

# COMPUTATIONAL INTELLIGENCE: CURRENT STATE AND CHALLENGES

Pedro Larrañaga

`pedro.larranaga@fi.upm.es`  
`http://cig.fi.upm.es/`

Computational Intelligence Group  
Departamento de Inteligencia Artificial  
Universidad Politécnica de Madrid



EVIA 2014 · A Coruña, September 3, 2014

# Outline

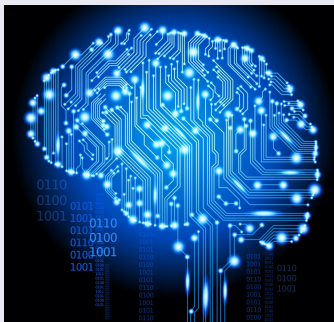
- 1 Computational Intelligence
- 2 Current State
- 3 Challenges
- 4 Conclusions

# Outline

- 1 **Computational Intelligence**
- 2 Current State
- 3 Challenges
- 4 Conclusions

# Computational Intelligence

## Towards a Definition



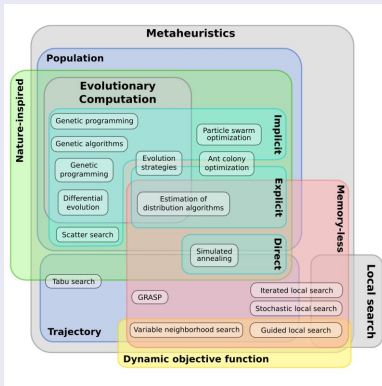
- Set of **nature-inspired computational methodologies** and approaches
- Address **complex real-world problems**
- **Traditional approaches**, i.e., explicit statistical modeling, are **ineffective or infeasible**
- Many such **real-life problems** are **not** considered to be **well-posed** problems mathematically



# Computational Intelligence

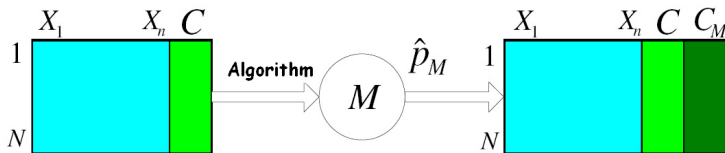
## Machine Learning + Metaheuristic-based Optimization

- **Machine learning** as the field of study that gives computers the ability to learn without being explicitly programmed (Samuel, 1959): **supervised classification, clustering, associations**
- **Metaheuristic-based optimization**



# Computational Intelligence. Machine Learning

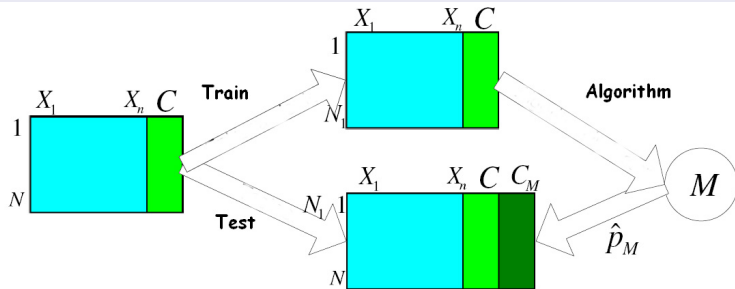
## Supervised classification. Estimation methods. No honest



$$\hat{p}_M = \frac{1}{N} \sum_{i=1}^N \delta(c^{(i)} = c_M^{(i)})$$

# Computational Intelligence. Machine Learning

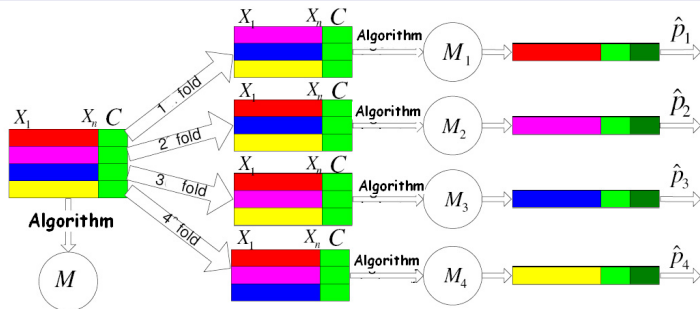
## Supervised classification. Estimation methods. Train and test



$$\hat{p}_M = \frac{1}{N - N_1} \sum_{i=1}^{N - N_1} \delta(c^{(N_1+i)} = c_M^{(N_1+i)})$$

