Perceptron Networks and Applications

M. Ali Akcayol
Gazi University
Department of Computer Engineering

Content

- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Self organizing maps

- Perceptron, Adaline, backpropagation networks, radial basis function networks are based on supervised learning.
- The self organizing map (SOM) network is the most prominent type of artificial neural network model which uses unsupervised learning.
- SOM is trained using an unsupervised learning algorithm.
- SOMs (also called as Kohonen SOM) are the category of competitive learning networks.
- No human intervention is required during the training phase.
- A very little information about the characteristics of the input data is needed.

Self organizing maps

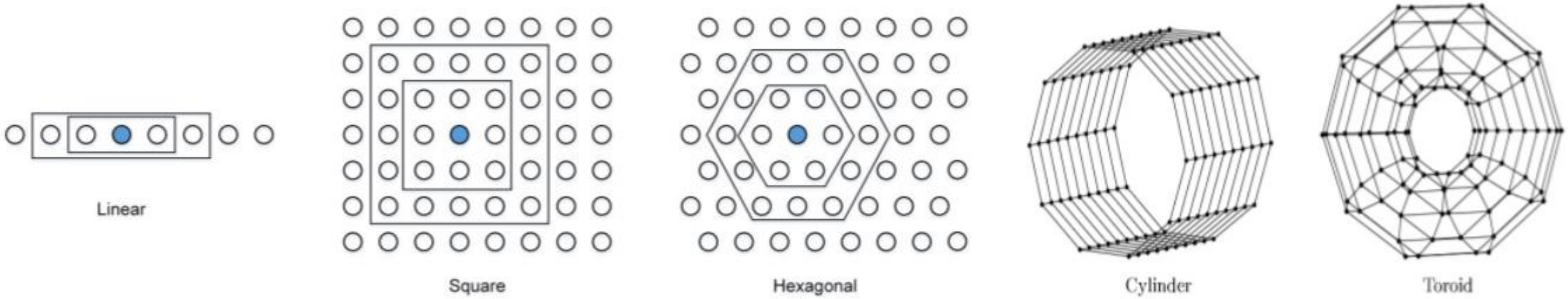
- Unlike other ANNs, SOMs use a neighborhood function to conserve the topological properties of the input space.
- In other words, it provides a topological preserving mapping from the high dimensional space to lower dimensional space.
- The relative distance between the points are preserved during the mapping.
- Points that are closest to each other in the input space are mapped to adjacent map portions in the SOM.
- The SOM can thereby serve as a cluster analyzing technique for high dimensional data.
- Also, the SOM is capable of generalizing.

Self organizing maps

- SOMs are also include two operating modes: training and mapping.
- Training refers to the building of map using input examples (vector quantization).
- Mapping refers to the classification of a new input vector.
- A SOM consists of elements called nodes or neurons.
- Each node is assigned with a weight vector and a map space is associated along with it.
- The nodes are usually arranged in a two-dimensional regular spacing, such as hexagonal or rectangular grid pattern.

Self organizing maps

- In simple words, the SOMs can be considered as an arrangement of two dimensional assembly of neurons.
- SOM topologies can be in one, two (most common) or even three dimensions.
- Two most used two dimensional grids in SOMs are rectangular and hexagonal grid.
- Three dimensional topologies can be in form of a cylinder or toroid shapes.

