Business Intelligence Course
Submitted by: Zehavit Harush Ganon

SOM : Self Organizing Maps

Teuvo Kohonen

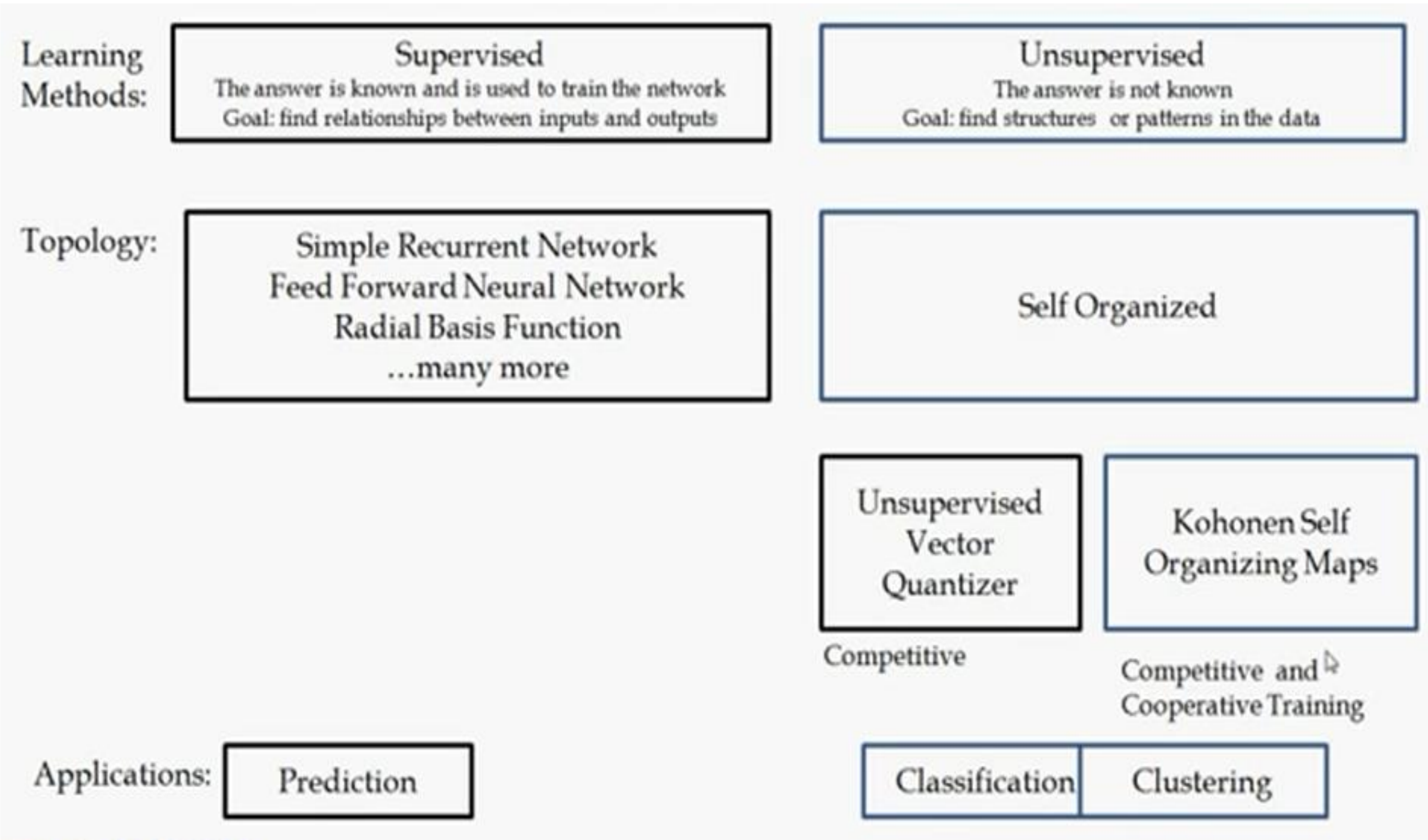
Roadmap

- Introduction
- Artificial Neural Networks
- Competitive learning
- Self-Organizing Map (SOM)
- Examples
- Survey of Practical Application of The Map
- Uses & Application
- Conclusion & Summery

Introduction

- In the last lecture , we talked about supervised learning.
- This is not biologically plausible: In a biological system, there is no external “teacher” who manipulates the network’s weights from outside the network.
- Self organization is a basic property of the brain’s computational structure.
- Biologically more adequate: unsupervised learning.
- We will speak about Self-Organizing Maps (SOMs) - unsupervised learning (Kohonen, 1980).