# Computational Intelligence. Machine Learning

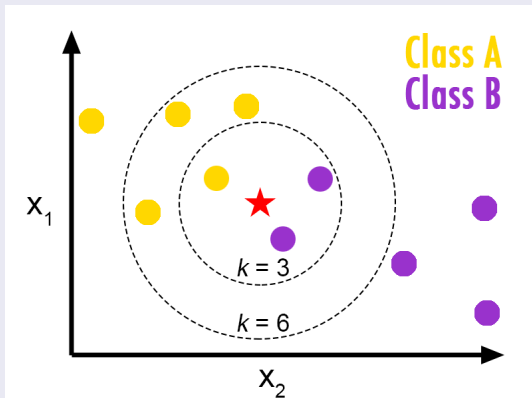
## Supervised classification. Estimation methods. $k$ -fold cross validation



$$\hat{p}_M = \frac{1}{k} \sum_{i=1}^k \hat{p}_i$$

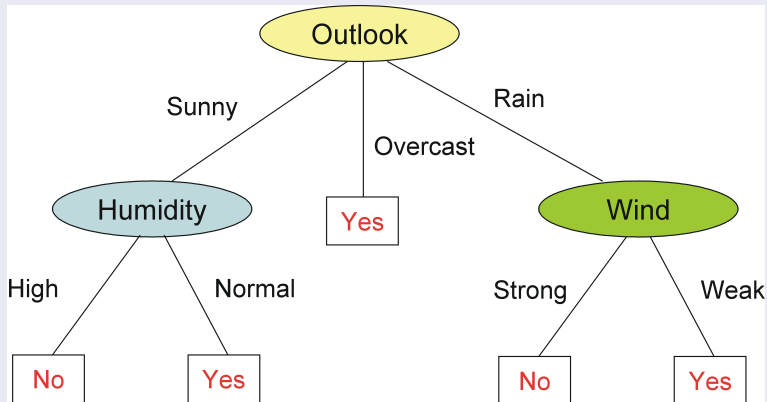
# Computational Intelligence. Machine Learning

## Supervised classification. $k$ -NEAREST NEIGHBORS



# Computational Intelligence. Machine Learning

## Supervised classification. CLASSIFICATION TREE



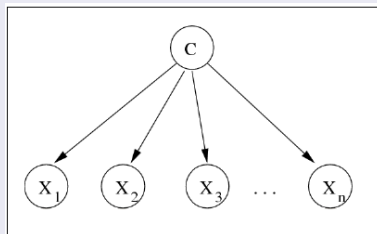
# Computational Intelligence. Machine Learning

## Supervised classification. NAIVE BAYES

Predictor variables are **conditionally independent** given  $C$

$$P(c|x_1, \dots, x_n) \propto P(C = c) \prod_{i=1}^n P(X_i = x_i | C = c)$$

$$\Rightarrow c^* = \arg \max_c P(C = c) \prod_{i=1}^n P(X_i = x_i | C = c)$$

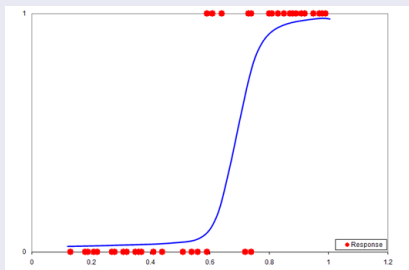


# Computational Intelligence. Machine Learning

## Supervised classification. LOGISTIC REGRESSION

$$\pi_j = P(C = 1 | \mathbf{x}_j) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{j1} + \dots + \beta_n x_{jn})}}$$

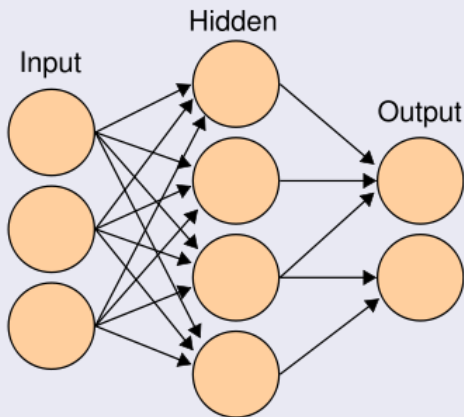
$$\Rightarrow 1 - \pi_j = P(C = 0 | \mathbf{x}_j) = \frac{1}{1 + e^{(\beta_0 + \beta_1 x_{j1} + \dots + \beta_n x_{jn})}}$$





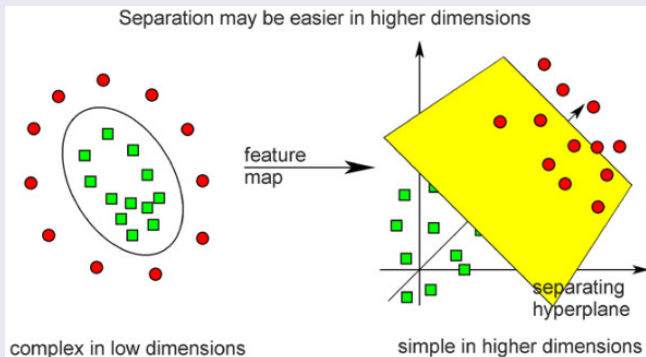
# Computational Intelligence. Machine Learning

## Supervised classification. NEURAL NETWORK



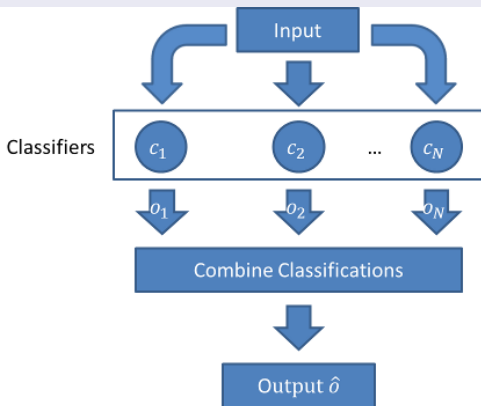
# Computational Intelligence. Machine Learning

