HUGE

Hello

Softcomputing and machine learning

July 26, 2018

Agenda.

1. Definitions
2. How can ML be seen?
3. Problems
3. Approaches
4. Techniques
5. Libraries
6. The course

William E. Gómez

Web Engineer at HUGE
PhD. Student at Universidad de Antioquia
Adjunct professor at Universidad de Antioquia

¿What was the most intelligent you did the last week?

What about:
Made the breakfast
University coming
Identify your mate

¿Are these intelligent activities? ¿Are they examples of intelligence?

¿What is the intelligence?

Definition:

A very general mental capacity, which allows to reason, plan, solve problems, think abstractly, understand complex ideas, learn quickly, and learn from experience.

More than a simple encyclopedic knowledge

When would we give the qualifier of intelligent to a program / machine?

Is a program that "plays chess" smart?

Is a program that diagnoses cancer intelligent?

Is a program that answers questions in natural language intelligent?

Systems that think like humans

Bellman (1978): IA is the automation of activities that we link with human thought processes, activities such as decision making, problem solving, and learning. Cognitive science: understanding human intelligence.

Systems that act like humans

Minsky (1986): IA is the art of building machines capable of doing things that would require intelligence in case they were made by human beings. Natural language, vision, robotics, expert systems, etc.

Systems that act rationally

Luger (1989): IA is the branch of computing that deals with the automation of intelligent behavior. Agents (interaction, reasoning, etc.)

Systems that think rationally, that is, that do the right thing.

Winston (1992): IA deals with the study of computations that allow to perceive, reason and act. Logic and automatic reasoning.

Study of cognitive processes in general.

- Model human intelligence, brain functioning.
- Study of intelligence as computing

Construcción de sistemas automáticos capaces de llevar a cabo tareas hasta ahora reservadas exclusivamente a los seres humanos.

Emulate the capacities of living beings through computers.

Existing intelligent systems lack common sense and the generality of beings.

1

Definitions

Artificial intelligence

Study of design of intelligence agents to create machines that can mimic human intelligence.

Soft computing

It is a subdiscipline of AI that focuses on heuristics, imperfect solutions to complex problems. Uncertainty.

Machine Learning

To make the machine learn by itself to solve the problems using a large quantity of data.

Greek philosophers, degrees of uncertainty

Century XIX and half XX: Biological and phenomenological analogies, notion of computation.

1943 McCulloch and Pitts: Neurological properties of neurons.

1949 Hebb: Learning by synapse modification.

1950 Alan Turing: Turing machine.

1950-1960: John Von Neumann: computers should be designed taking as model the human brain, physical-chemical structure.

1955: Study the brain functions, capabilities as an information processor.

McCulloch

1956 Congress in Darthmouth: Founding Fathers of the discipline.

Thought can occur outside the brain, in machines

Thought can be understood formally and scientifically

The best way to understand it is through digital computers

1958 Newell, Shaw and Simon: First intelligent program based on its information processing model. Dominant theory in cognitive psychology.

1958 Rosenblat. The perceptron.

1959 Minsky: Imitation of the brain at the molecular level should be abandoned

1962 McCarthy and Raphael: Home mobile robot "Shakey". Learn from experience, recognize visual patterns, modeling, manipulating symbols, etc.

1965 Zadeh: Fuzzy sets (Imprecision increases according to the complexity of the system).

1968 Russel and Black: Vagueness term.

1969 Minsky and Papert. Limitations of the perceptron.

Newell-Shaw-Simon: First "intelligent" machine. Capable of memorizing and learning, was able to demonstrate some of the theorems.

Early 70s some U. Japanese; Study of Fuzzy Logic

1974 Mandani: First fuzzy controller

Circa 1975. Interest by different areas.

1977 Boden: Programs with linguistic abilities. Inferences and "phrases" about the known world. Answering is nof the same as understanding.

Dendral -Stanford: Medical diagnoses-human error rate.

1979 McCorduck: First program capable of learning from experience.

1980 Mandani controller application in cement factory in Denmark.

1981 Boom Industrial applications, Japanese (smart computers), USA (human interfaces).

1982 Hopfield: Dynamic networks with symmetric weights.

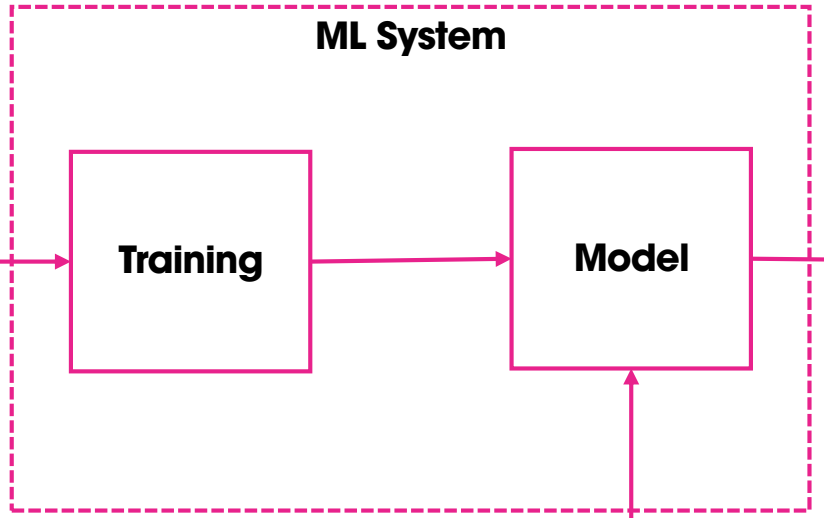
1982 Kohonen: Self-organizing maps.

1983 Barto, Sutton and Anderson: Learning by reinforcement.

1998 Bartlett: Theory of RNA generalization.

So what is ML actually?