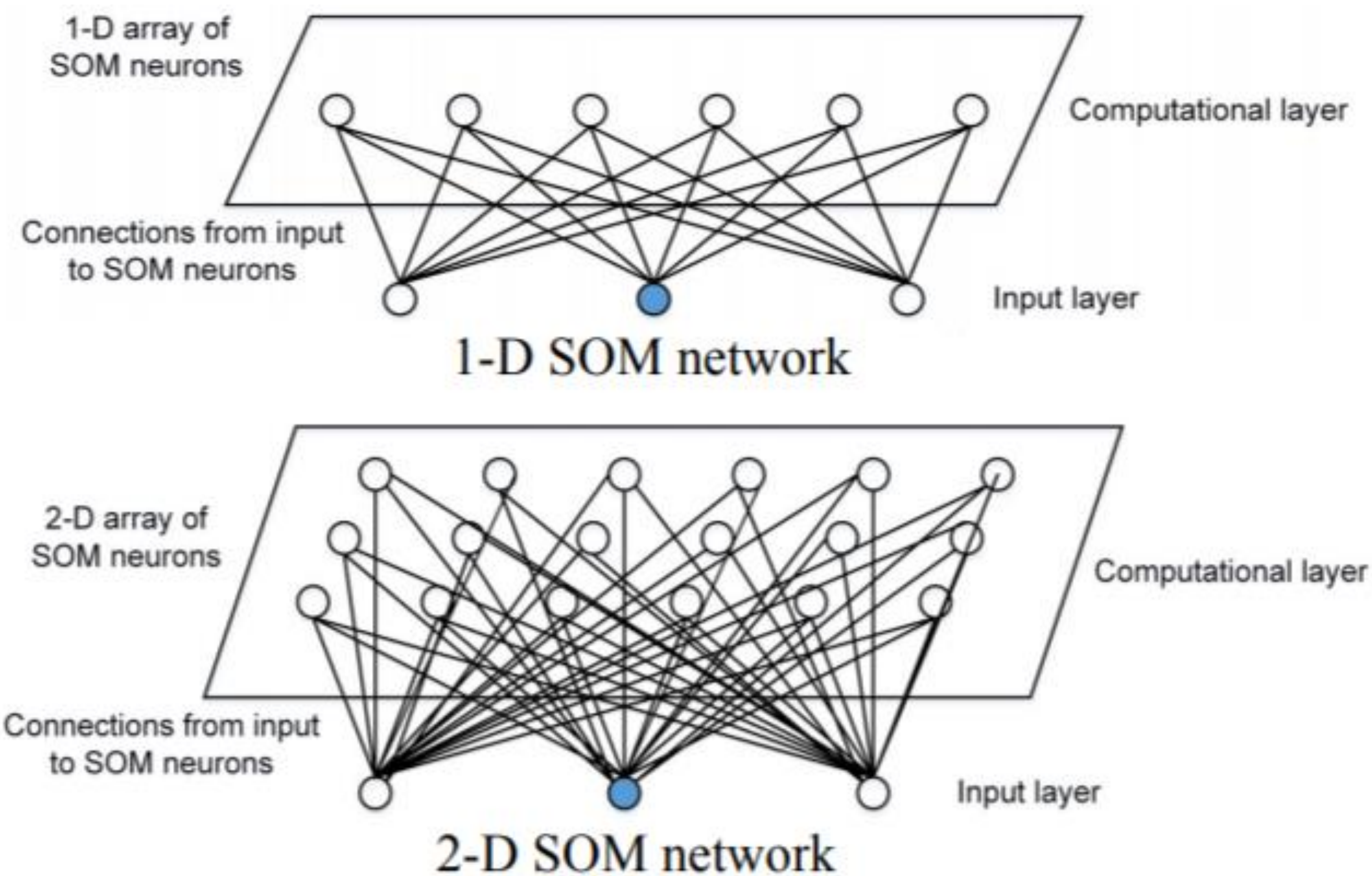


Content

- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Architecture

- SOM networks are mostly designed as 1-D and 2-D.



Content

- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Training

- A SOM does not require a target output value unlike many other types of networks.
- Instead, the area is selectively optimized to more closely resemble the data for the class. The member of the class is the input vector.
- From the initial classification of random weights, and after going through several iterations, the SOM finally settles into a map of stable sectors.
- Each sector is a feature classifier.
- The graphical output can be concluded as a type of feature map of the input space.

Training

- The algorithm utilized for the self organization of the network is based on three fundamental processes:
  - Competitive process
  - Cooperation process
  - Synaptic adaptation

Content

- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Competitive process

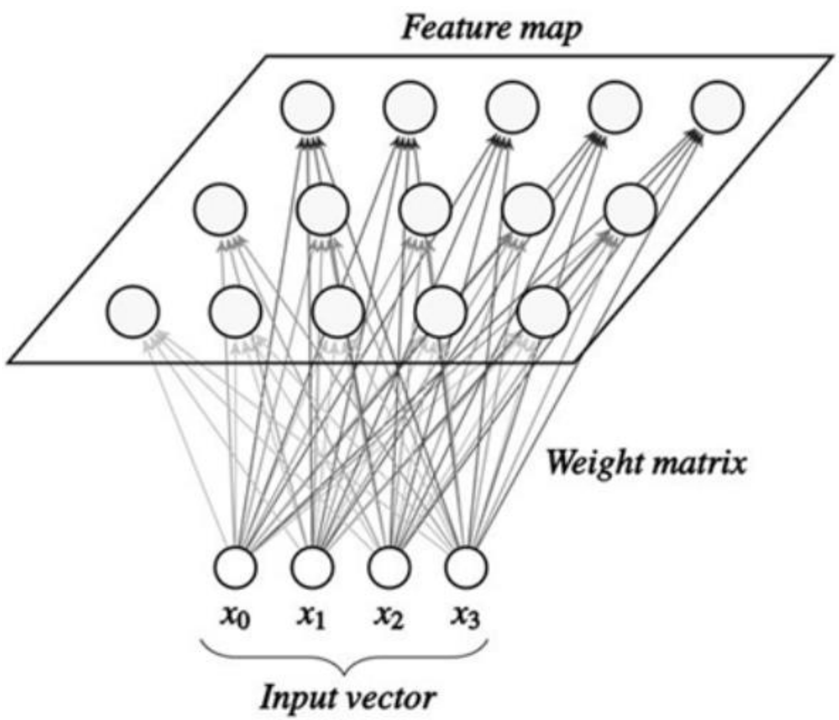
- First of all, each node’s weights are initialized.
- A vector is selected randomly from the set of training data and introduced over the lattice.
- Each node’s weight is investigated to determine whose weights are most alike the input vector.
- The winning node is referred to as the Best Matching Unit (BMU).
- To determine the BMU, the Euclidean distance for all the nodes is calculated.
- The node having the weight vector nearest to the input vector is categorized as the winning node or the BMU.