Artificial Neural Networks



Competitive learning

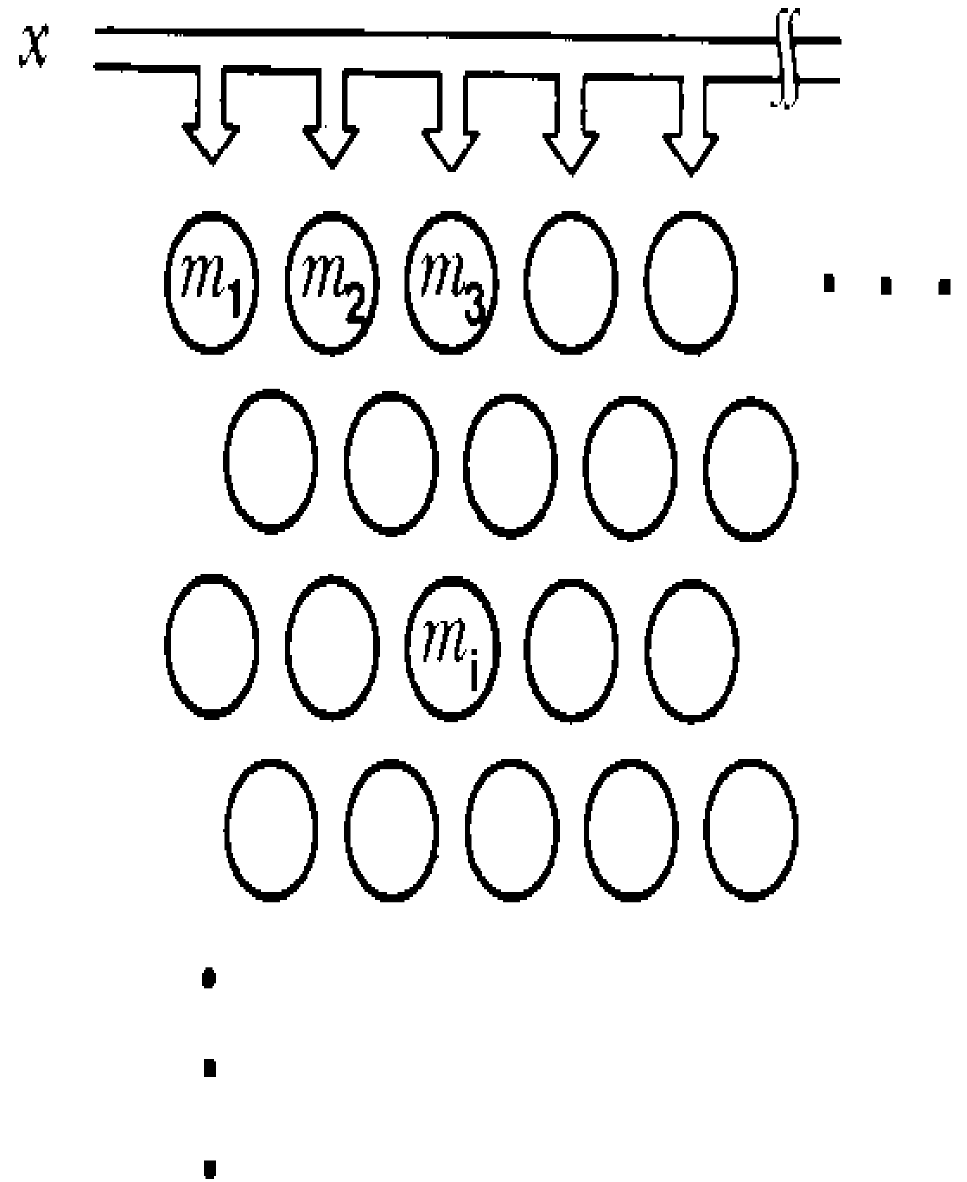
- Assume :
  - a sequence of statistical samples of vectorial observable $x = x(t) \in \mathbb{R}^n$ t-time coordinate (t>0).
  - Set of variables references vectors $\{m_i(t) : m_i \in \mathbb{R}^n, i = 1, 2, \cdots, k\}$
  - Distance measure $d(x, m_i)$
  - i=c index of the best matching reference vector
  - The smallest distance $d(x, m_c)$
  - P(x)=the probability density function of the sample x
- Function of the input vector x:

$$\|x - m_c\| = \min_i \{\|x - m_i\|\}$$

- The expected rth power reconstruction error

$$E = \int \|x - m_c\|^r p(x)\, dx$$

- Cell arrangement for the map and definition of variables.



Competitive learning

- $\alpha(t)$ monotonically decreasing sequence of scalar valued gain coefficients, $0<\alpha(t)<1$
- Simplest analytical description of competitive learning:

$$\boldsymbol{m}_c(t+1) = \boldsymbol{m}_c(t) + \alpha(t)\,[\boldsymbol{x}(t) - \boldsymbol{m}_c(t)],$$

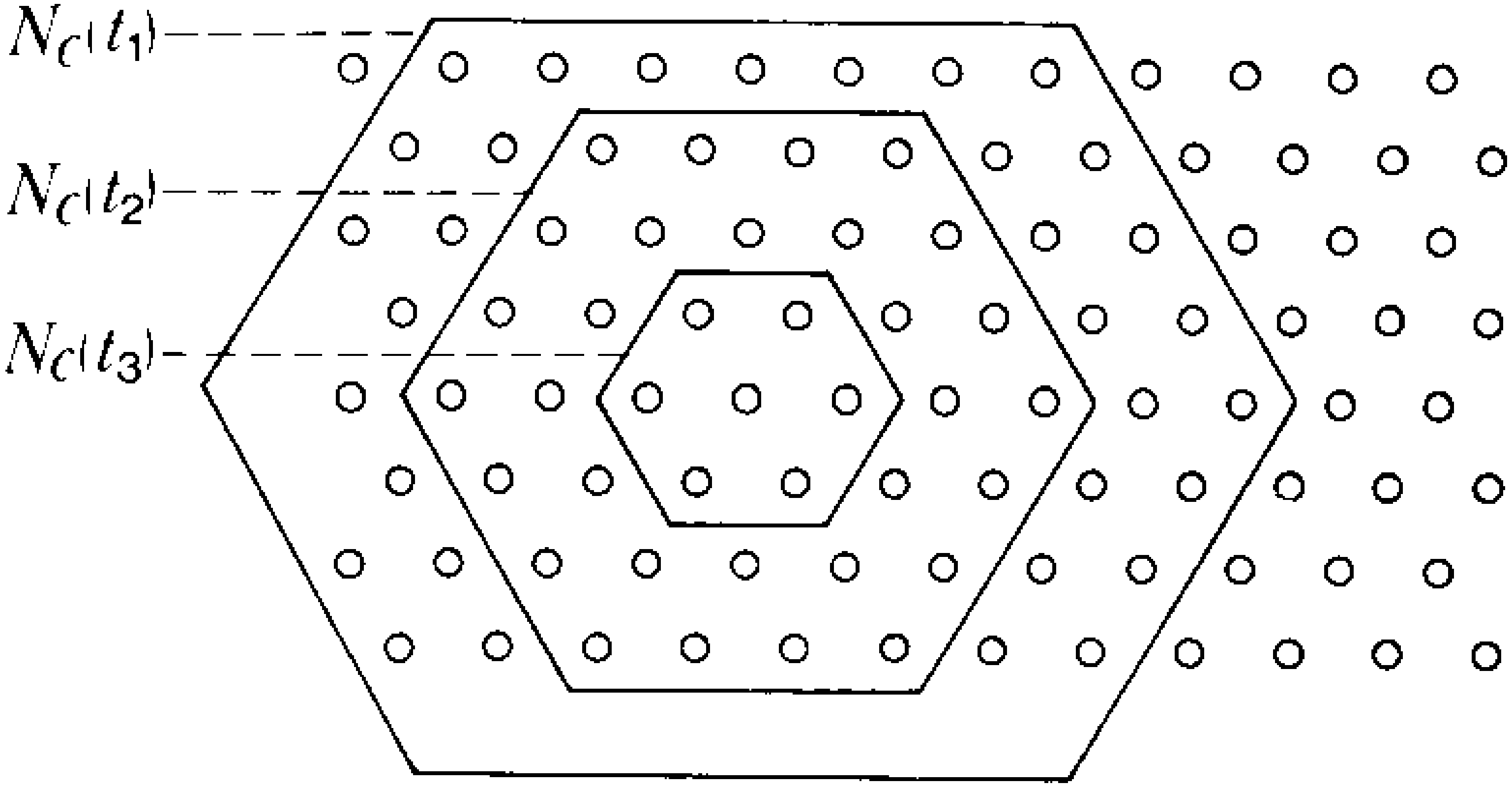
$$\boldsymbol{m}_i(t+1) = \boldsymbol{m}_i(t) \qquad \text{for } i \neq c$$

