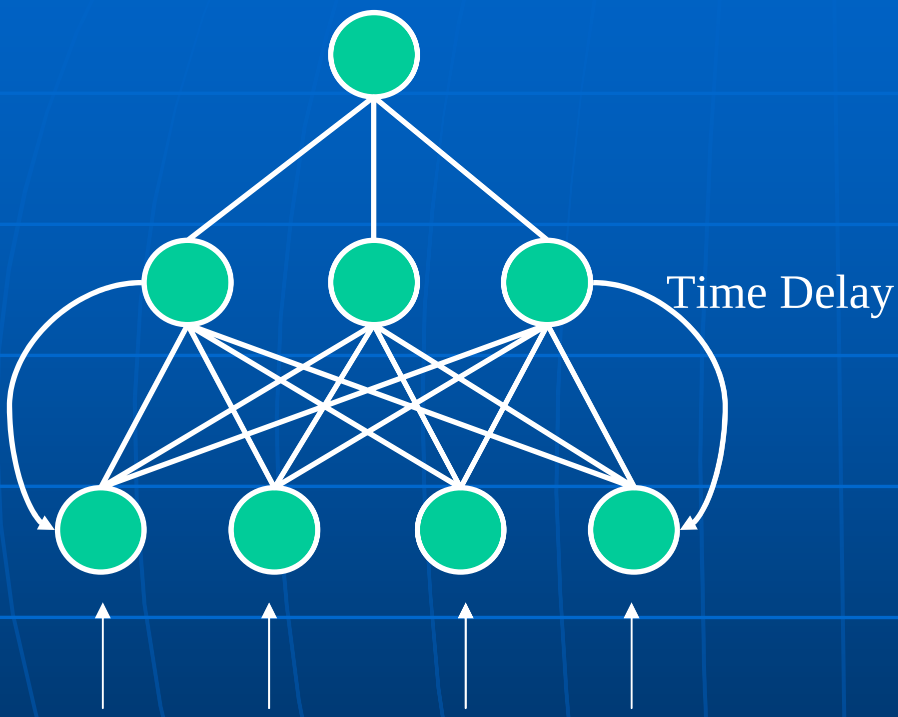
# Today

- Variations on neural networks
  - Recurrent Neural Networks
  - Radial Basis Function Networks (very brief)
- Unsupervised Learning
  - Kohonen neural networks

# Static vs Dynamic Problems

- Multilayer Perceptrons are useful for processing data that is static
- Many problems are dynamic or have a temporal component e.g stock market prediction
- These can be processed using recurrent neural networks

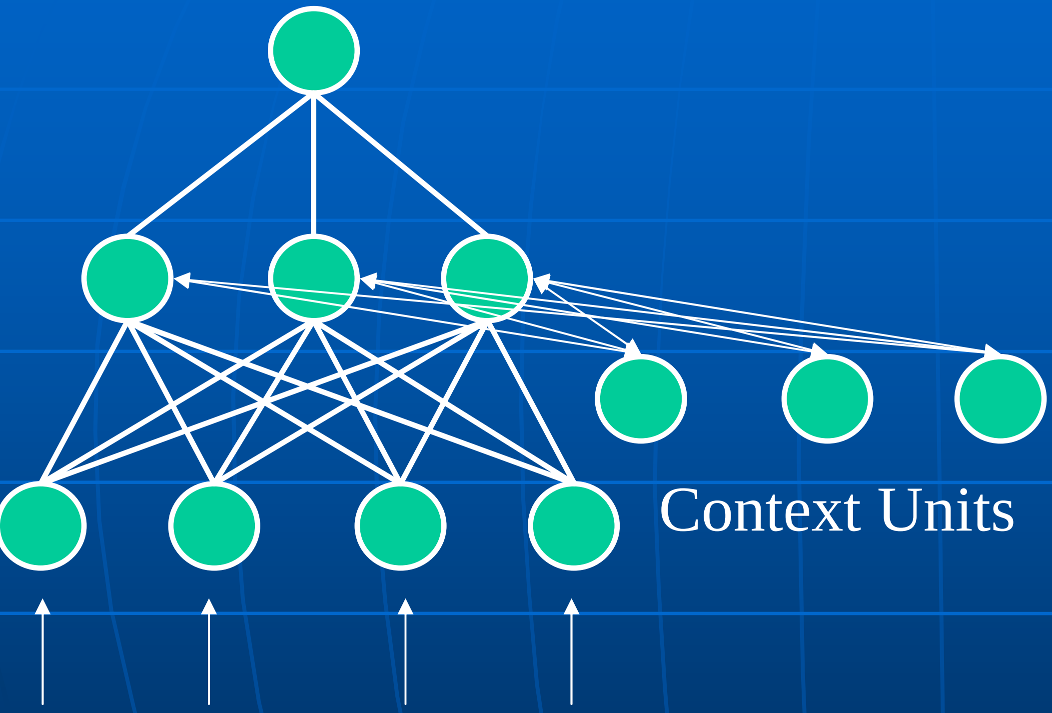
# Recurrent Neural Networks (RNNs)