Competitive process

- The steps for the competition are:
  - Let the dimension of the input space is denoted by m.
  - A pattern chosen randomly from input space.

    $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$
  - Each node's synaptic weight in the output layer has the same dimension as the input space. The weight of neuron *j* is denoted as:

    $\mathbf{w}_j = [w_{j1}, w_{j2}, \dots, w_{jm}]^T, \quad j = 1, 2, \dots, n$ n = number of neurons at output layer



Competitive process

- The steps for the competition are:
  - In order to determine the best match of the input vector *x* with the synaptic weights w_j the Euclidean distance is used.
  - The Euclidean distance between a pair of vectors ***X*** and $\boldsymbol{W}_j$ is represented by,

$$d = \| x - w_{ij} \| = \left[\sum_{j=1}^{n} (x_i - w_{ij})^2 \right]^{1/2}$$

where, *x* is the input vector (current) and *w* is the node's weight vector.

Competitive process

- The steps for the competition are:
  - For example, to calculate the distance between the input vector, the color red (1, 0, 0) with an arbitrary weight vector (0.1, 0.4, 0.5),

$$\text{distance, d} = [\,(1 - 0.1)^2 + (0 - 0.4)^2 + (0 - 0.5)^2\,]^{1/2}$$
$$= [(0.9)^2 + (-0.4)^2 + (-0.5)^2]^{1/2}$$
$$= [0.81 + 0.16 + 0.25]^{1/2}$$
$$= (1.22)^{1/2} = 1.1061.106$$

  - The neuron with the infinitesimal distance is called i(**x**):

$$i(\mathbf{x}) = \arg\min_j \|\mathbf{x} - \mathbf{w}_j\|,\ j = 1, 2, \ldots, l$$

  - The neuron (i) that fulfills the above condition is called **best-matching** or **winning neuron** for the input vector x.

Content

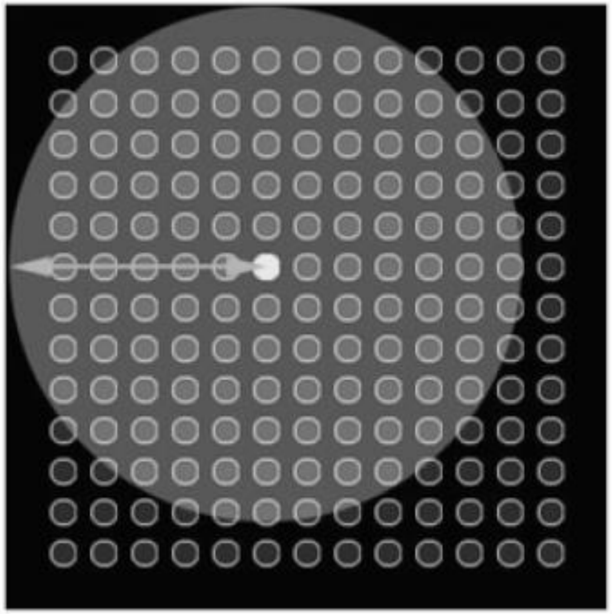
- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Cooperative process

- In this step, the radius of the neighborhood of the best matched unit is calculated.
- The value initially starts at large and typically set to the ‘radius’ of the lattice, but diminishes with each iteration.
- All the nodes that are present within this radius are considered to be inside the BMU’s neighborhood.
- After determining the BMU, next we have to observe the other nodes within the BMU’s neighborhood.
- The weight vectors for all nodes will be adjusted in the next step.
- This can be done by simply calculating the radius of the neighborhood.

Cooperative process

- Almost all nodes are taken within radius from the start of training.



- The winning neuron effectively locates the center of a topological neighborhood.
- A winning neuron excites other neurons.
- The neurons that exist in its immediate neighborhood are more excited.
- Inhibits more the neurons that they are in longer distances.

Cooperative process

Cooperation - continue

- In the neighborhood are included only excited neurons, while inhibited neurons exist outside of the neighborhood.
- Let d_{ij} is the lateral distance between neurons i and j (assuming that i is the winner and it is located in the center of the neighborhood).
- h_{ji} denotes the topological neighborhood around neuron i, then h_{ji} is a unimodal function of distance.

$$h_{ji(x)} = \exp(-d_{ij}^2 / 2\sigma^2)$$

where, σ is the effective width of the neighborhood.

- It measures the degree up to which the excited neurons in the area of the winning neuron participate in the learning process.

Cooperative process

Cooperation - continue

- The distance among neurons is defined as the Euclidean metric.

$$d_{ij}^2 = \| r_j - r_i \|^2$$

  where the discrete vector r_j describes the position of excited neuron j and r_i defines the position of the winning neuron in the lattice.