Samples used to learn



Prediction

Cat or Dog?

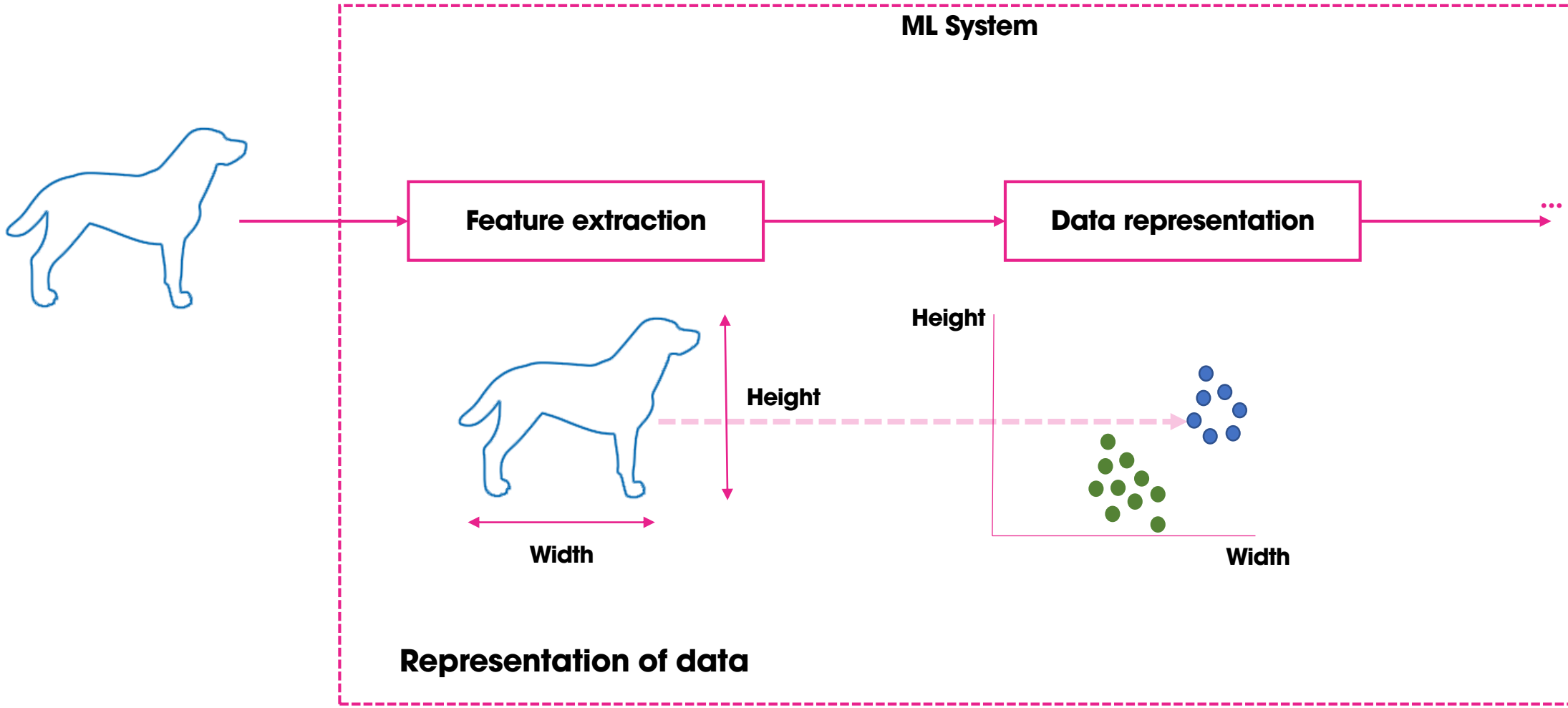
Cat !

New individual

How can ML be seen?

2

How can ML be seen?



How can ML be seen?

ML System

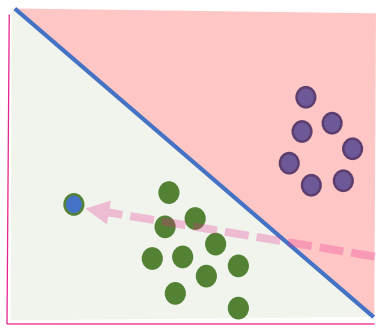
Algorithm (Training)

Model

Prediction

Cat or Dog?

Cat !