One Example of an RNN
(there are many)

- RNNs have a similar structure to MLPs but also have connections from activity in previous timesteps
- In the example, the output of the hidden units at time t-1 is fed back to the inputs at time t1
- Learning is achieved through Backpropagation through time (BPTT)

# Recurrent Neural Networks (RNNs)




One Example of an RNN (there are many)

- A different example shows the use of context units which 'store' activations from previous timesteps
- Many different examples of RNNs exist – Hopfield and Jordan-Elman Networks

# RNN Summary

- Recurrent networks process not only current patterns, but have some short-term memory
- Have applications where sequences or remembering previous events are important – protein sequences, stock market values, robotics
- Generally more difficult to train than standard ANNs
- Can develop chaotic behaviour

# Radial Basis Function (RBF) Neural Networks

- RBFs are the next-best-used type of ANN
- They utilise a basis function (often a gaussian function) which is combined with the input data to give a response for each hidden unit
- Learning modifies the centres and widths of the basis function

# Self-Organisation

- Although the idea of the ‘Grandmother Cell’ is highly biologically implausible, the brain is organised into separate sections

- Each section can also have a local structure
  - In the auditory cortex, the cells are lined up in logarithmic order of frequency

- There are some aspects of brain organisation and learning which are based on self-organisation

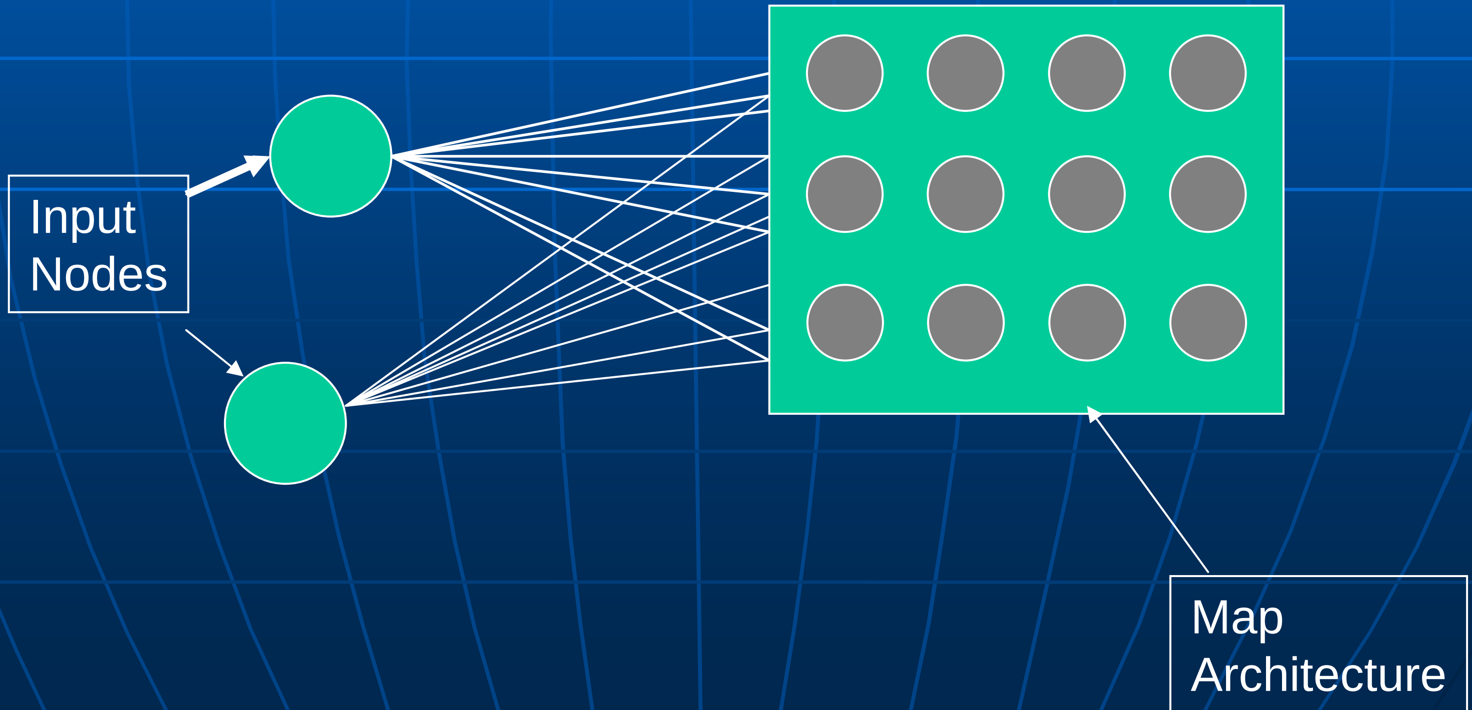
# Unsupervised Learning

- Unsupervised learning requires the network to learn the underlying structure in the data with no 'desired signal'

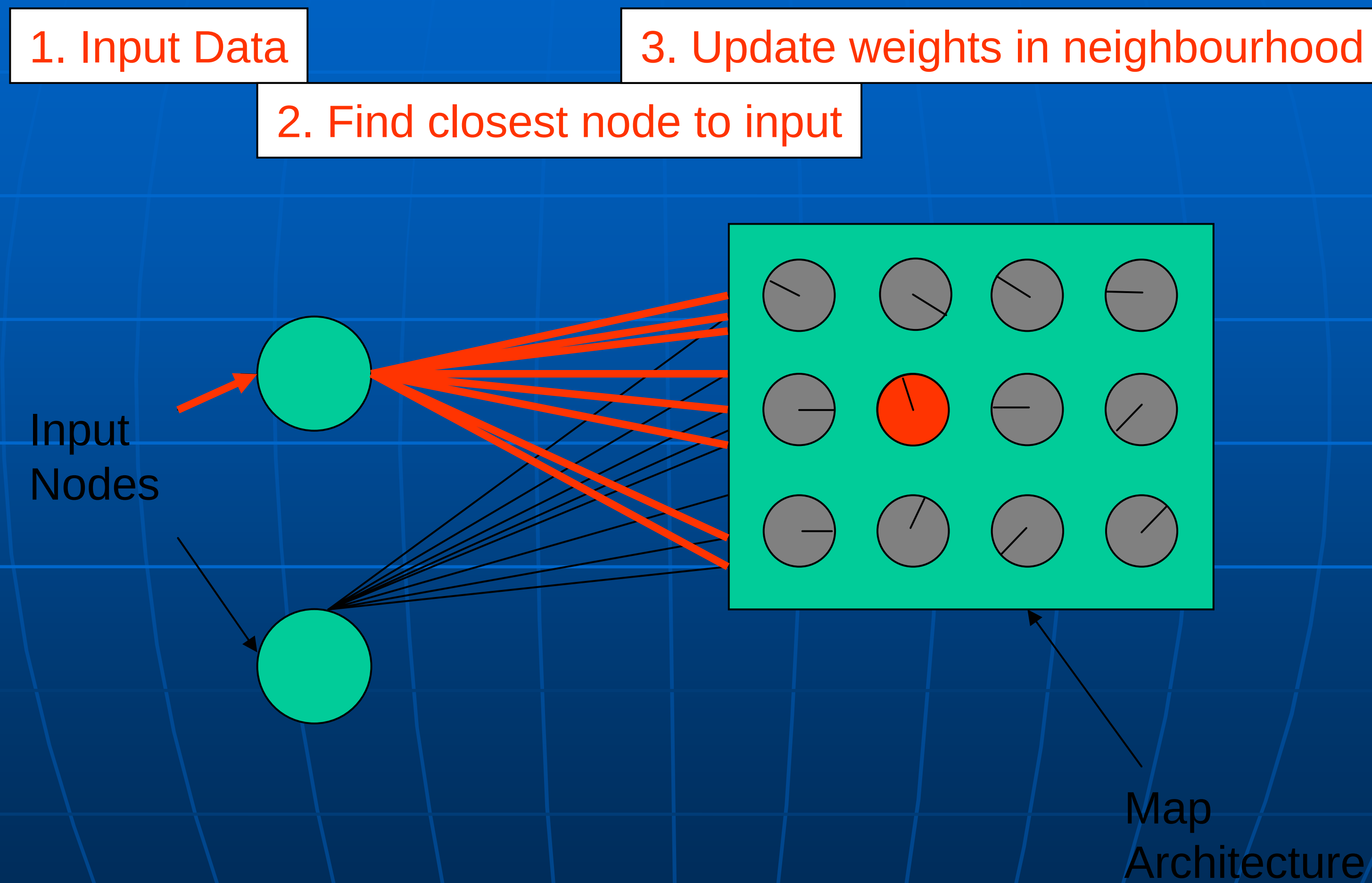
- It assumes that:
  - That input data that belongs to the same class share some common features
  - That the network will be able to identify the features across a number of data points