## Supervised classification. SUPPORT VECTOR MACHINE



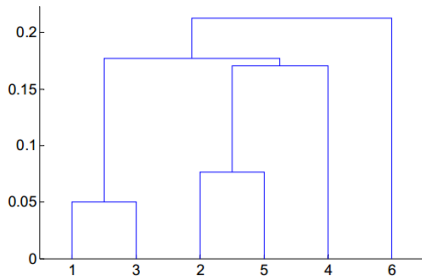
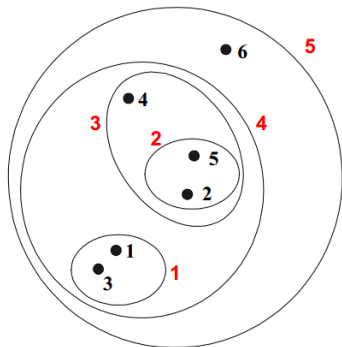
# Computational Intelligence. Machine Learning

## Supervised classification. METACLASSIFIERS



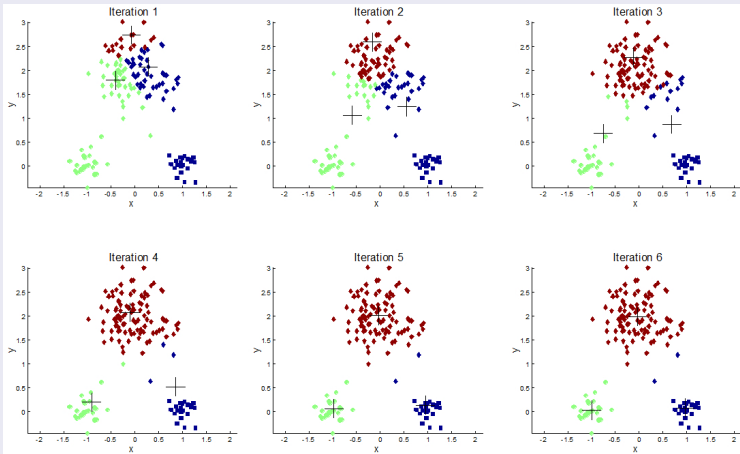
# Computational Intelligence. Machine Learning

## Clustering. HIERARCHICAL CLUSTERING



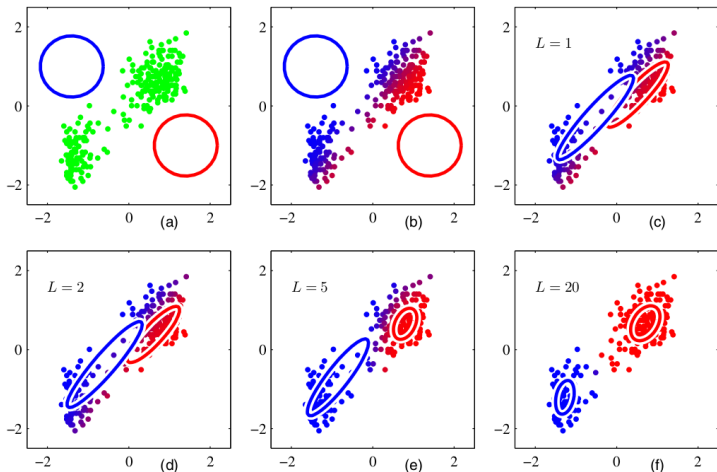
# Computational Intelligence. Machine Learning

## Clustering. PARTITIONAL CLUSTERING: $k$ -means



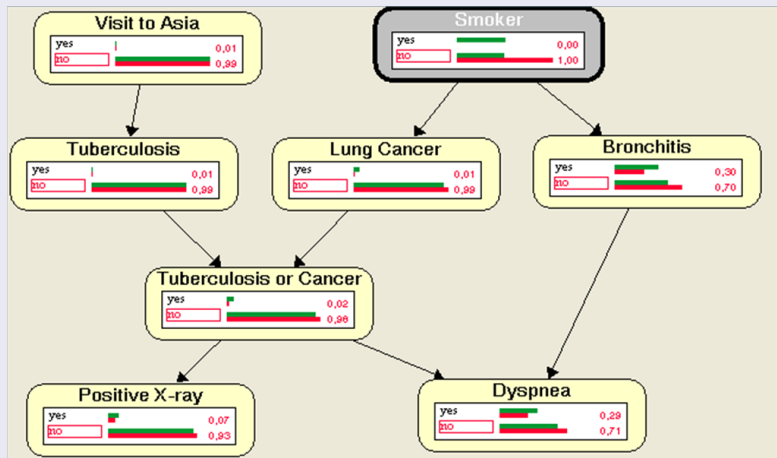
# Computational Intelligence. Machine Learning

## Clustering. PROBABILISTIC CLUSTERING: finite mixture models with EM



# Computational Intelligence. Machine Learning

## Associations. BAYESIAN NETWORKS



Evidence: "Smoker = no"

