RHEI'2004

"SCIENTIFIC ACHIEVEMENTS FOR WELLBEING AND DEVELOPMENT OF SOCIETY".

Pēteris Grabusts

NEURAL NETWORKS METHODS OF KNOWLEDGE EXTRACTION.

NEURAL NETWORKS METHODS OF KNOWLEDGE EXTRACTION

The problem:

How do we turn information into useful knowledge?

Solution:

Data mining & knowledge discovery

Data Mining & Knowledge Discovery

This class provides

- Tools & techniques for producing useful knowledge from information

- Experience in using these tools

Data -- Information -- Knowledge

The set of values:-

12345	1000.00	AS
32320	228.29	KC

has no meaning. It is **data** but it is **NOT information**.

Information: Information is the result of organizing data into meaningful quantities.

The following relational table helps turns the data into information since it associates meaning with the data:

Account Number	Balance	type
12345	1000.00	AS
32320	228.29	KC

What Is Data Mining?

How Does It Differ From Existing Technologies?

Data Sources: **Databases, data warehouses, Internet**

Decision Support Systems

Tools for asking questions & doing analyses when you know what you want to ask and where you are going.

Data Mining

Process of **discovering knowledge** (meaningful new correlations, patterns and trends) in data by sifting through large amounts of data using pattern pattern recognition as well as statistical and mathematical techniques.

Other Names Used in Conjunction with Data Mining

- Knowledge discovery(mining) in databases (KDD)
- Knowledge extraction
- Data/pattern analysis
- Data archeology
- Data dredging
- Information harvesting
- What is not data mining
  - (Deductive) query processing
  - Expert systems or small statistical programs

Data Mining Example

Customer*

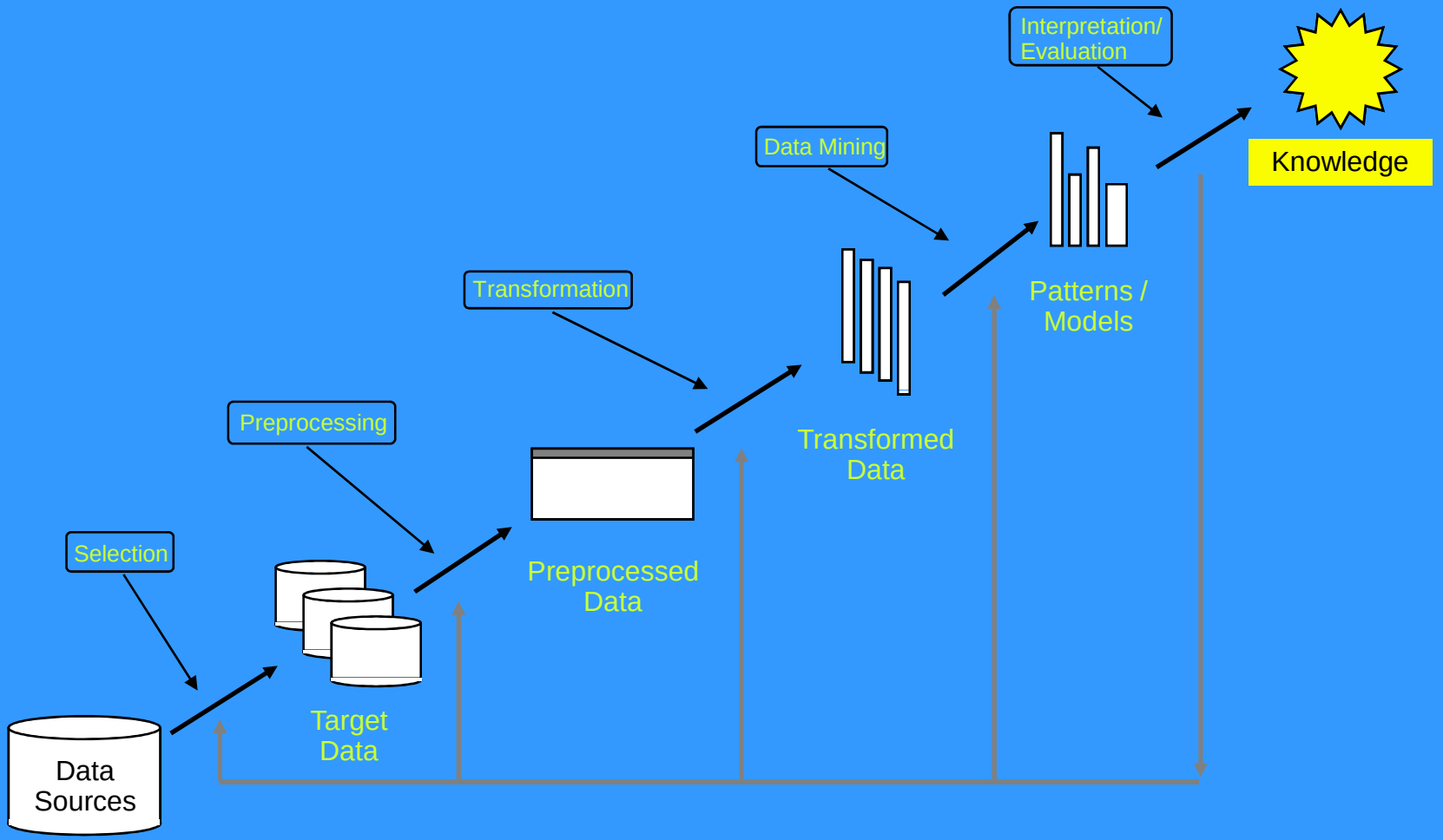
Person	Age	Sex	Income	Customer
Ann Smith	32	F	10,000	yes
Joan Gray	53	F	1,000,000	yes
Mary Blythe	27	F	20,000	no
Jane Brown	55	F	20,000	yes
Bob Smith	30	M	100,000	yes
Jack Brown	50	M	200,000	yes