- Another characteristic feature of the SOM algorithm is that the size of the neighborhood shrinks with time.
- This requirement is satisfied by making the width of the Gaussian function decreasing with time.

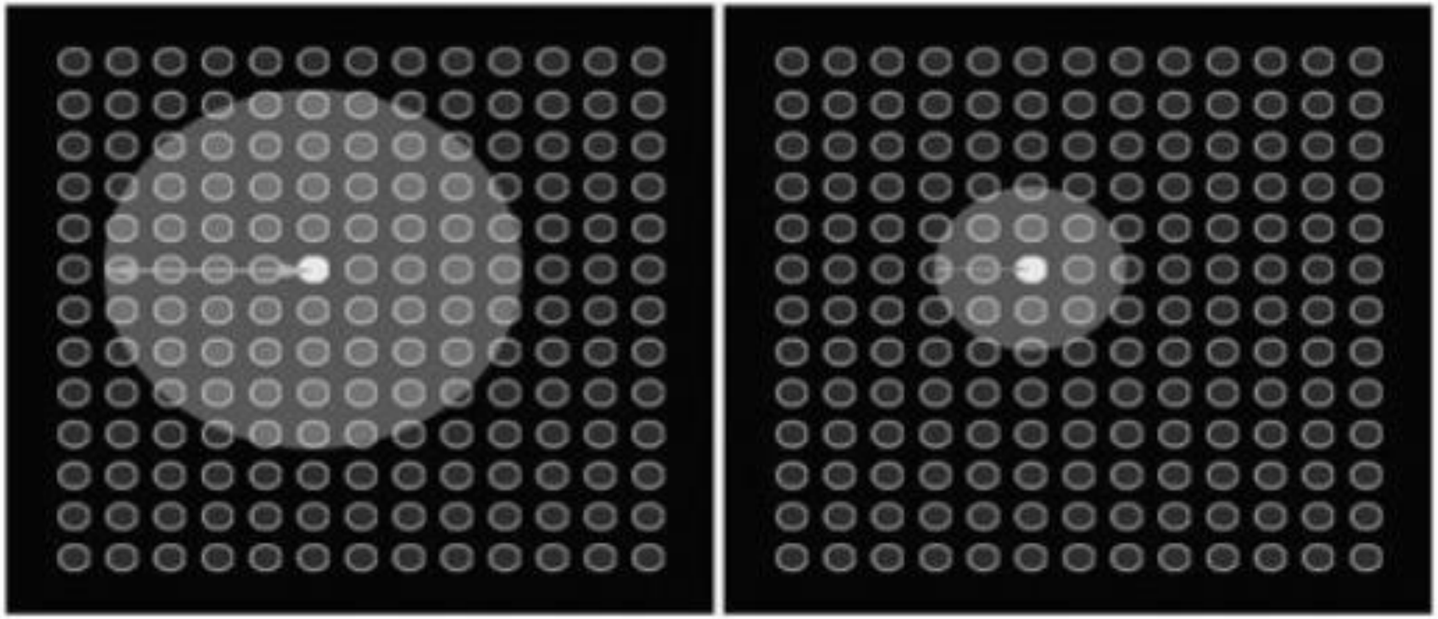
Cooperative process

- A popular choice is the exponential decay function.

  $\sigma(n) = \sigma_0 \exp(-n / \tau_1);\ n = 0, 1, 2, \ldots$

  where σ_0 is the value of σ at the initialization of the SOM algorithm and τ_1 is a time constant.
- Correspondingly the neighborhood function assumes a time dependent form:

$$h_{ji(x)} = \exp(-d_{ij}^2 / 2\sigma(n)^2)$$



Content

- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Synaptic adaptation

- All the nodes in the lattice to be determined whether they lay within the radius or not.
- If a node is present in the neighborhood then its weight vector is adjusted.
- The rate of alteration in the weights of the node, depends on its distance from the BMU.
- The adaptive process modifies the weights of the network so as to achieve the self-organization of the network.
- Just the weight of the winning neuron and that of neurons inside its neighborhood are adapted.
- All the other neurons have no change in their weights.

Synaptic adaptation

- The weight of the winning neuron and neurons inside its neighborhood are changed.

$$\Delta w_j = \alpha\, h_{ji(x)}\, (x - w_j)$$

$$w_j(n+1) = w_j(n) + \eta(n)\, h_{ji(x)}(n)\, (x - w_j(n))$$

- The learning rate is also required to be time varying.