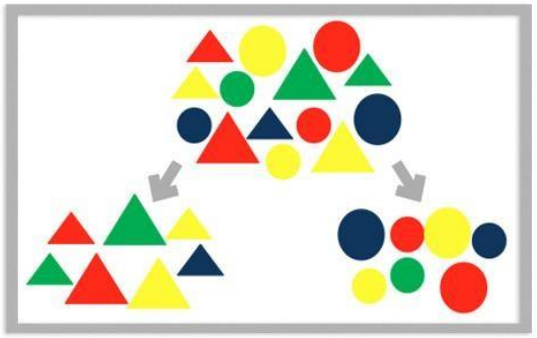


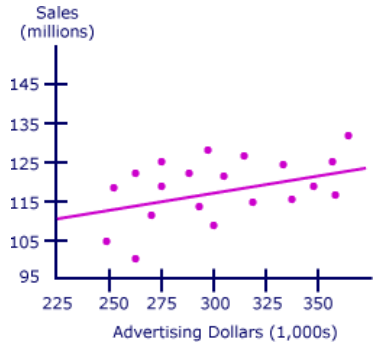
New individual

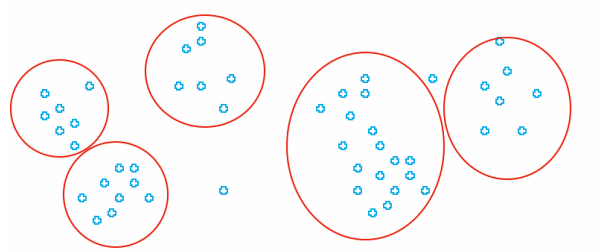
3

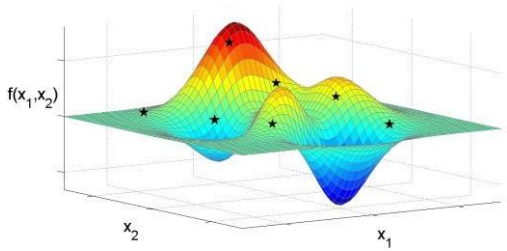
Problems

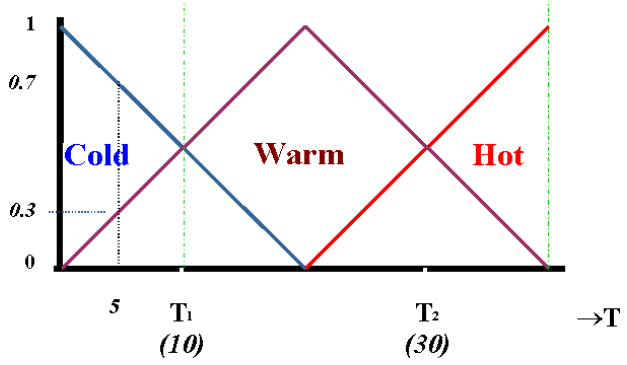
Problems we can find:

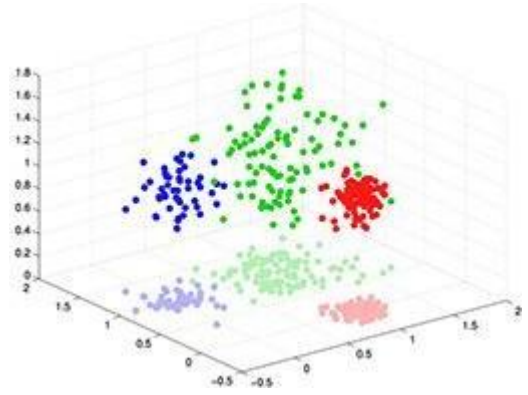












4

Approaches

Supervised

There is an expert knowledge that is desired to reproduced.

Labels

$$f^*\left(\begin{bmatrix}\vec{X}_1\\ \vec{X}_2\\ \vdots\\ \vec{X}_k\end{bmatrix},\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_k\end{bmatrix}\right)=\begin{bmatrix}y_1^*\\ y_2^*\\ \vdots\\ y_k^*\end{bmatrix} \qquad \min\left(\begin{bmatrix}y_1^*\\ y_2^*\\ \vdots\\ y_k^*\end{bmatrix}-\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_k\end{bmatrix}\right)$$

Prediction

Unsupervised

There is an expert knowledge that is desired to reproduced.

$$f\begin{bmatrix}\vec{X}_1\\ \vec{X}_2\\ \vdots\\ \vec{X}_k\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_k\end{bmatrix}$$

~~Labels~~

Prediction

5

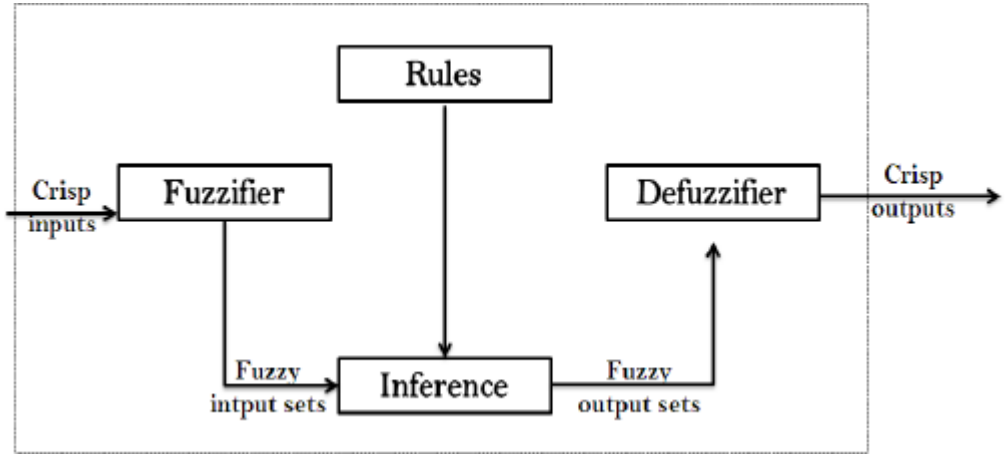
Techniques

1. Fuzzy logic.
2. Regression.
3. Classification.
4. Evolutionary algorithms.
5. Clustering.
6. Dimensionality reduction.
7. Feature selection.

Fuzzy logic

Aproximation to human reasoning, management of uncertainty in decisions.

Representation of knowledge.



Controllers (cars, planes, altitude, traffic)

Autonomous systems

Disease detection

In itself, any system to which rules can be applied.