# Kohonen Self-Organising Map

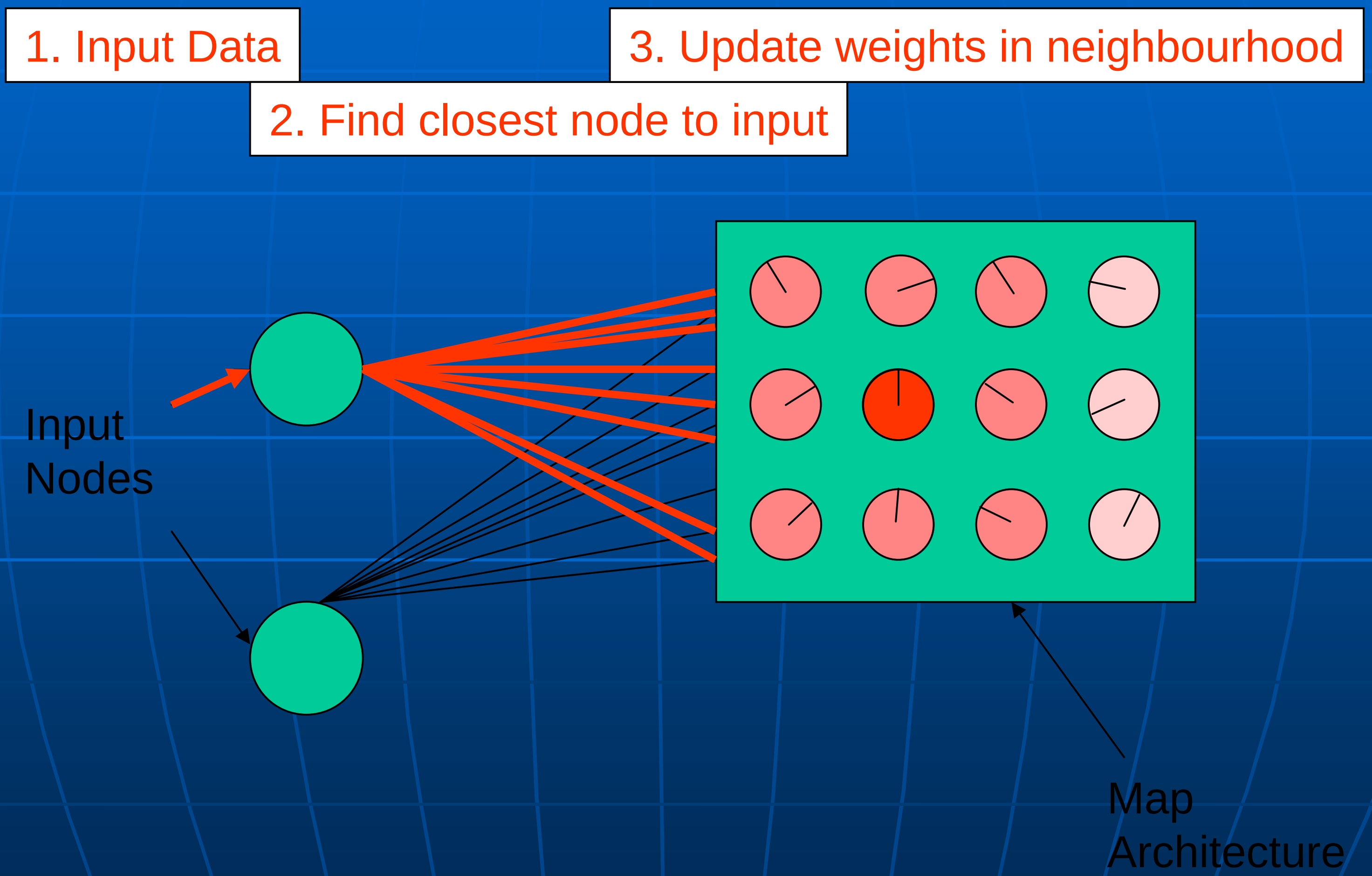
- Network architecture differs from MLPs
- There are as many input nodes as 'features' in the data
- Input nodes are connected to a 'map' of interconnected nodes
- Output is interpreted from the map