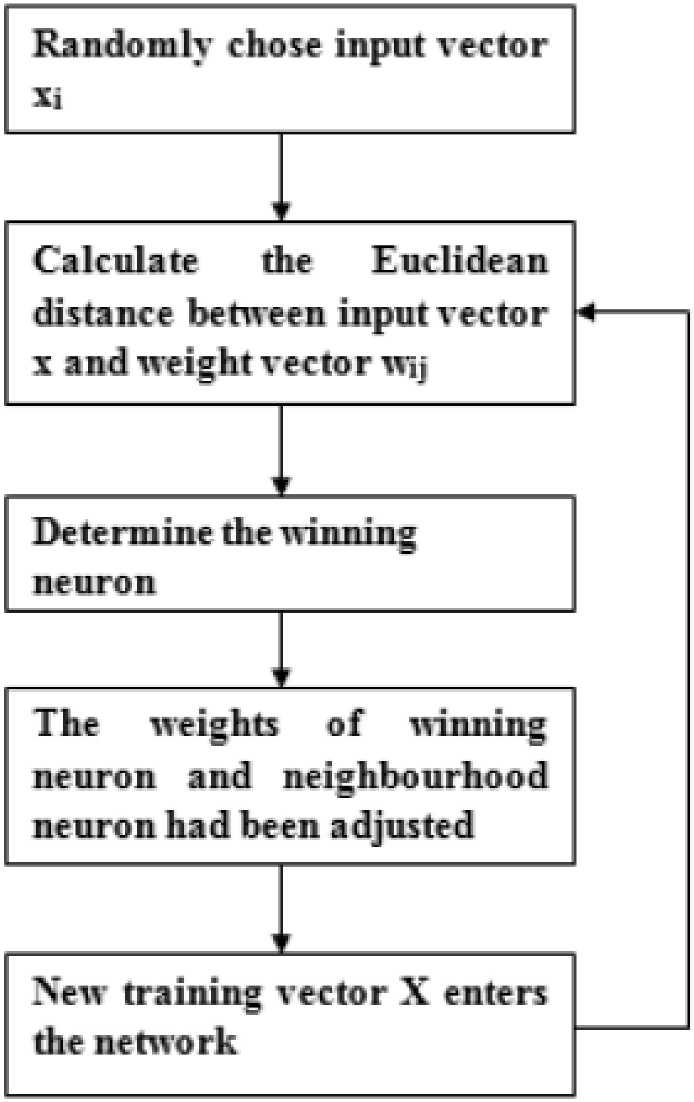
$$\eta(n) = \eta_0 \exp(-n/\tau_2);\; n = 0,1,2.....$$

$$\eta_0 = 0.1,\; \tau_2 = 1000$$

- Convergence phase is needed to fine tune the feature map.
- In general the number of iterations needed for this phase is 500 times the total number of neurons.
- During this phase, η must be a small value, (e.g. 0.01).

Training

- Flow diagram of Kohonen SOM algorithm.



Content

- Self organizing maps
- Architecture
- Training
  - Competitive process
  - Cooperative process
  - Synaptic adaptation
- Applications

Applications

World Poverty Map

- The Self-Organizing Map (SOM) can be used to portray complex correlations in statistical data.
- In this example, the data consisted of World Bank statistics of countries in 1992.
- 39 indicators describing various quality-of-life factors, such as state of health, nutrition, educational services, etc, were used.
- The complex joint effect of these factors can be visualized by organizing the countries using the SOM.