- The "winner" m_c: $d(\boldsymbol{x}, \boldsymbol{m}_c) = \min_i \{d(\boldsymbol{x}, \boldsymbol{m}_i)\}$

- Updating rule: the correction $\delta \boldsymbol{m}_i$ of m_i :

$$[\mathrm{grad}_{\boldsymbol{m}_i} d(\boldsymbol{x}, \boldsymbol{m}_i)]^T \cdot \delta \boldsymbol{m}_i < 0$$

Neighborhood $N_c(t)$



- Examples of topological neighborhood $N_c(t)$, where $t_1<t_2<t_3$.

Self-Organizing Maps (Kohonen Maps)

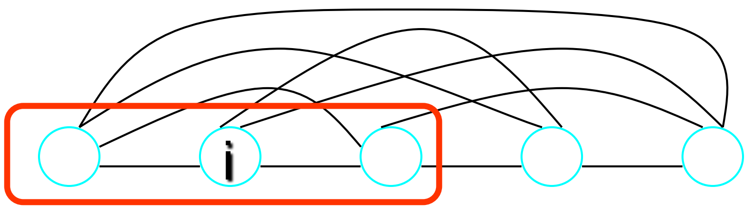
- In the human cortex, multi-dimensional sensory input spaces (e.g., visual input, tactile input) are represented by two-dimensional maps.
- The projection from sensory inputs onto such maps is topology conserving.
- This means that neighboring areas in these maps represent neighboring areas in the sensory input space.
- For example, neighboring areas in the sensory cortex are responsible for the arm and hand regions.

Self-Organizing Maps (Kohonen Maps)

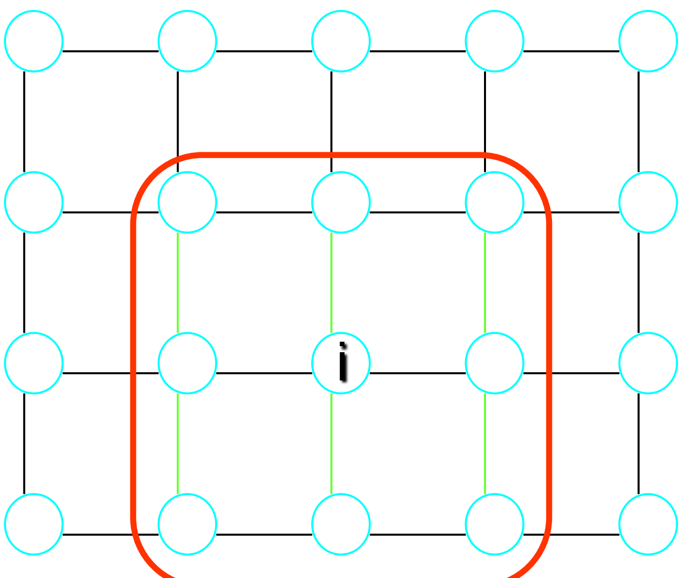
- Such topology-conserving mapping can be achieved by SOMs:
- Two layers: input layer and output (map) layer
- Input and output layers are completely connected.
- Output neurons are interconnected within a defined neighborhood.
- A topology (neighborhood relation) is defined on the output layer.

Self-Organizing Maps (Kohonen Maps)

Common output-layer structures:



One-dimensional
(completely interconnected)

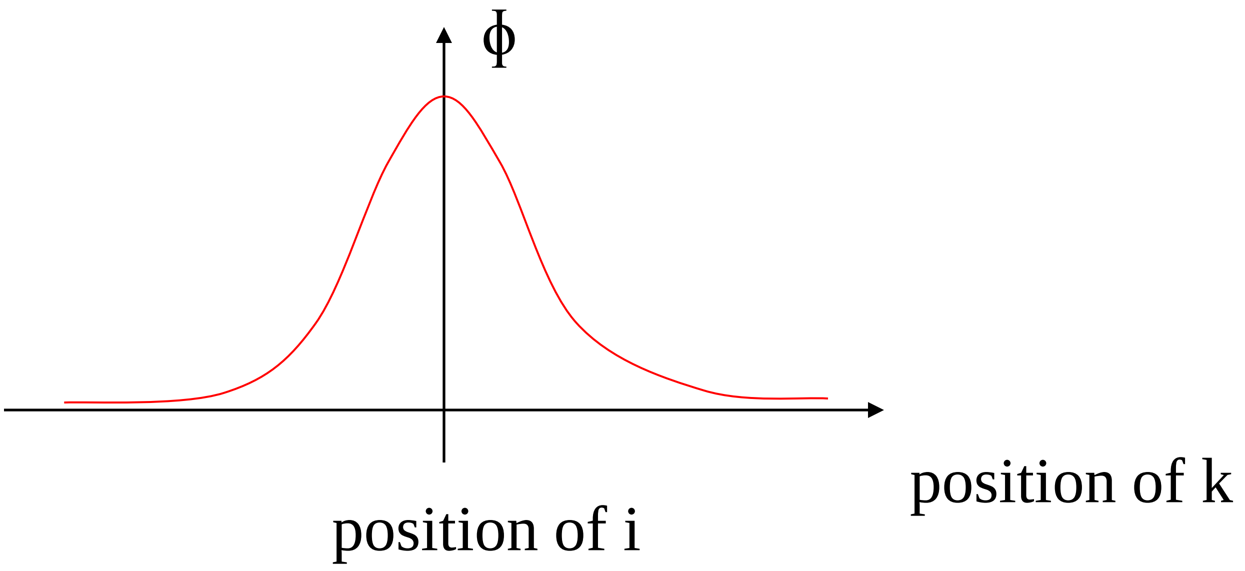


Two-dimensional
(connections omitted, only neighborhood relations shown [green])