Knowledge Within A Relation

IF Sex(Person) = F AND age (Person) ≥ 32 THEN Customer(Person)

* Dzeroski, Saso, *Inductive Logic Programming and Knowledge Discovery in Databases*, **Advances in Knowledge Discovery and Data Mining**, Ed. U. Fayyad, G.Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy, AAAI Press, 1996, pp. 117-152.

The Knowledge Discovery Process



Source: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P, *From Data Mining To Knowledge Discovery In Databases*, **AI Magazine**, Fall 1996.

Data Mining Methods

Data Mining study

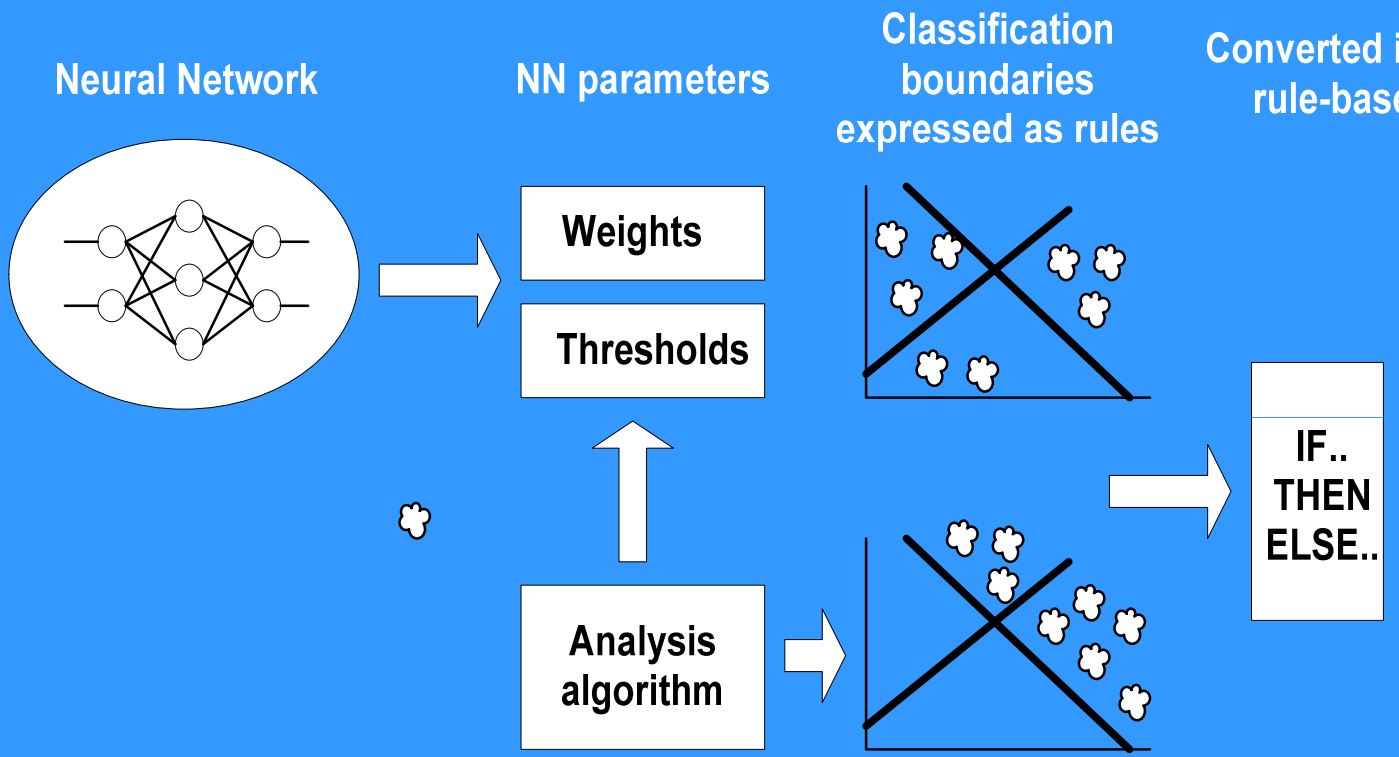
- Data warehouses
- Classification & Association rule miners C4.5
- Neural networks (BP, RBF, SOM)
- Classical tools:

  Correlation

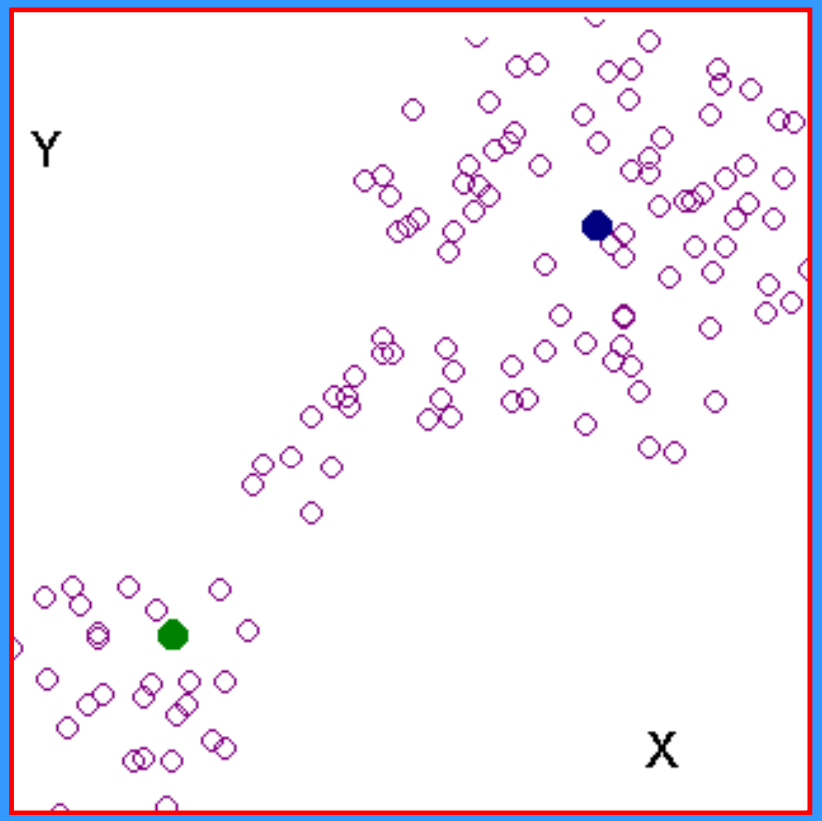
  Regression

  Clustering

Rule extraction process

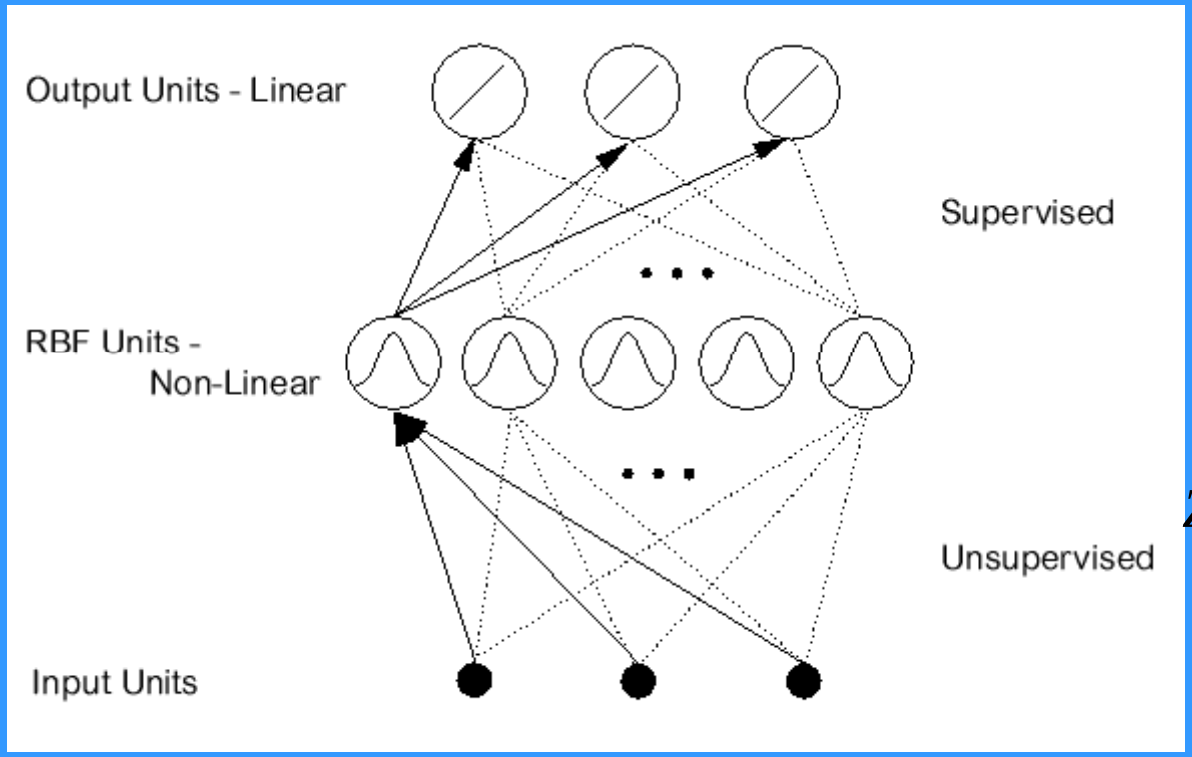


Simple idea



IF x is high THEN y is high
IF x is low THEN y is low

RBF Network Architecture



$$y = \sum_{j=i}^{J} W_{ij} Z_j(x)$$

Gaussian function

$$Z_j(x) = \exp\left(-\frac{\| x - \mu \|^2}{\delta_j^2}\right)$$

The RBF rule extraction algorithm RULEX

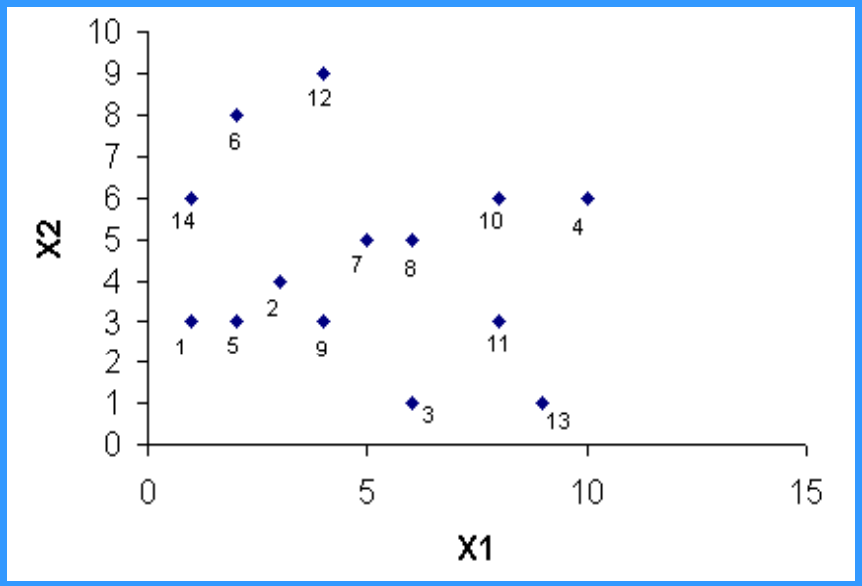
Input:	Hidden weights μ (center positions) Gaussian radius spread σ Steepness S