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		

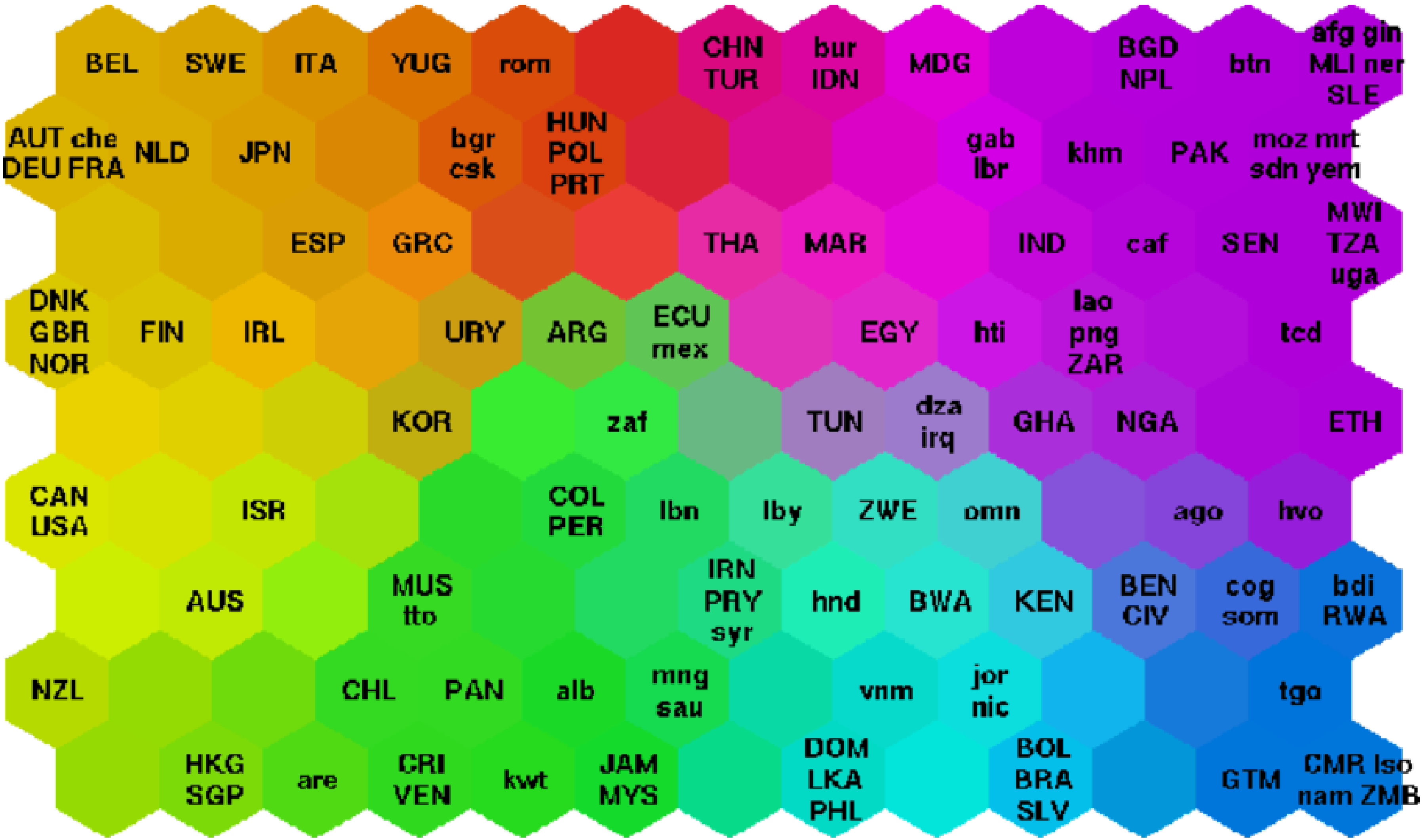
Applications

World Poverty Map – cont.

- Countries that had similar values of the indicators found a place near each other on the map.
- The different clusters on the map were automatically encoded with different bright colors.
- Each country was in fact automatically assigned a color describing its poverty type in relation to other countries.
- The poverty structures of the world were visualized so that each country on the geographic map has been colored according to its poverty type.

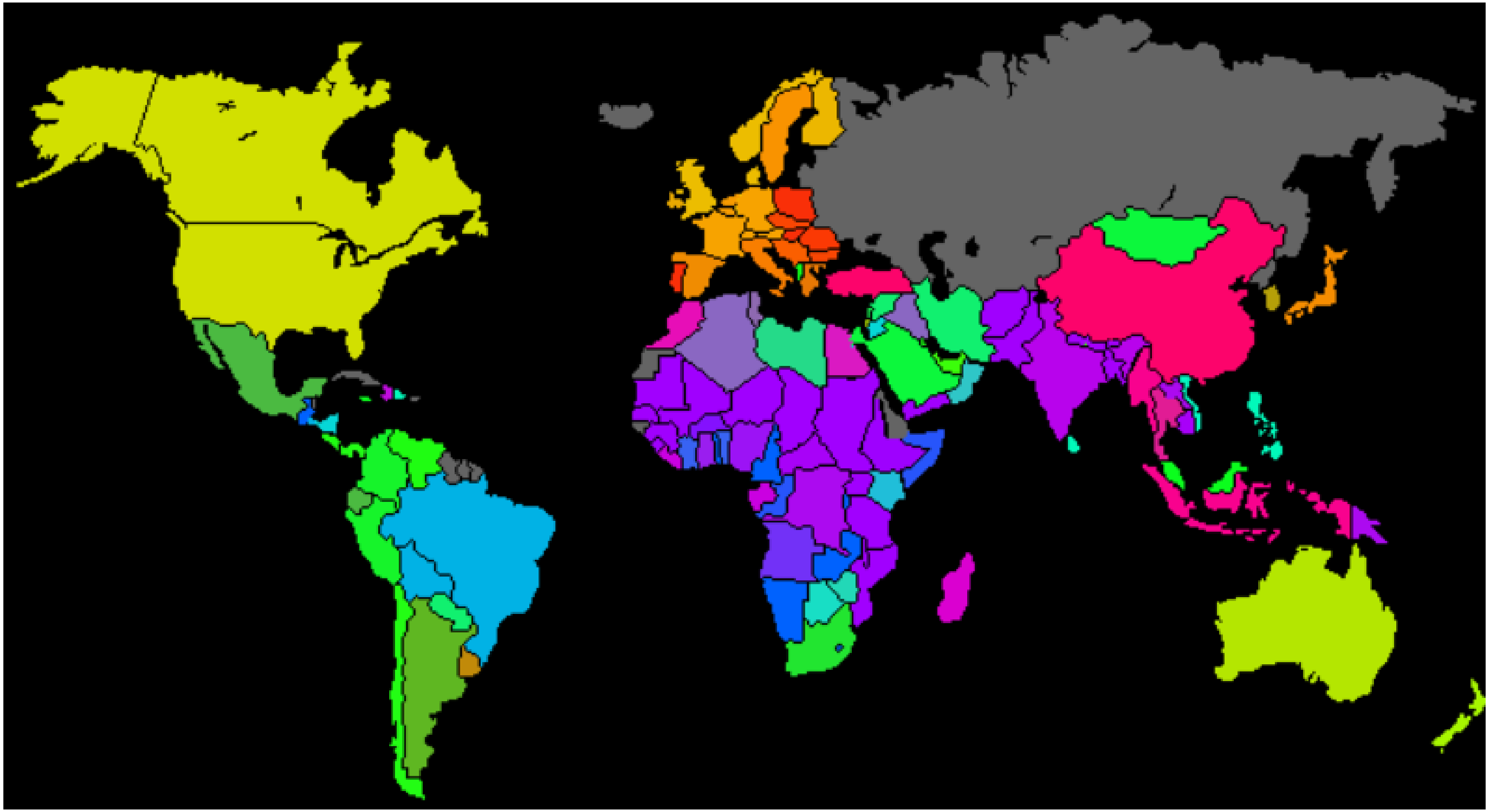
World Poverty Map – cont.