Neighborhood of neuron i

Self-Organizing Maps (Kohonen Maps)

- A neighborhood function $\phi(i, k)$ indicates how closely neurons i and k in the output layer are connected to each other.
- Usually, a Gaussian function on the distance between the two neurons in the layer is used:



Unsupervised Learning in SOMs

For n-dimensional input space and m output neurons:

(1) Choose random weight vector w_i for neuron i, i = 1, ..., m

(2) Choose random input x

(3) Determine winner neuron k:

$$||w_k - x|| = \min_i ||w_i - x|| \quad \text{(Euclidean distance)}$$

(4) Update all weight vectors of all neurons i in the neighborhood of neuron k: $w_i := w_i + \alpha \cdot \phi(i, k) \cdot (x - w_i)$
(w_i is shifted towards x)

(5) If convergence criterion met, STOP.
Otherwise, narrow neighborhood function ϕ and learning parameter α and go to (2).

Update rule

- The following update rule is used for each neuron i in the the neighborhood of winner neuron b

$$\boxed{m_i(t+1) = m_i(t) + \alpha(t)\, h_{bi}(t)\left[x - m_i(t)\right]}$$

t *is the discrete time coordinate*

$m_i(t+1)$ *is a prototype vector at* $t+1$

$$h_{bi}(t) = \exp\left[-\frac{\left\|r_b - r_i\right\|^2}{2\sigma^2(t)}\right] \textit{ is a neighbourhood kernel}$$

r_b, r_i *radius vectors of b,i neurons*

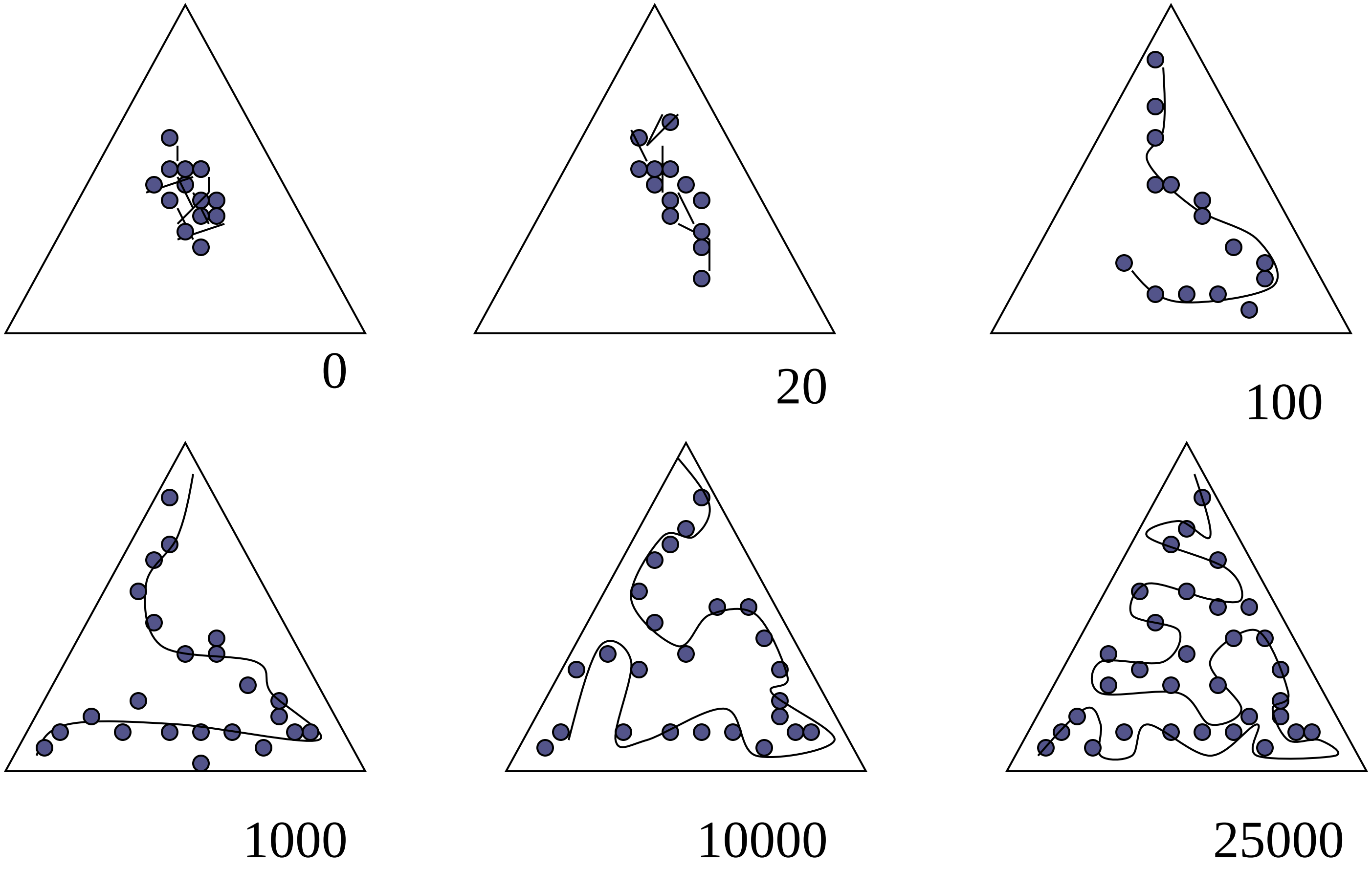
$\sigma(t)$ *is the width of the* kernel