# Learning in a Kohonen Network



# Learning in a Kohonen Network



# Learning in a Kohonen Network II

- Initialise Network
  - Set weights to small random values
  - Define (large) neighbourhood size
- Present input ($x_i(t)$) to the network
- Compute the distance of each output node from the current input using:

$$d_j = \sum_{i=0}^{n-1} (x_i(t) - w_{ij}(t))^2$$

- Find the node closest to the input (Best Matching Unit – BMU)
- Adjust the weights of BMU and its neighbours using:

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t)(x_{ij}(t) - w_{ij}(t))$$

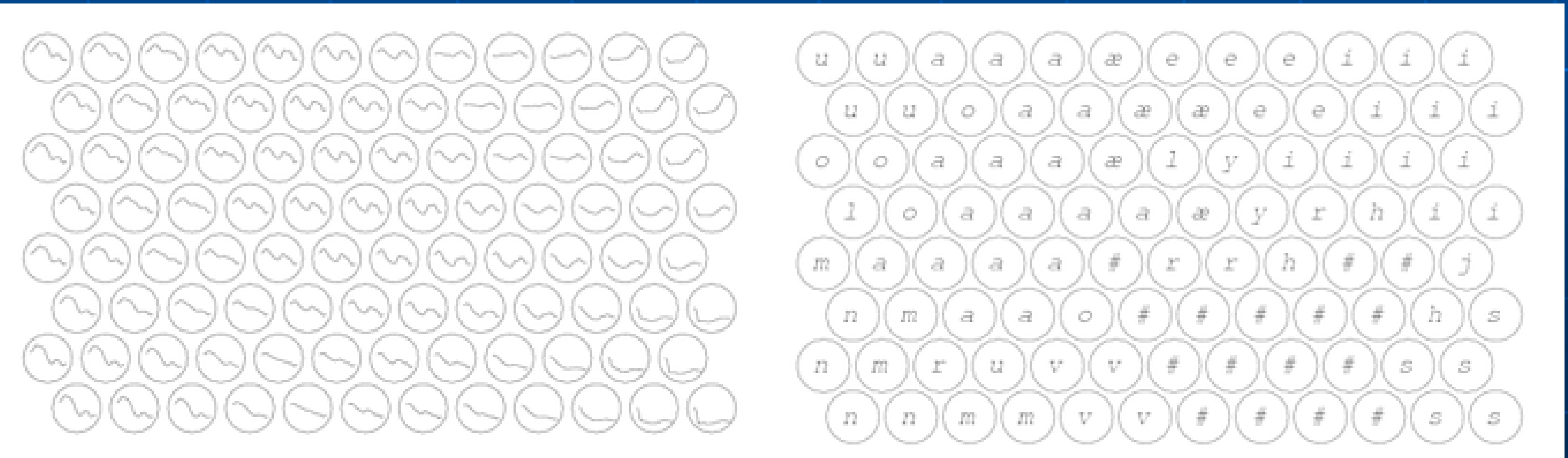
# Parameters

- $\eta$ is a learning rate which decreases with time