Output:	One rule per hidden unit
Process:	Train RBF network on data set For each hidden unit: For each μ_i $X_{lower} = \mu_i - \sigma_i + S$ $X_{upper} = \mu_i + \sigma_i - S$ Build rule by: antecedent=[X_{lower}, X_{upper}] Join antecedents with AND Add class label Write rule

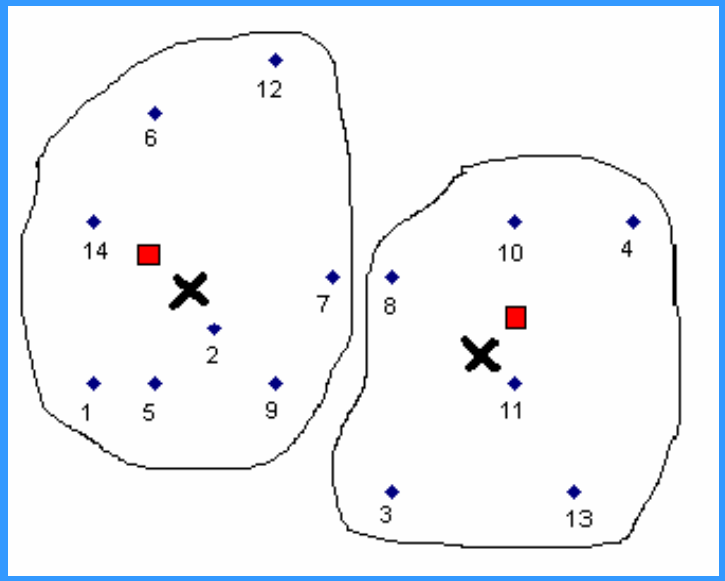
Experiments

(a)RBF neural network learning

(b) rule extraction from trained RBF network.

Data





Two clusters with centers at points (-0.73; 0.26) and (0.97; -0.35)