$\alpha(t)$ *is a scalar valued learning rate of the map*

$\sigma(t), \alpha(t)$ *are monotonically decreasing with time*

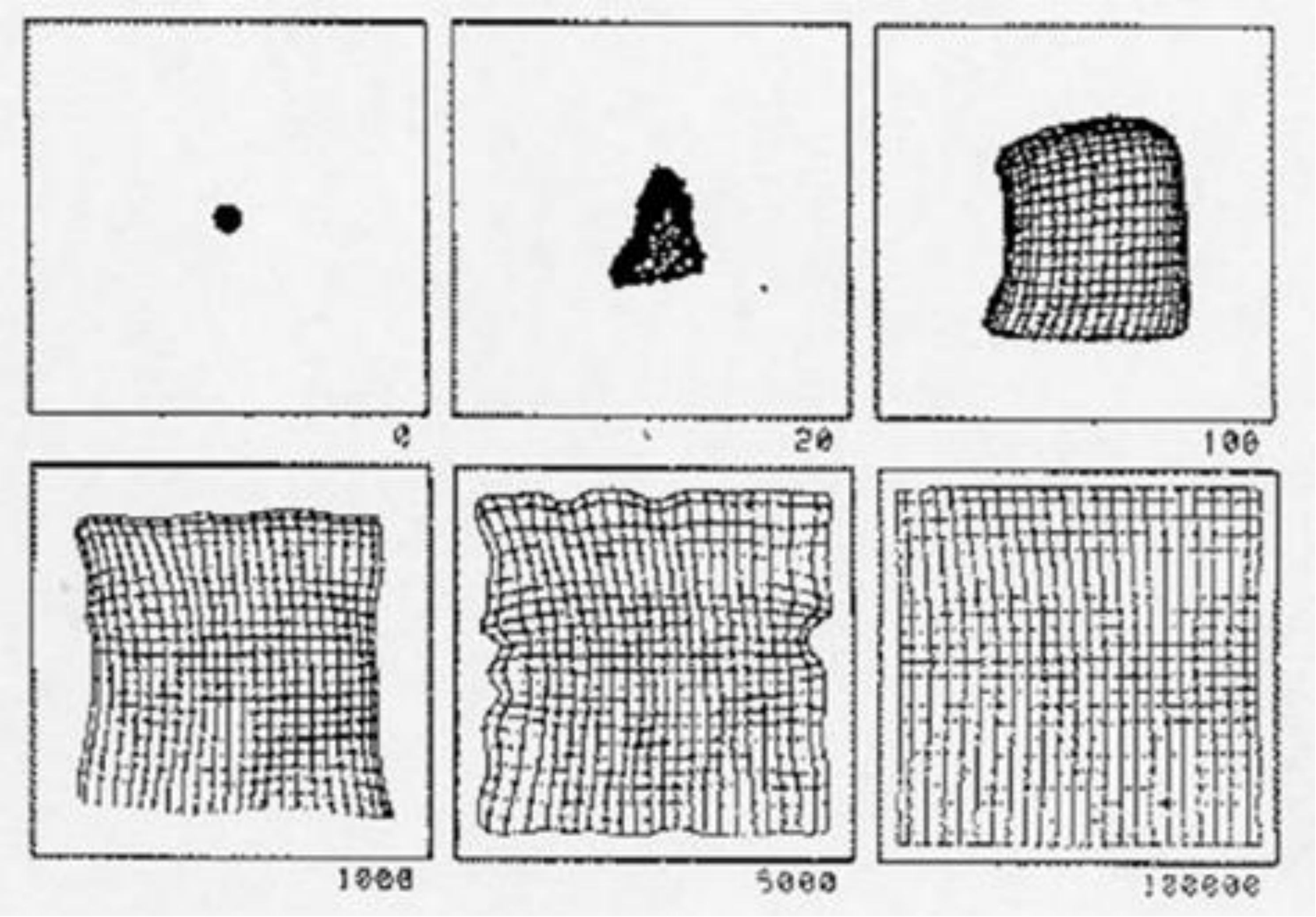
One dimension

Example : Learning a one-dimensional representation of a two-dimensional (triangular) input space: weight vector during the ordering process, one dimensional array.

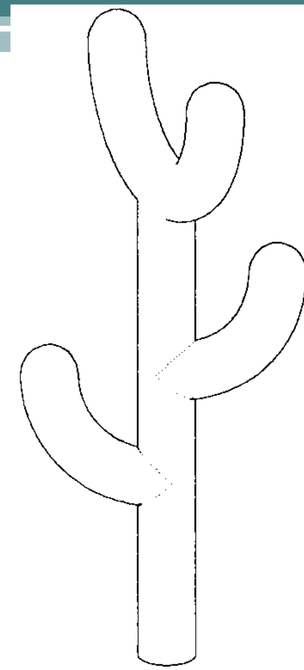


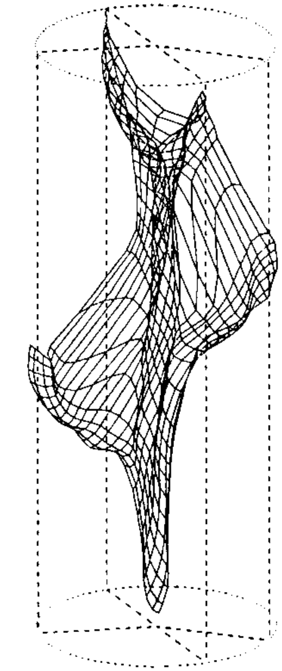
Two-dimensions

Example : Learning a two-dimensional representation of a two-dimensional (square) input space: weight vectors during the ordering process, two dimensional array.