# Outline

- 1 Computational Intelligence
- 2 Current State**
- 3 Challenges
- 4 Conclusions



# Successful Applications

## CHESS











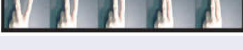
# Successful Applications

## JEOPARDY!



# Successful Applications

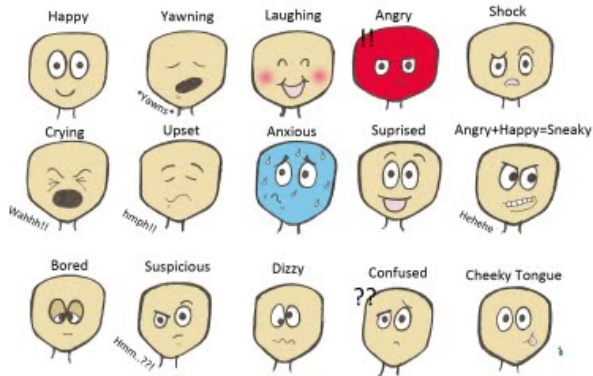
## GESTURE RECOGNITION

Flat/Leftward		class1
Flat/Rightward		class2
Flat/Contract		class3
Spread/Leftward		class4
Spread/Rightward		class5
Spread/Contract		class6
V-shape/Leftward		class7
V-shape/Rightward		class8
V-shape/Contract		class9