radius values $\sigma_1^2 = 1.07$ and $\sigma_2^2 = 1.04$

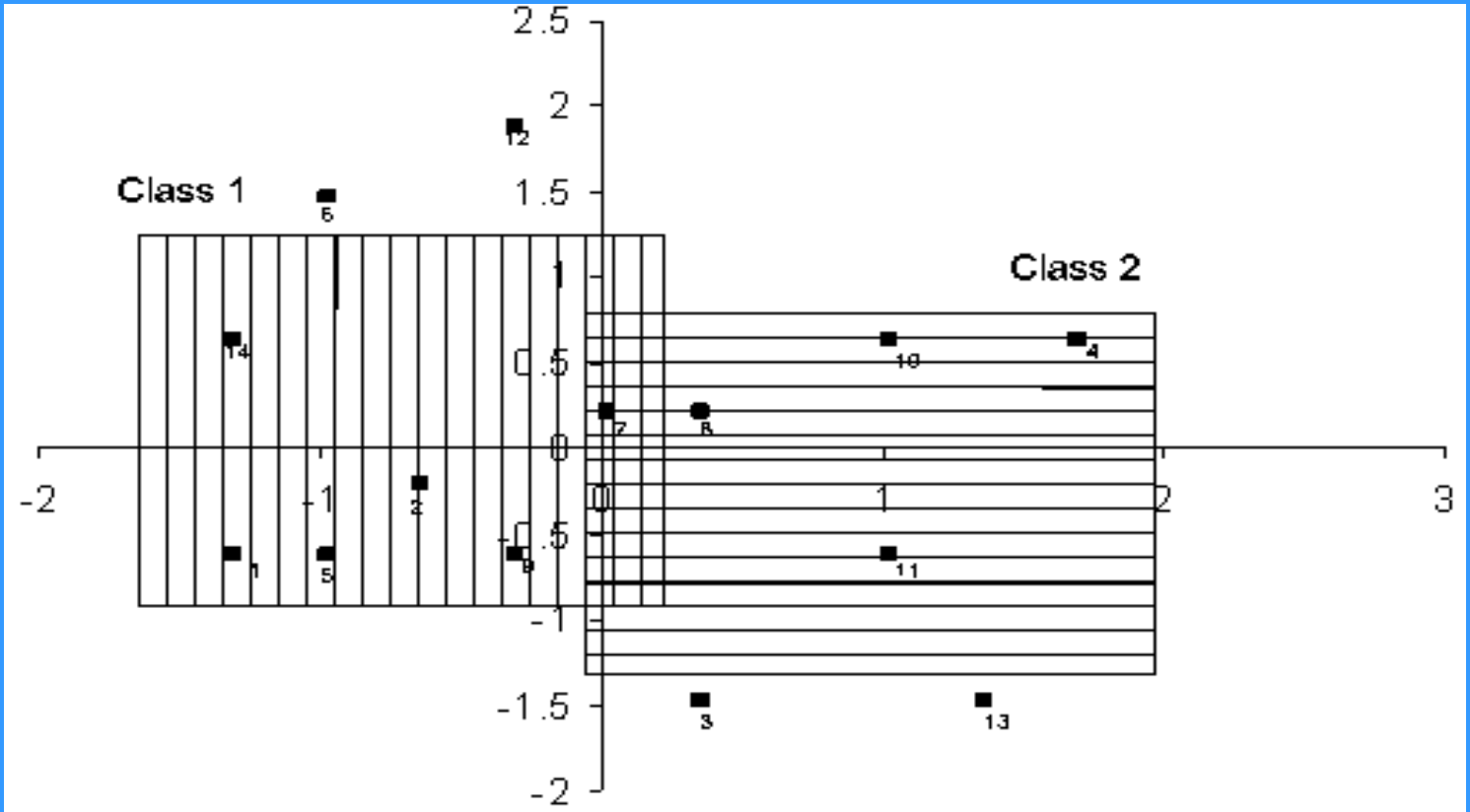
Results

Errors: 28% (4 input vectors out of 14 - Points 3, 6, 12 and 13 in fig.

RULES: **(Steepness=0)**

IF (x1 ≥ -1.76 AND ≤ 0.3) AND IF (x2 ≥ - 0.77 AND ≤ 1.29) THEN CLASS 1

IF (x1 ≥ -0.04 AND ≤ 1.98) AND IF (x2 ≥ - 1.36 AND ≤ 0.66) THEN CLASS 2.



Conclusions

1) After training the RBF classifier, the rules will be extracted through analyzing the parameters of the classifier.

2) One hidden unit corresponds to one rule.

3) It is desirable to reduce the number of hidden units of RBF neural networks while maintaining high classification accuracy.

4) The extracted rules can help discover and analyze the hidden knowledges in data sets further.

Thanks !