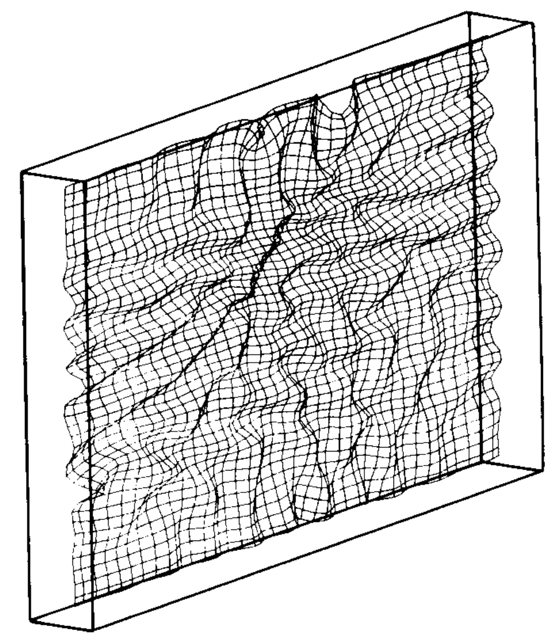


Three-dimensions





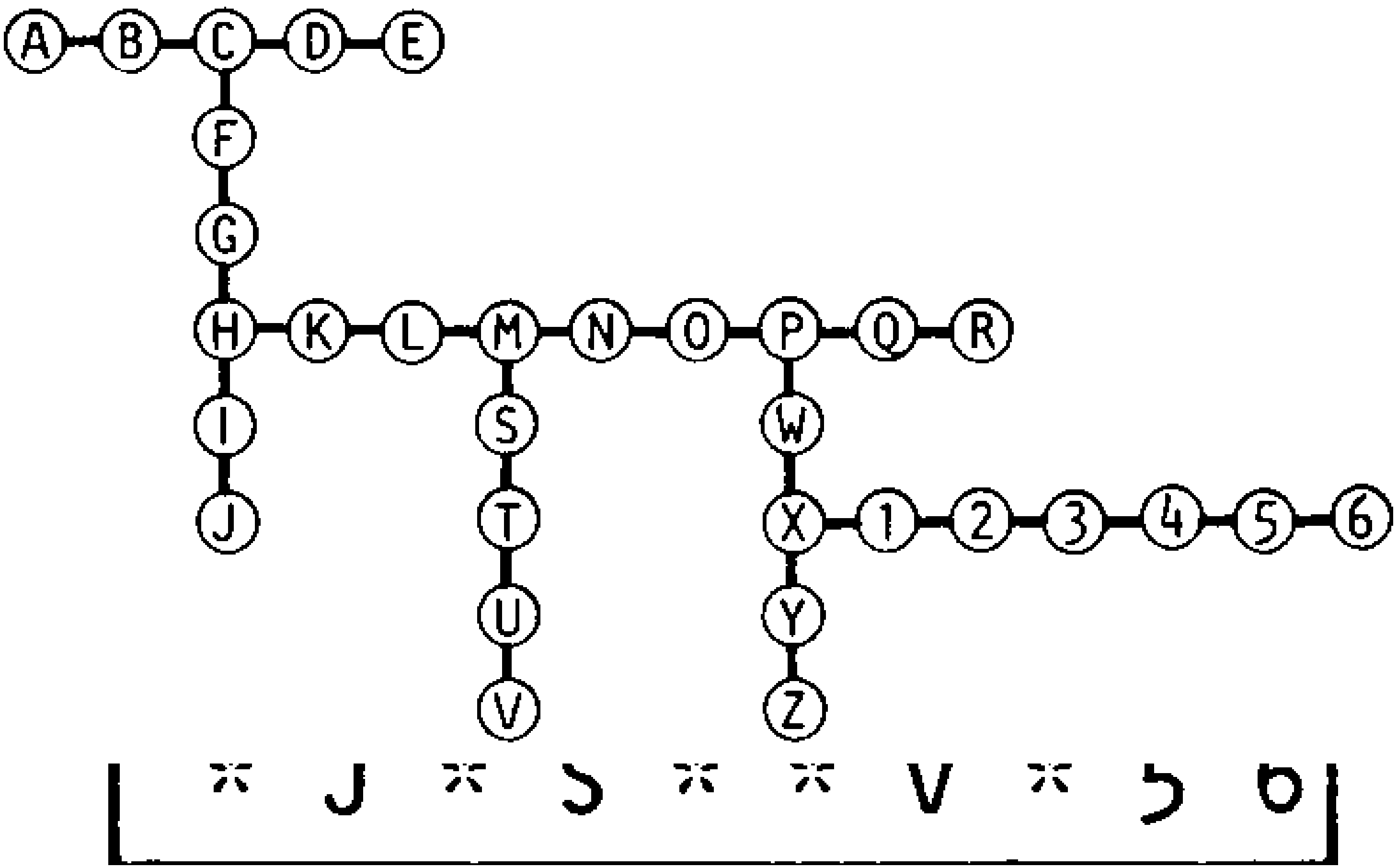
- Representation of three-dimensional (uniform) density functions by two-dimensional maps.



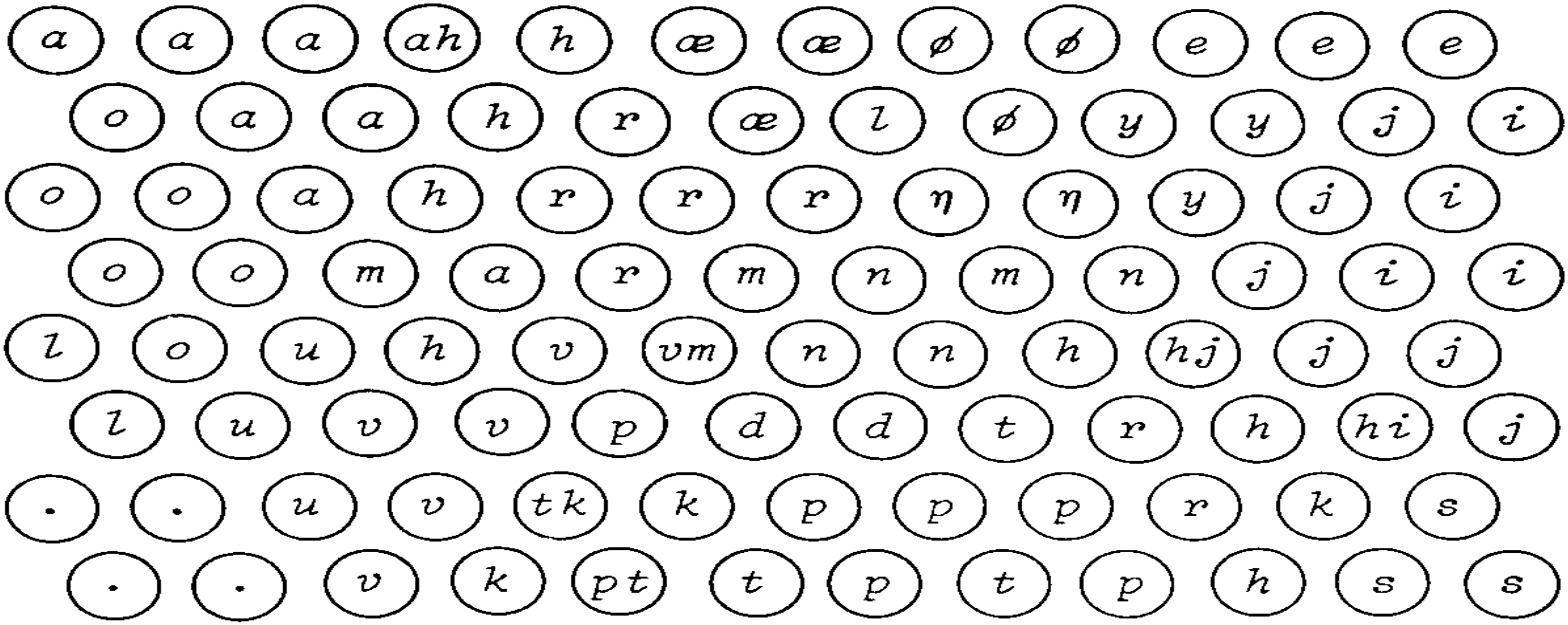
Taxonomy (Hierarchical clustering) of abstract data