# Successful Applications

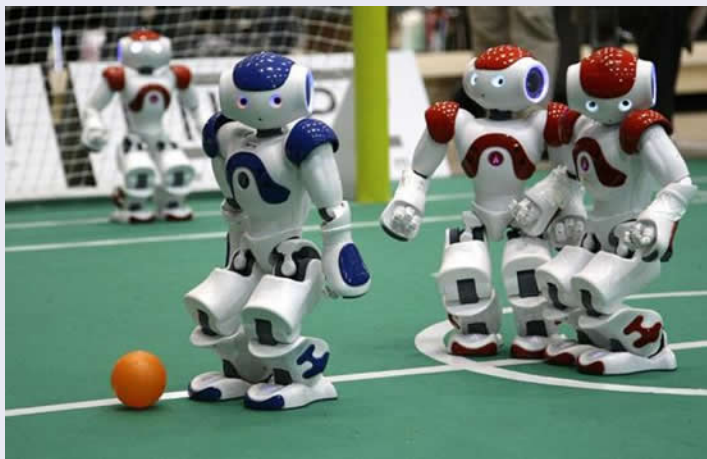
## FACIAL EXPRESSIONS

### Facial Expressions



# Successful Applications

## ROBOT SOCCER



# Successful Applications

## BRAIN COMPUTER INTERFACE



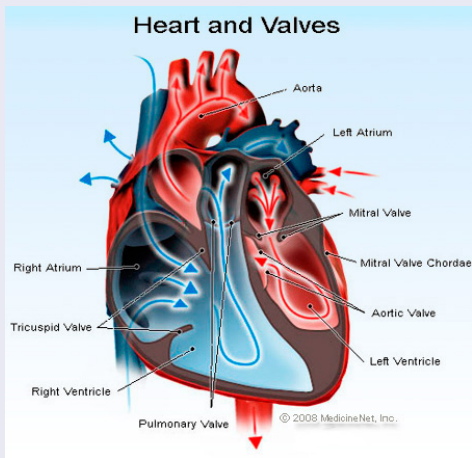
# Successful Applications

## AUTONOMOUS CAR



# Successful Applications

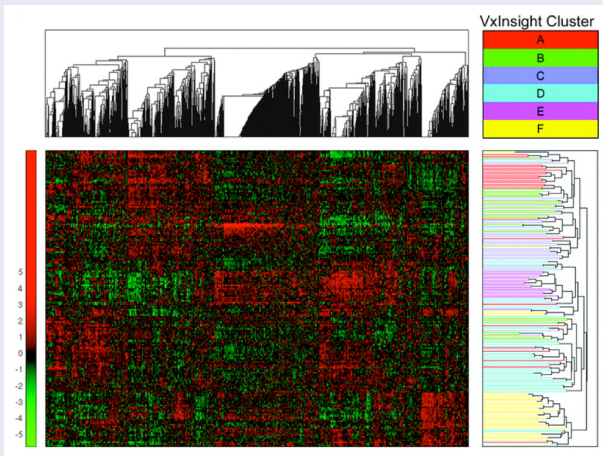
## MEDICAL DIAGNOSIS





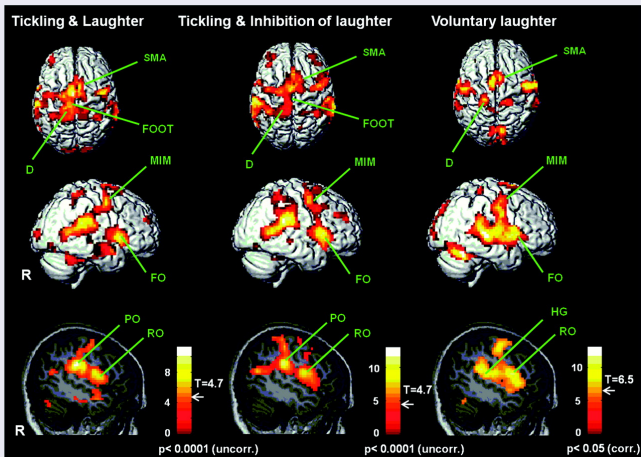
# Successful Applications

## BIOINFORMATICS



# Successful Applications

## NEUROSCIENCE



# Top 10 Algorithms in Data Mining

## Knowledge Information Systems (2008) 14:1-37

- 1 C4.5 (Quinlan, 1993)
- 2 The *k*-means algorithm (Lloyd, 1957)
- 3 Support vector machines (Vapnik, 1995)
- 4 The Apriori algorithm (Agraval and Srikant, 1994)
- 5 The EM algorithm (Dempster et al., 1977)
- 6 PageRank algorithm (Brin and Page, 1998)
- 7 AdaBoost (Freund and Schapire, 1995)
- 8 *k*-nearest neighbors (Fix and Hodges, 1951)
- 9 Naive Bayes (Minsky, 1961)
- 10 CART: Classification and Regression Trees (Breiman et al., 1984)

# Outline

- 1 Computational Intelligence
- 2 Current State
- 3 Challenges**
- 4 Conclusions

## 10 Challenging Problems



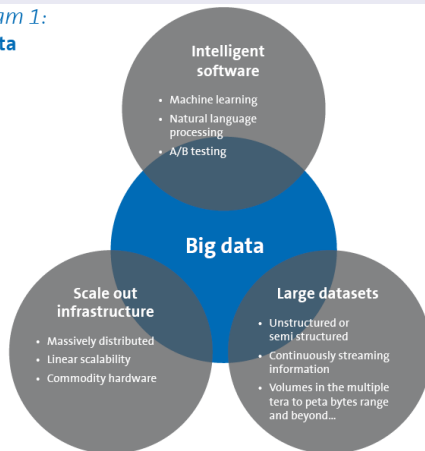
## 10 Challenging Problems



# 10 Challenges

## 1. Big Data: The five V's: volume, variety, velocity, viability and value

*Diagram 1:*  
**Big data**

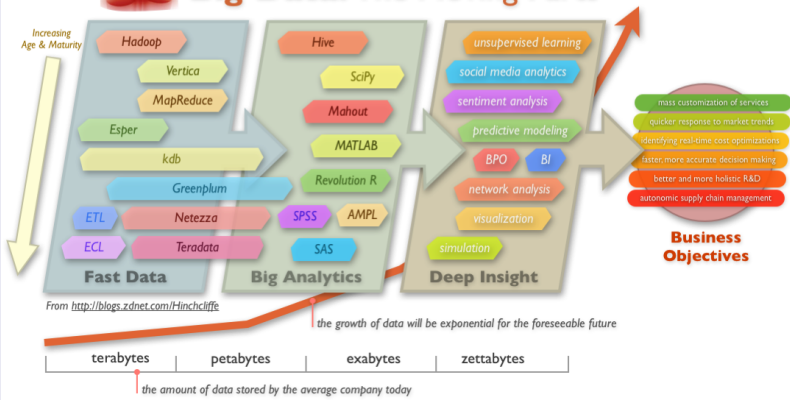


# 10 Challenges

## 1. Big Data. Adaptation of machine learning algorithms



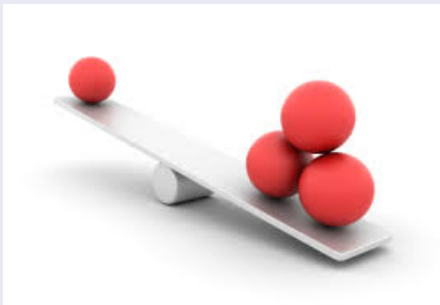
### Big Data: The Moving Parts





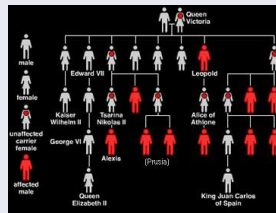
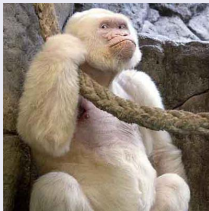
# 10 Challenges