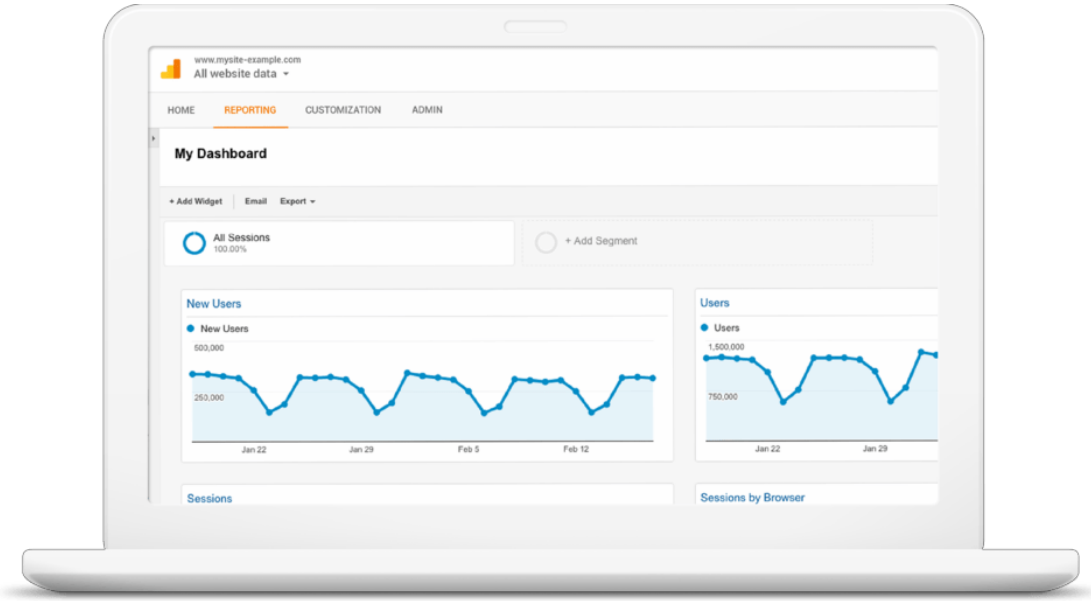
Applications video:

https://www.youtube.com/watch?v=2d_7GqoINJg

Regression

To find relation between to or more variables.

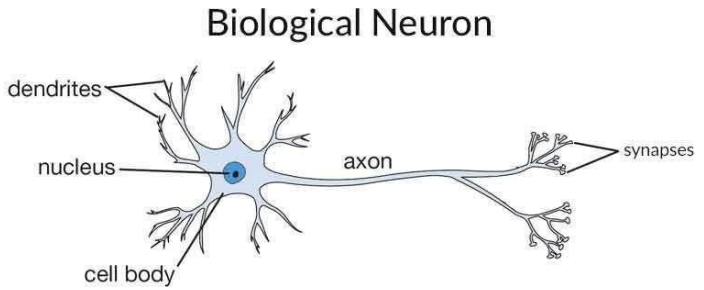
Forecasting future opportunities.
Predicting house costs.
Predicting forest fires.
Estimate web traffic

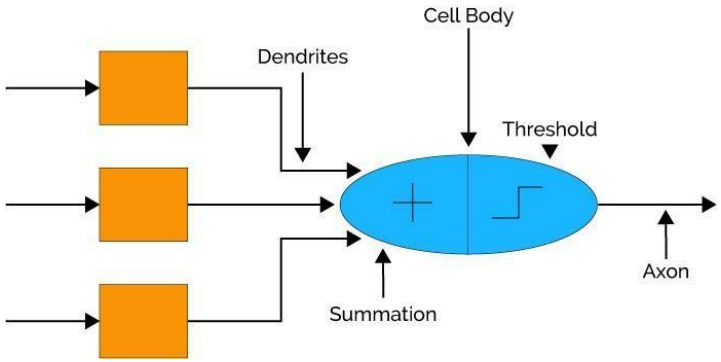


Classification

From a set of features the algorithms discriminate between classes.

- Spam or not spam.
- Failure prediction.
- Diseases prediction.
- Object detection.
- Face recognition





Applications video:

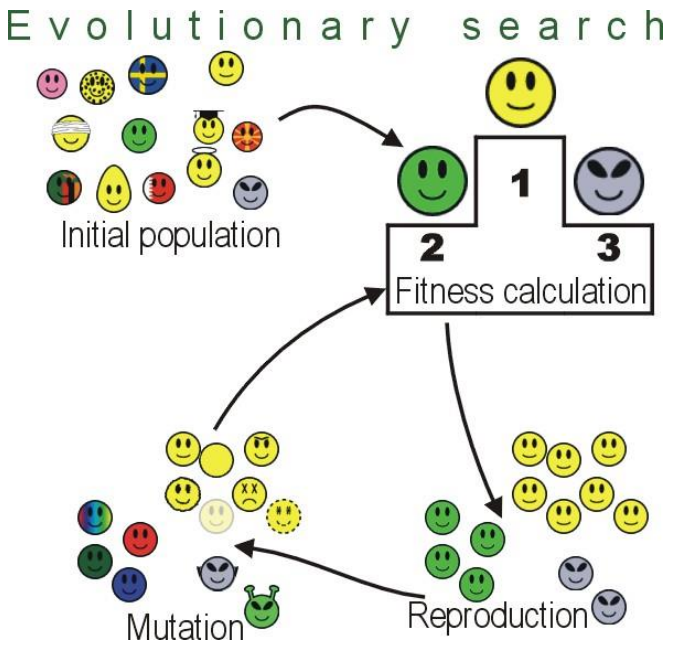
https://www.youtube.com/watch?v=20dErCwfxTY

https://www.youtube.com/watch?v=hPKJBXkyTKM

Evolutionary algorithms

Search for the optimal response by mixing the best "parents". Natural selection theory.