Table 1 Input Data Matrix

Attribute	Item A	B	4	5	6
a_1	1	2	3	3	3
a_2	0	0	3	3	3
a_3	0	0	6	6	6
a_4	0	0	2	2	2
a_5	0	0	4	5	6



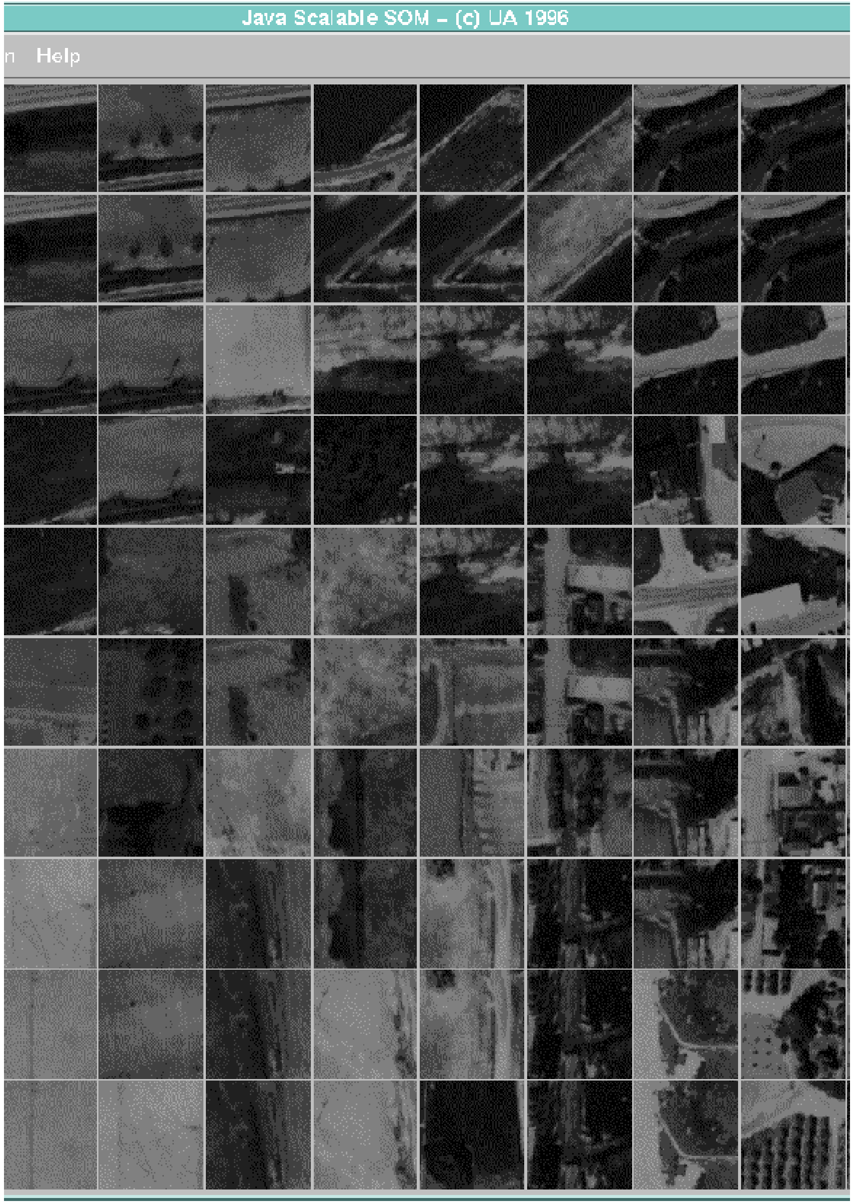
Phoneme Map



An example of phoneme map. Natural Finnish speech was processed by a model of the inner ear which performs its frequency analysis. The resulting signals were then connected to an artificial network, the cells which are shown in this picture as circles. The cells were tuned automatically, without any supervision or extra information given, to the acoustic units of speech known as phonons. The cells are labeled by the symbols of those phonemes to which the “learned” to give responses; most cells give a unique answer, whereas the double labels show which cells respond two phonemes.

Unsupervised Learning in SOMs

Example :
Learning a two-dimensional mapping of texture images



Survey of Practical Application of The Map

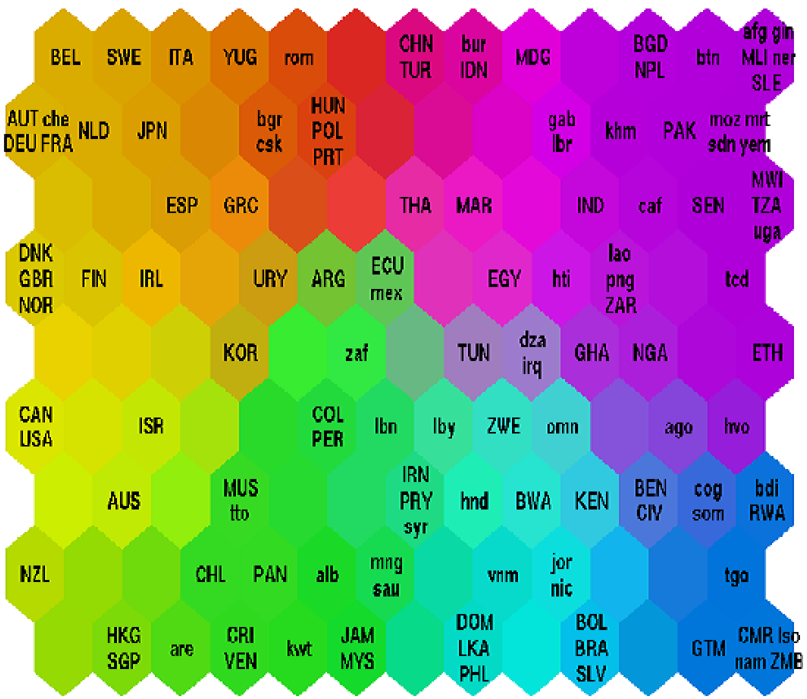
- Statistical pattern recognition, especially recognition of speech.
- Control of robot arms, and other problems in robotics.
- Control of industrial processes, especially diffusion processes in the production of semiconductor substrates.
- Automatic synthesis of digital systems.
- Adaptive devices for various telecommunications tasks.
- Image compression.
- Radar classification of sea-ice.
- Optimization problems.
- Sentence understanding.
- Application of expertise in conceptual domains.
- Classification of insert courtship songs.