## 2. Variants of Supervised Classification. Class imbalance



# 10 Challenges

## 2. Variants of Supervised Classification. Positive Labels



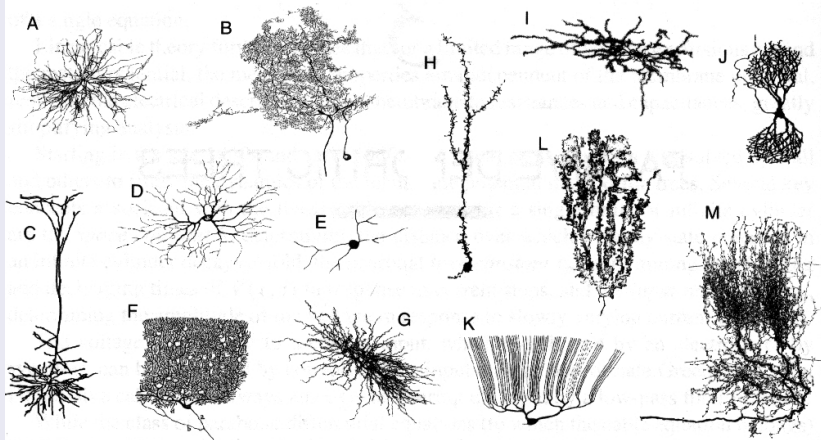
# 10 Challenges

## 2. Variants of Supervised Classification. Positive Labels

genes	$X_1$	...	$X_n$	$C$
$g^{(1)}$	$x_1^{(1)}$	...	$x_n^{(1)}$	+
$g^{(2)}$	$x_1^{(2)}$	...	$x_n^{(2)}$	+
$g^{(3)}$	$x_1^{(3)}$	...	$x_n^{(3)}$	?
$g^{(4)}$	$x_1^{(4)}$	...	$x_n^{(4)}$	+
...	...	...	.....	
$g^{(N-1)}$	$x_1^{(N-1)}$	...	$x_n^{(N-1)}$	?
$g^{(N)}$	$x_1^{(N)}$	...	$x_n^{(N)}$	?

# 10 Challenges

## 2. Variants of Supervised Classification. Semi-supervised

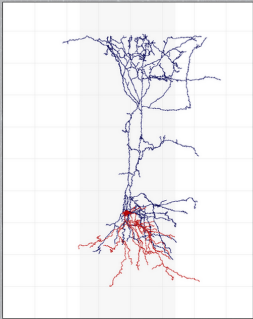


# 10 Challenges

## 2. Variants of Supervised Classification. Semi-supervised

A GARDENER CLASSIFICATION Home Log out Help

**Neuron 3/320**  
Mouse, Visual, Layer V (150-300µm)



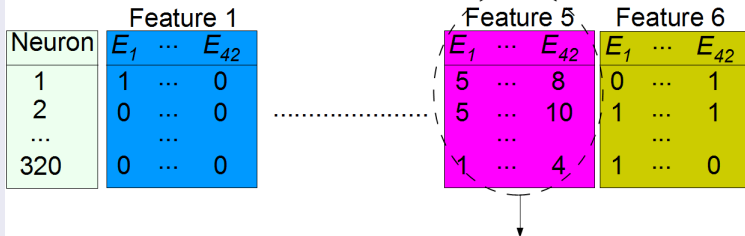
3D Visualization

- 1.  Intralaminar  Translaminar
- 2.  Intracolumnar  Transcolumnar
- 3.  Centered  Displaced
  - 4.  Ascending  Descending  Both
- 5.  Arcade  Cajal-Retzius  Chandelier  Common Basket  Horse-tail  Large Basket  Martinotti  Neurogliaform  Common type  Other
- 6.  Uncharacterized: not enough morphological axonal features

Neuron1  
Neuron2  
Neuron3  
Neuron4  
Neuron5  
Neuron6  
Neuron7  
Neuron8  
Neuron9  
Neuron10  
Neuron11  
Neuron12  
Neuron13  
Neuron14  
Neuron15  
Neuron16  
Neuron17  
Neuron18  
Neuron19  
Neuron20  
Neuron21  
Neuron22  
Neuron23  
Neuron24  
Neuron25  
Neuron26  
Neuron27  
Neuron28  
Neuron29  
Neuron30  
Neuron31  
Neuron32  
Neuron33  
Neuron34  
Neuron35  
Neuron36  
Neuron37  
Neuron38  
Neuron39  
Neuron40  
Neuron41  
Neuron42  
Neuron43  
Neuron44  
Neuron45  
Neuron46

# 10 Challenges

## 2. Variants of Supervised Classification. Semi-supervised



# 10 Challenges

## 2. Variants of Supervised Classification. Semi-supervised

Neuron	Feature 1			...	Feature 5			...	Feature 6		
	$E_1$	...	$E_{42}$		$E_1$	...	$E_{42}$		$E_1$	...	$E_{42}$
1	1	...	0		5	...	8		0	...	1
2	0	...	0		5	...	10		1	...	1
...	...	...	...		...	...	...		...	...	...
320	0	...	0		1	...	4		1	...	0

Neuron	128 Axon, 86 Dendrite, 10 Soma				$\geq 26$ votes
	$X_1$	$X_2$	...	$X_{224}$	<b>C</b>
1	8.4	7.1	...	2.1	Martinotti (27)
2	2.0	5.4	...	4.0	?
3	7.3	2.0	...	1.1	?
...	...	.....	...	...	...
<b>240</b>	3.7	2.8	...	5.2	Basket (32)

# 10 Challenges

## 2. Variants of Supervised Classification. Probabilistic labels

cell	morpholog. variables			classif. 42 experts			class
	$X_1$	...	$X_{2885}$	$E_1$	...	$E_{42}$	$C$
1	10.1	...	6.6	trans	...	intra	38-4
2	3.7	...	7.7	intra	...	trans	24-18
3	5.9	...	9.2	intra	...	intra	35-7
4	11.2	...	10.1	intra	...	intra	10-32
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
240	13.6	...	5.7	intra	...	intra	3-39