- **Automatic design**
- **Robotics**
- **Optimization (Connection routing, traffic)**
- **Computer games**
- **Strategies**

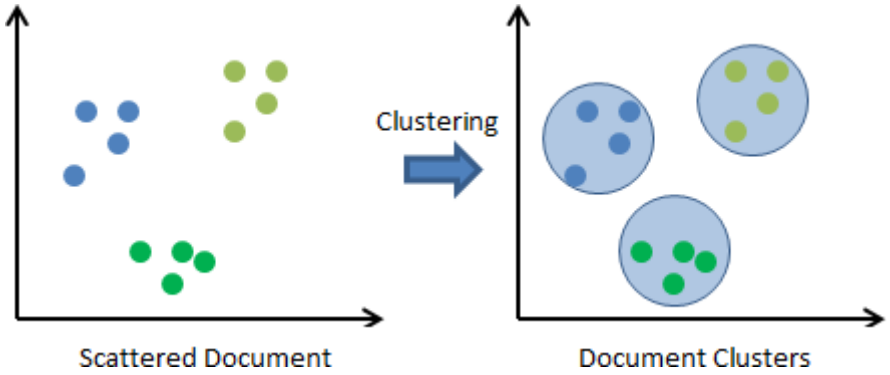


Applications video:

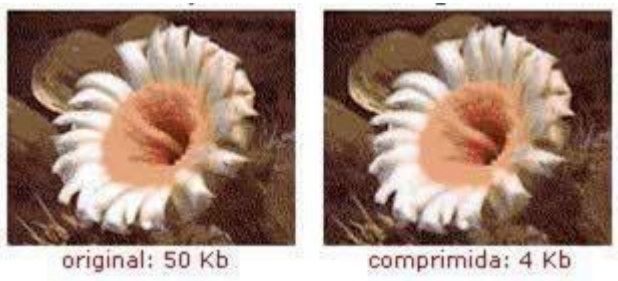
https://www.youtube.com/watch?v=yQTurXpXd1M

Clustering

Identify similarities between data and identify "natural" groups in the data.



Search results
Customer segmentation
Discovering patterns
Compression of information

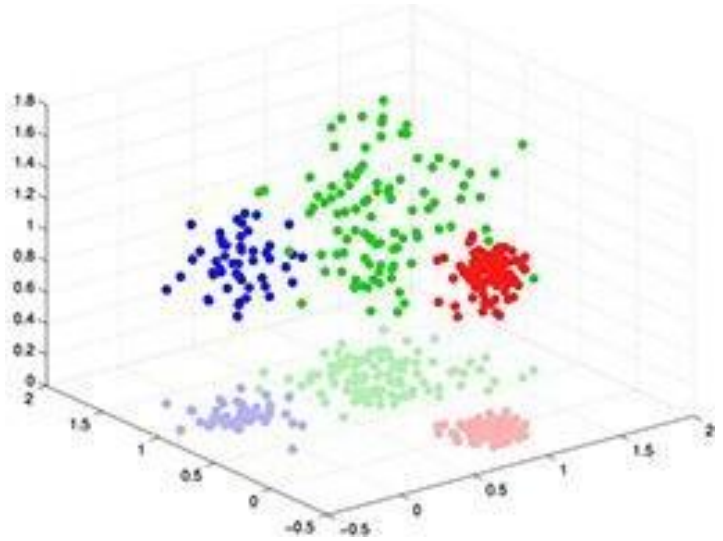


Dimensionality reduction

Reduce the number of variables under consideration, but representing the same information.

Creating indicators.

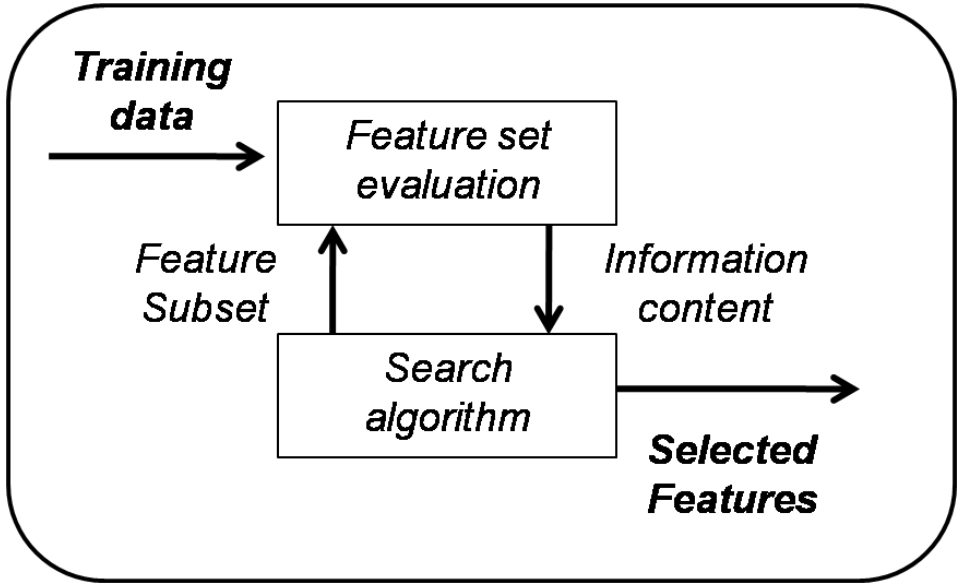
Visualize the information.



Feature selection

Selection of a smaller group of descriptors, keeping those that improved a task.

Improve performance of the algorithms.
Hypothesis generation.