Uses & Applications

- CSSCP –Classification System for Serial Criminal Patterns – Chicago police.

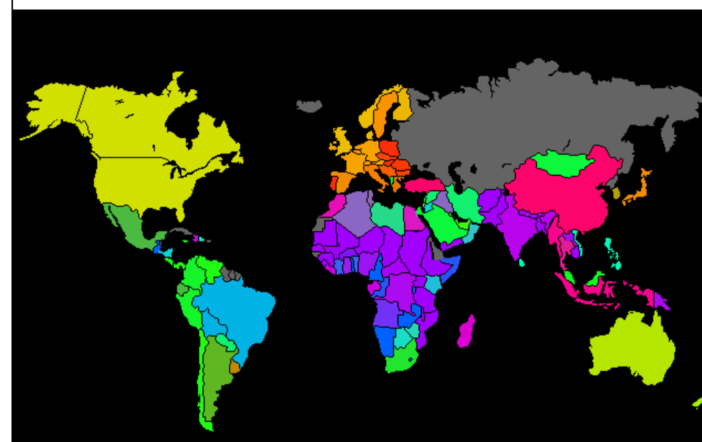
Uses & Applications

- Classifying World Poverty



The Country Names

AFG	Afghanistan	GTM	Guatemala	NZL	New Zealand
AGO	Angola	HKG	Hong Kong	OAN	Taiwan, China
ALB	Albania	HND	Honduras	OMN	Oman
ARE	United Arab Emirates	HTI	Haiti	PAK	Pakistan
ARG	Argentina	HUN	Hungary	PAN	Panama
AUS	Australia	HVO	Burkina Faso	PER	Peru
AUT	Austria	IDN	Indonesia	PHL	Philippines
BDI	Burundi	IND	India	PNG	Papua New Guinea
BEL	Belgium	IRL	Ireland	POL	Poland
BEN	Benin	IRN	Iran, Islamic Rep.	PRT	Portugal
BGD	Bangladesh	IRQ	Iraq	PRY	Paraguay
BGR	Bulgaria	ISR	Israel	ROM	Romania
BOL	Bolivia	ITA	Italy	RWA	Rwanda
BRA	Brazil	JAM	Jamaica	SAU	Saudi Arabia
BTN	Bhutan	JOR	Jordan	SDN	Sudan
BUR	Myanmar	JPN	Japan	SEN	Senegal
BWA	Botswana	KEN	Kenya	SGP	Singapore
CAF	Central African Rep.	KHM	Cambodia	SLE	Sierra Leone
CAN	Canada	KOR	Korea, Rep.	SLV	El Salvador
CHE	Switzerland	KWT	Kuwait	SOM	Somalia
CHL	Chile	LAO	Lao PDR	SWE	Sweden
CHN	China	LBN	Lebanon	SYR	Syrian Arab Rep.
CIV	Cote d'Ivoire	LBR	Liberia	TCD	Chad
CMR	Cameroon	LBY	Libya	TGO	Togo
COG	Congo	LKA	Sri Lanka	THA	Thailand
COL	Colombia	LSO	Lesotho	TTO	Trinidad and Tobago
CRI	Costa Rica	MAR	Morocco	TUN	Tunisia
CSK	Czechoslovakia	MDG	Madagascar	TUR	Turkey
DEU	Germany	MEX	Mexico	TZA	Tanzania
DNK	Denmark	MLI	Mali	UGA	Uganda
DOM	Dominican Rep.	MNG	Mongolia	URY	Uruguay
DZA	Algeria	MOZ	Mozambique	USA	United States
ECU	Ecuador	MRT	Mauritania	VEN	Venezuela
EGY	Egypt, Arab Rep.	MUS	Mauritius	VNM	Viet Nam
ESP	Spain	MWI	Malawi	YEM	Yemen, Rep.
ETH	Ethiopia	MYS	Malaysia	YUG	Yugoslavia
FIN	Finland	NAM	Namibia	ZAF	South Africa
FRA	France	NER	Niger	ZAR	Zaire
GAB	Gabon	NGA	Nigeria	ZMB	Zambia
GBR	United Kingdom	NIC	Nicaragua	ZWE	Zimbabwe
GHA	Ghana	NLD	Netherlands		
GIN	Guinea	NOR	Norway		
GRC	Greece	NPL	Nepal		



Roadmap

- ✓ Introduction
- ✓ Artificial Neural Networks
- ✓ Competitive learning
- ✓ Self-Organizing Map (SOM)
- ✓ Examples
- ✓ Survey of Practical Application of The Map
- ✓ Uses & Application
- Conclusion & Summery

Summary

- The SOM help classify data
- The SOM help identify clusters and outlying data points
- The exact meaning of the clusters and outliers are left for interpretation
- This is a tool that can be added to the more conventional seismic interpretation process
- Reduces human bias in the analysis

Conclusion & Summery

- SOM is the most popular artificial neural network algorithm in the unsupervised learning category.
- About 4000 research articles on it have appeared in the open literature, and many industrial projects use the SOM as a tool for solving hard real-world problems.
- Many fields of science have adopted the SOM as a standard analytical tool: statistics, signal processing, control theory, financial analysis, experimental physics, chemistry and medicine.
- The SOM solves difficult high-dimensional and nonlinear problems such as feature extraction and classification of images and acoustic patterns, adaptive control of robots, and equalization, demodulation, and error-tolerant transmission of signals in telecommunications.
- A new area is the organization of very large document collections.
- The SOM is one of the most realistic models of the biological brain function.

Questions

Thank you for your attention...

References

- The Self-Organizing Map ,*T Kohonen - Proceedings of the IEEE, 1990 - ieeexplore.ieee.org*