# 10 Challenges

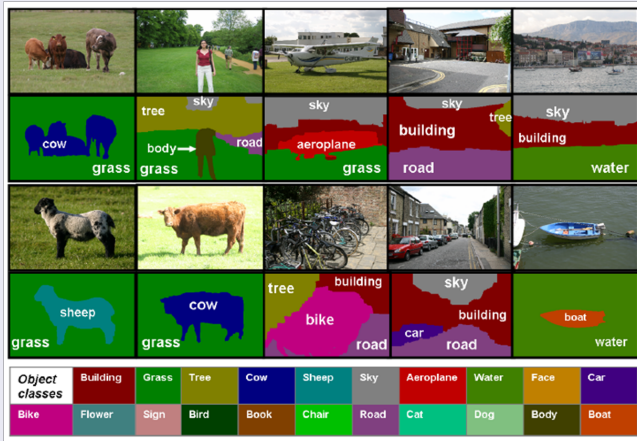
## 3. Multitarget Prediction. Multilabel Classification

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$C$
3.2	1.4	4.7	7.5	3.7	1
2.8	6.3	1.6	4.7	2.7	0
7.7	6.2	4.1	3.3	7.7	1
9.2	0.4	2.8	0.5	3.9	0
5.5	5.3	4.9	0.6	6.6	1

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$C_1$	$C_2$	$C_3$	$C_4$
3.2	1.4	4.7	7.5	3.7	1	0	1	1
2.8	6.3	1.6	4.7	2.7	0	0	1	0
7.7	6.2	4.1	3.3	7.7	1	0	1	1
9.2	0.4	2.8	0.5	3.9	0	1	0	0
5.5	5.3	4.9	0.6	6.6	1	1	0	1

# 10 Challenges

## 3. Multitarget Prediction. Multilabel Classification



# 10 Challenges

## 3. Multitarget Prediction. Multilabel Classification

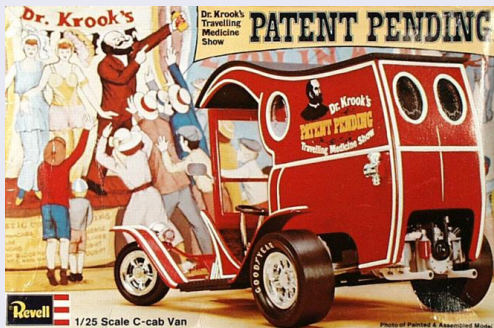
© Original Artist  
 Reproduction rights obtainable from  
[www.CartoonStock.com](http://www.CartoonStock.com)



"That's not what it says on the Web."

© Mike Salsbery / Corbis

search | id: rmban1897



# 10 Challenges

## 3. Multitarget Prediction. Multidimensional Classification



# 10 Challenges

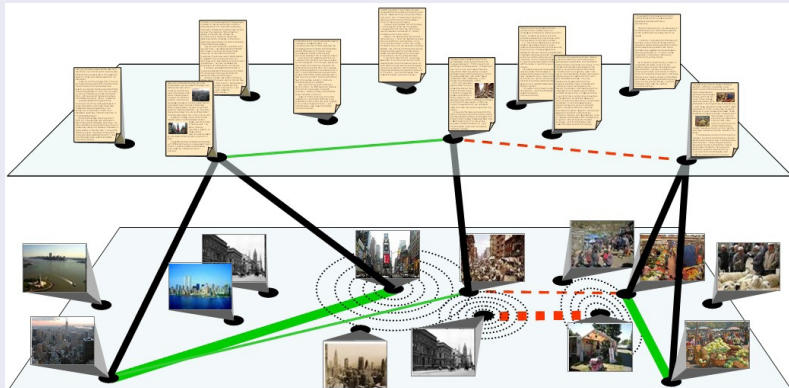
## 3. Multitarget Prediction. Multioutput Regression

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$Y$
3.2	1.4	4.7	7.5	3.7	1.7
2.8	6.3	1.6	4.7	2.7	0.4
7.7	6.2	4.1	3.3	7.7	1.9
9.2	0.4	2.8	0.5	3.9	0.2
5.5	5.3	4.9	0.6	6.6	1.7

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$Y_1$	$Y_2$	$Y_3$	$Y_4$
3.2	1.4	4.7	7.5	3.7	1.8	0.2	1.8	1.2
2.8	6.3	1.6	4.7	2.7	0.3	0.4	1.1	0.1
7.7	6.2	4.1	3.3	7.7	1.3	0.4	1.1	1.9
9.2	0.4	2.8	0.5	3.9	0.7	1.1	0.1	0.9
5.5	5.3	4.9	0.6	6.6	1.1	1.2	0.7	1.2

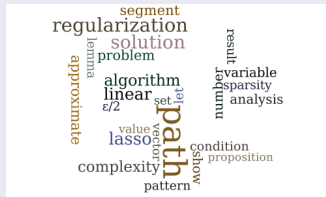
# 10 Challenges

## 4. Clustering. Multipartition. Subspace Clustering



# 10 Challenges

## 5. Regularization



lasso: 
$$\hat{\beta} = \operatorname{argmin}_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_1$$

adaptive lasso: 
$$\hat{\beta} = \operatorname{argmin}_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \sum_{j=1}^p w_j |\beta_j|$$