  - Network makes large changes early in training and smaller ones later on
- Neighbourhood size – initially large, decreases with time
  - Again, large spatial changes to map initially, decreasing over time

# Interpreting the Output

- Once the network has been trained, nodes next to each other in output space will be related

- The 'clusters' then have to be labelled (often manually) through observation of the input data and output node



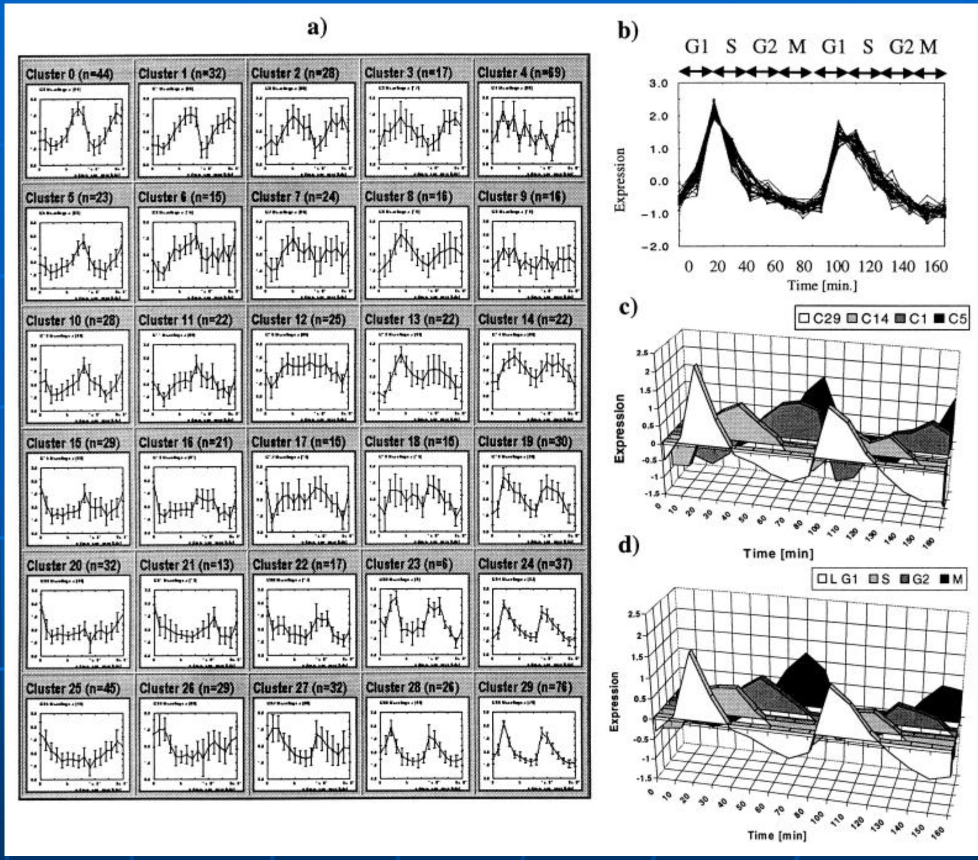
# Applications

- Kohonen originally used them for speech recognition
- Network learnt to classify phonemes from processed input waveforms of speech
- The phonemes are then presented to a rule-base which converts them into words in a word-processor