6

Libraries

SciKit-Learn

Features:

Classification, Regression, Clustering, Dimensionality reduction, Model selection, Preprocessing

TensorFlow

https://www.tensorflow.org

Features:

Neural networks, Deep learning, Mathematics

Requisites:

NumPy

SciPy

Pandas

Matplotlib

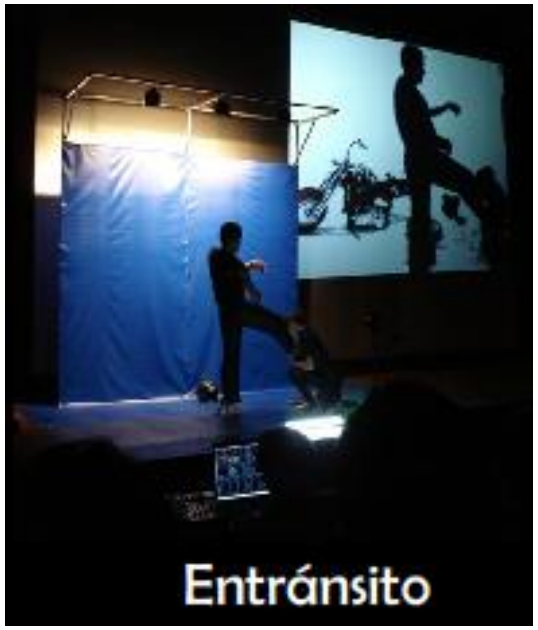
1

The course

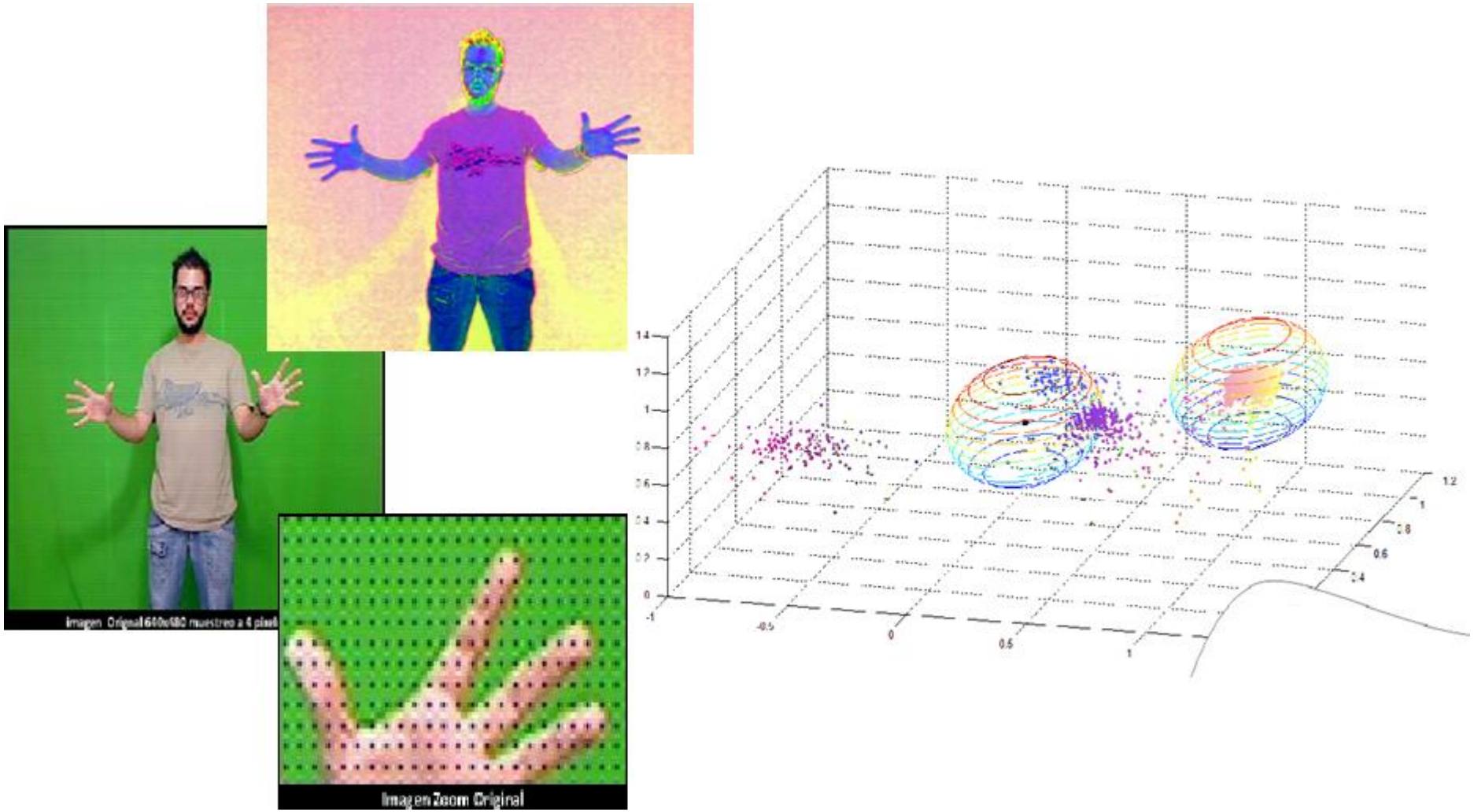
Course applications

Face and voice recognition (Cesar Uribe)
Background extraction algorithm (Lukas)
Position control (height) helicopter
Light control system using the voice (Efrain)
Web Intrusion Identification System (Andrés)
Cell identification system (Sebastián)
Helicopter control (Angela)
Inverted Pendulum Control: Fuzzy - RNA
RNA Gyroscope Control
Musical genres identification (Alejandro Arboleda)
Video game Powers selection using RNA (GEPAR students)
Car control to avoid obstacles (Giovanni-David)