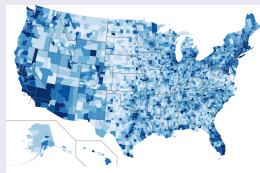
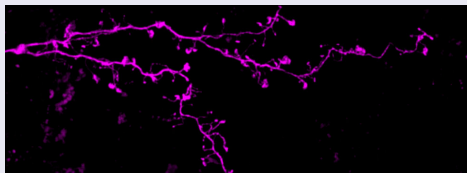
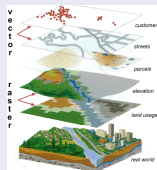
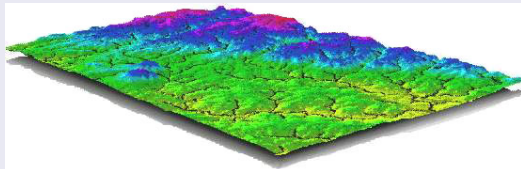
Dantzig: 
$$\hat{\beta} = \operatorname{argmin}_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_{\infty} + \lambda \|\beta\|_1$$

elastic net: 
$$\hat{\beta} = \operatorname{argmin}_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2$$

group lasso: 
$$\hat{\beta} = \operatorname{argmin}_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \sum_{j=1}^J \|\beta_j\| w_j$$

# 10 Challenges

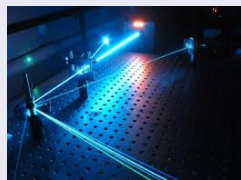
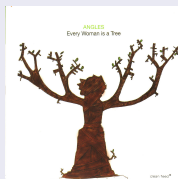
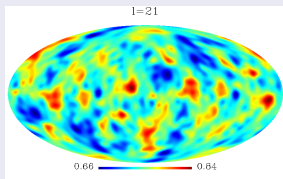
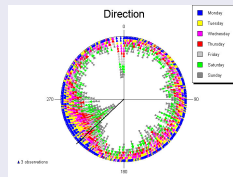
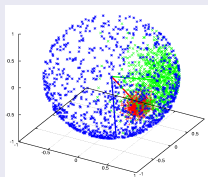
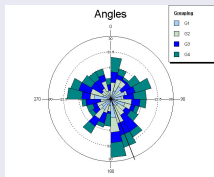
## 6. Spatial Data





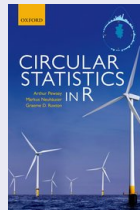
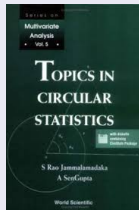
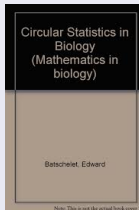
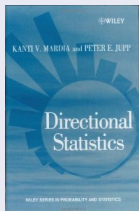
# 10 Challenges

## 7. Directional Data



# 10 Challenges

## 7. Directional Data. Books



# 10 Challenges

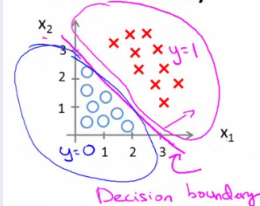
## 8. Data Streams. Concept Drift



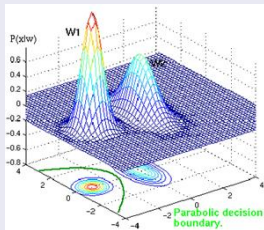
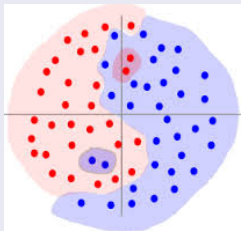
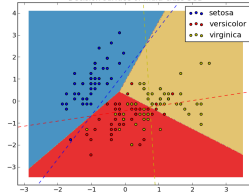
# 10 Challenges

## 9. Performance Measures. Decision Surfaces

Decision Boundary

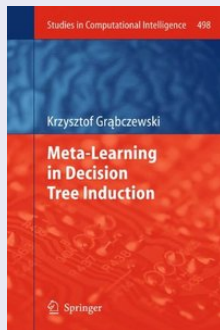
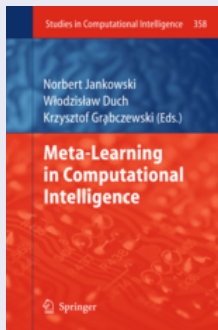


Decision surface of multi-class SGD



# 10 Challenges

## 10. Metalearning



# Outline

- 1 Computational Intelligence
- 2 Current State
- 3 Challenges
- 4 Conclusions**

# COMPUTATIONAL INTELLIGENCE

## 10 CHALLENGES

- 1 Big data
- 2 Variants of supervised classification
- 3 Multitarget prediction
- 4 Clustering
- 5 Regularization
- 6 Spatial data
- 7 Directional data
- 8 Data streams
- 9 Performance measures
- 10 Metalearning

# COMPUTATIONAL INTELLIGENCE: CURRENT STATE AND CHALLENGES

Pedro Larrañaga

`pedro.larranaga@fi.upm.es`  
`http://cig.fi.upm.es/`

Computational Intelligence Group  
Departamento de Inteligencia Artificial  
Universidad Politécnica de Madrid



EVIA 2014 · A Coruña, September 3, 2014