Countries organized on a self-organizing map based on indicators related to poverty:



World Poverty Map – cont.

- A map of the world where countries have been colored with the color describing their poverty type.



Applications

Animal similarity

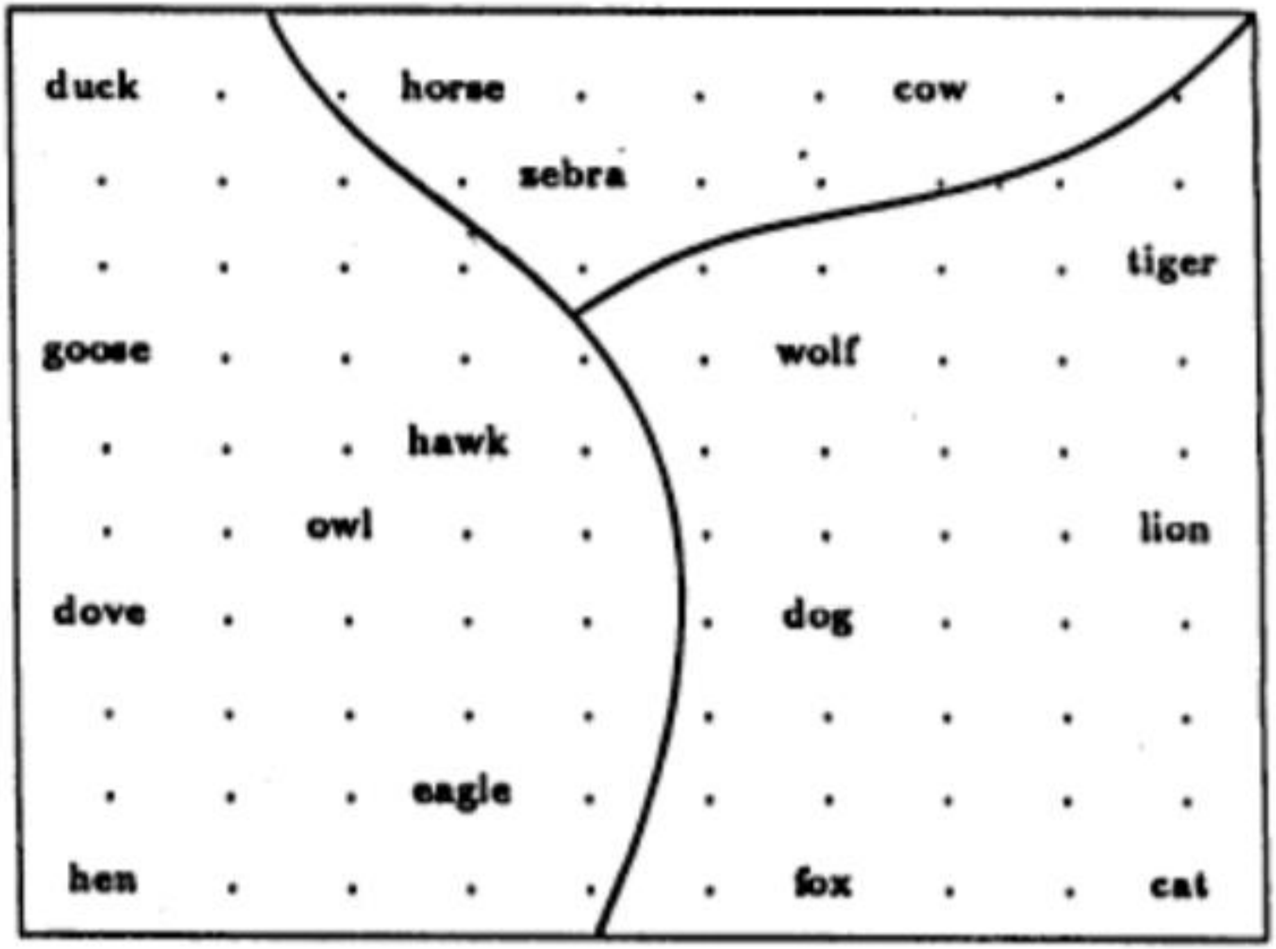
- Animal can be grouped using their attributes.
- In this example, 16 animals with 13 attributes are used.