# Other Applications

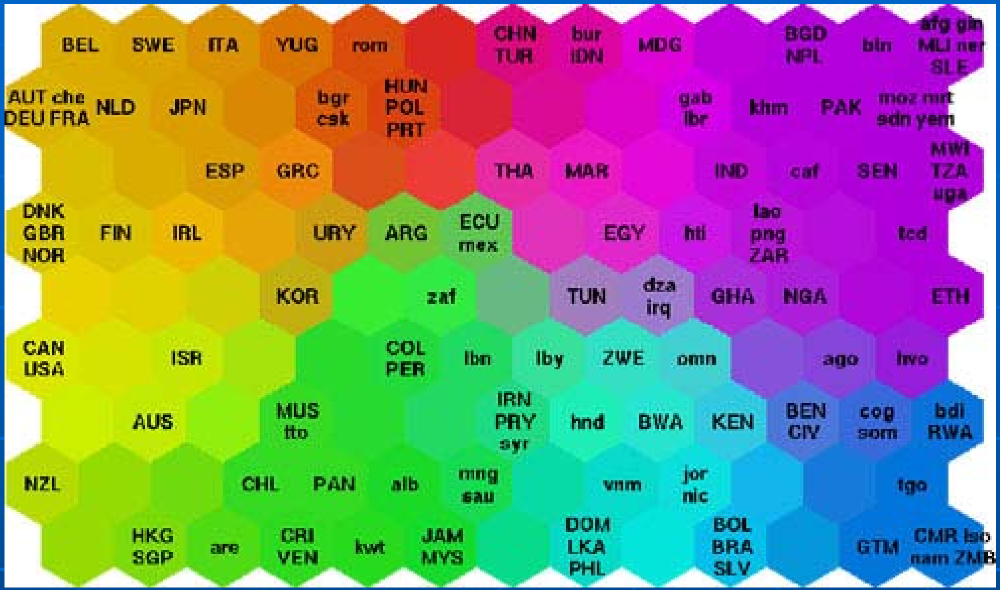
- Tamayo et al (1999) applied SOMs to gene expression data (snapshots of gene activity over time) of 828 genes in yeast cell cycle data

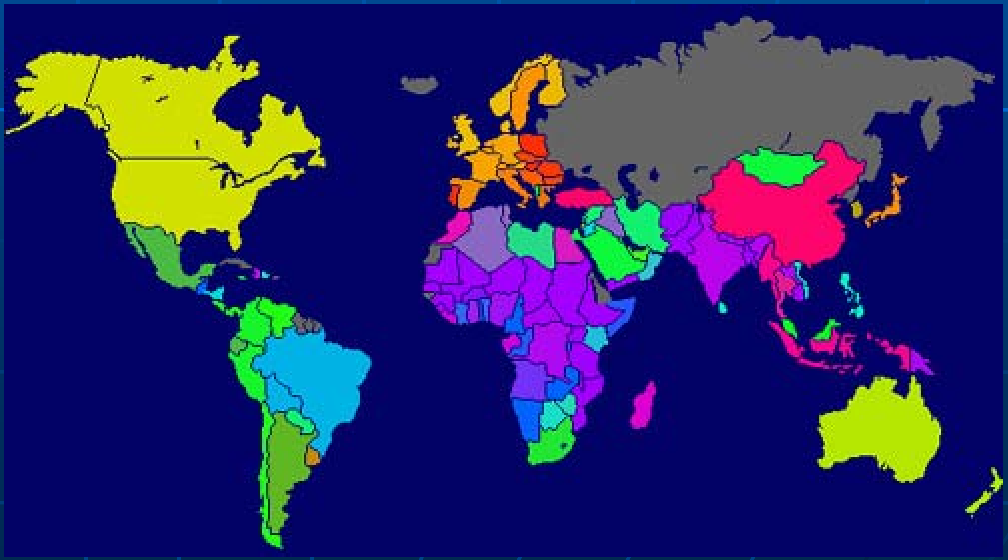
- The SOM found 30 clusters groups of which mapped to the phases of the cell G1, S, G2 and M



# Other Applications

- Classifying webpages by their textual content
- Grouping countries according to their ‘poverty’ based on 39 indicators about quality of life



http://www.cis.hut.fi/research/som-research/worldmap.html

# Summary

- Many variants of neural networks exist but these are some of the most popular
- The main distinction is between supervised and unsupervised learning
- Unsupervised techniques tend to be used for clustering and classification problems
- Must be able to label the clusters afterwards