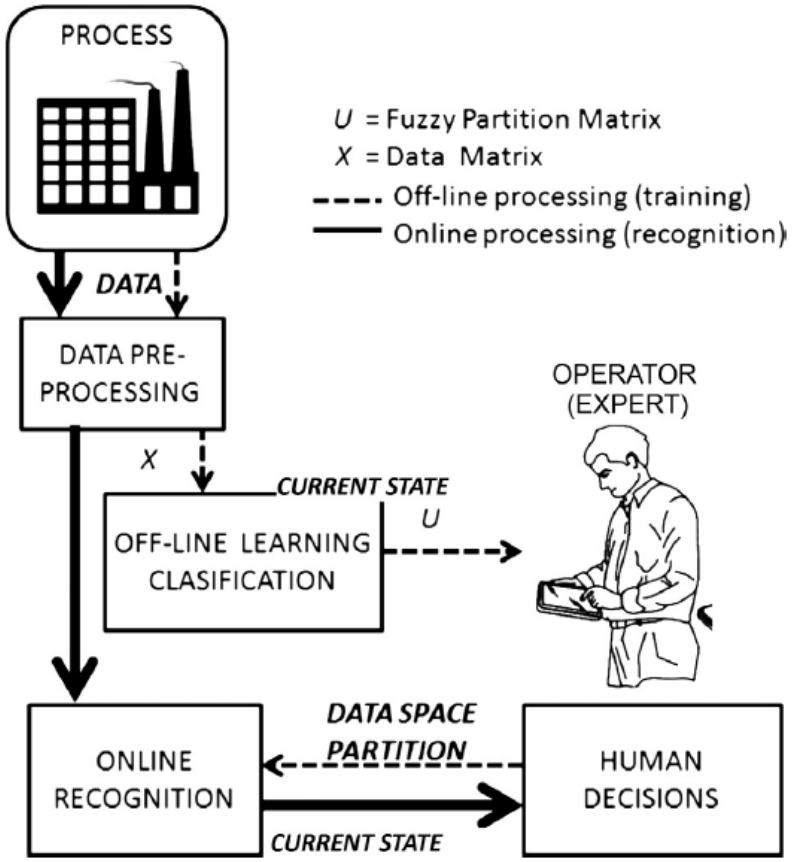
FC Means: background extraction



FC Means: background extraction



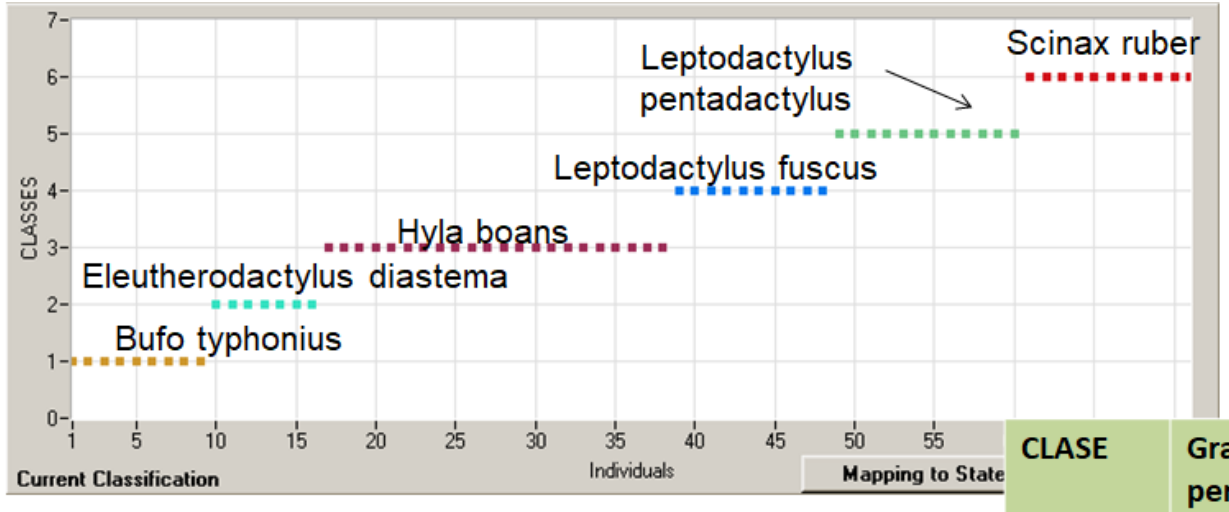
Automobile Diagnostics using clustering

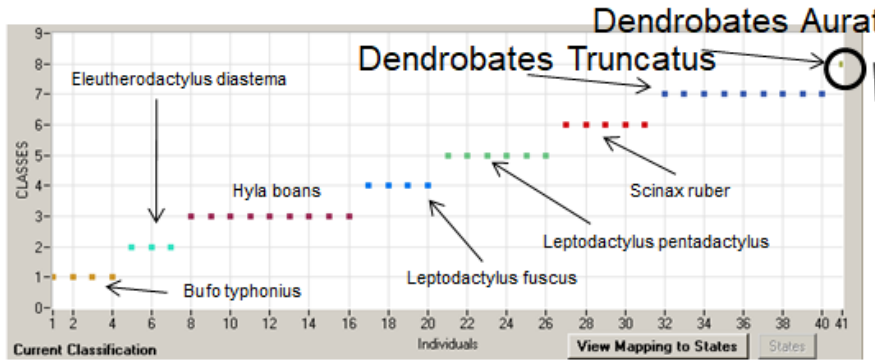


C. Isaza, H. Sarmiento, T. Kempowsky and M.V. Le Lann "Situation Prediction based on Fuzzy Clustering for Industrial Complex Processes", Information Sciences, ISSN: 0020-0255 Elsevier, Vol.279, 2014, pp. 785-804