		dove	hen	duck	goose	owl	hawk	eagle	fox	dog	wolf	cat	tiger	lion	horse	zebra	cow
is	small	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0
	medium	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
	big	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
has	2 legs	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	4 legs	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	hair	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	hooves	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	mane	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0
	feathers	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
likes to	hunt	0	0	0	0	1	1	1	1	0	1	1	1	1	0	0	0
	run	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0
	fly	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
	swim	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0

Applications

Animal similarity – cont.

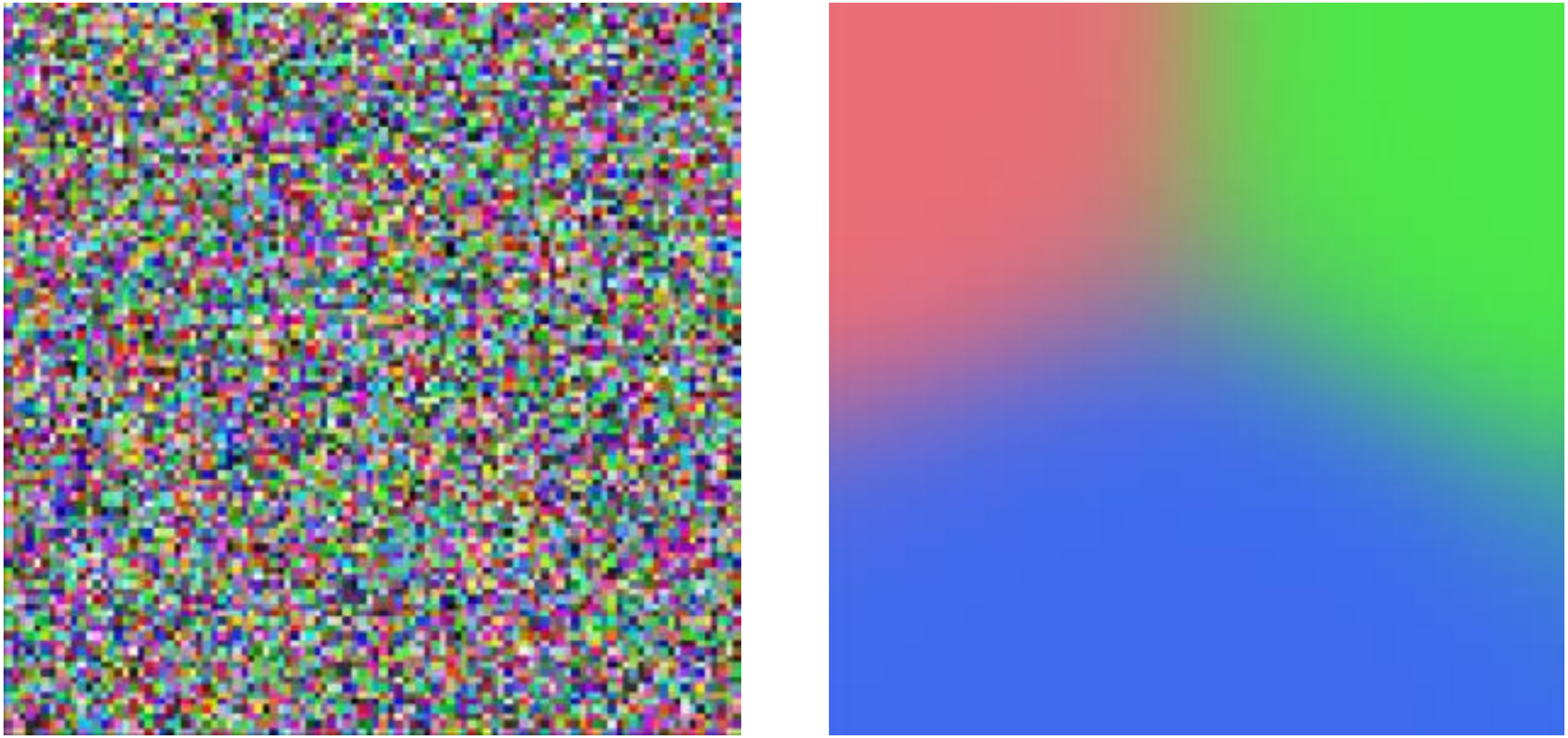
- After the SOM trained, the animals have been grouped according to the similarities.



Applications

Color similarity

- In the figure, the similarity between thousands of randomly distributed points, each with different properties can be found with SOM network.



Homework

- Prepare a report on the use of self organizing maps in medical applications.