Anuran recognition

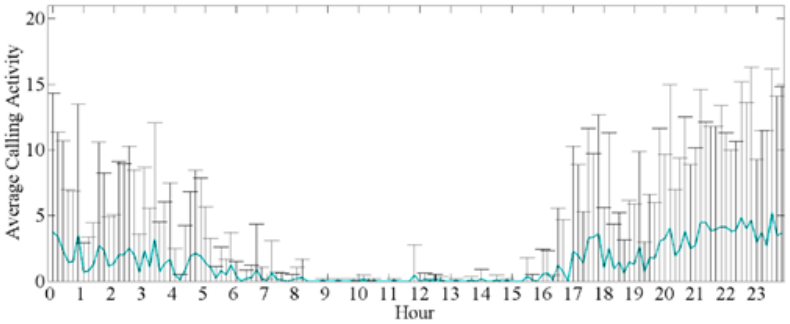
Anuran recognition





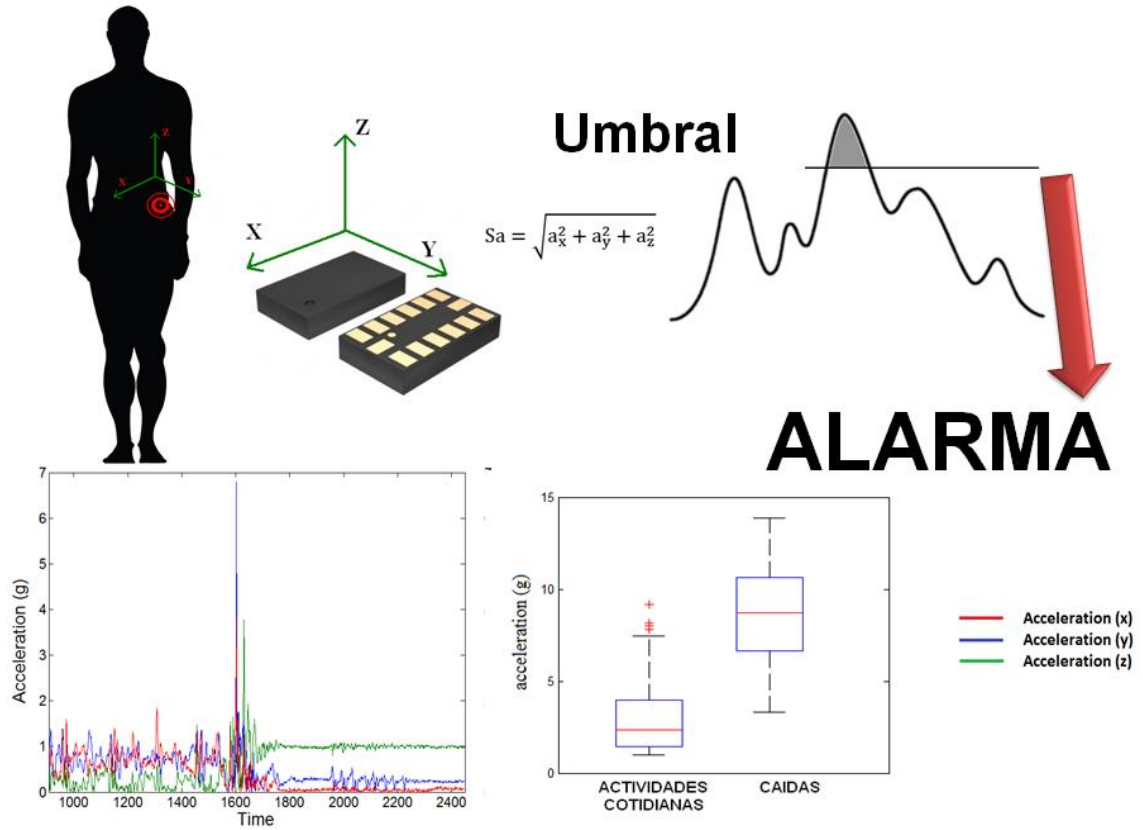
CLASE	Grado de pertenencia $x10^{-4}$
1 BT	0.015
2 ED	0.012
3 HB	0.001
4 LF	0.000
5 LP	0.000
6 SR	0.060
7 DT	**0.062**
8 DA	**0.451**

C. Bedoya, C. Isaza, J. Daza, J.D. Lopez. Automatic recognition of anuran species based on syllable identication. Ecological Informatics, Vol. 24, 200-209, 2014.



(a) *Diasporus gularis*

Fall detection using acelerometers



¿What do you propose?

Where to find ideas?

Youtube examples:

http://www.youtube.com/watch?v=KHV7fWvnn_0&feature=related
http://www.youtube.com/watch?v=3Dvw23AMZI4
http://www.youtube.com/watch?v=lPb2wV82D-A
http://www.youtube.com/watch?v=cFcQpC8vAr8&feature=related
http://www.youtube.com/watch?v=SVKKPD23eM4&feature=related
http://www.youtube.com/watch?v=OB6UMe9pyfI&feature=related
http://www.youtube.com/watch?v=2vueH30TbT4

Other pages:

https://www.ted.com/topics/robots
https://www.kaggle.com/competitions
https://www.kaggle.com/datasets
https://archive.ics.uci.edu/ml/index.php

Course methodology

1. Thirty-two academic meetings to present:
   The general concepts of the subject
   Solve some examples
   Assign jobs.
2. Four application works which will be developed in groups of two people.
3. During the semester, each working group will advance a small project applying the concepts seen in the course to a problem in the area of interest of each group: This project will have partial deliveries and a sustainment of results (conference style) at the end of the semester.
4. In the written documents and presentations, the spelling and writing will be evaluated.

Course methodology

Fuzzy logic	16%
Neural networks	16%
Genetic algorithms	16%
Clustering	16%
Approach Problem and objectives	0%
Avance 1 – Approach Problem - Background	5%
Avance 2 – Techniques selection	0%
Avance 3 – Solution desing and initial implementation	5%
Avance 4 - Implemented solution	6%
Course Project Delivery (presentation and article)	20%

HUGE

Done.