HUGE

Hello

Machine Learning Intro

April 9,2017

Agenda.

1. Definitions
2. Problems
3. Approaches
4. Techniques
5. Libraries

Who am I?

William E. Gómez

Web Engineer at HUGE

PhD. Student at Universidad de Antioquia

Adjunct professor at Universidad de Antioquia

2

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.

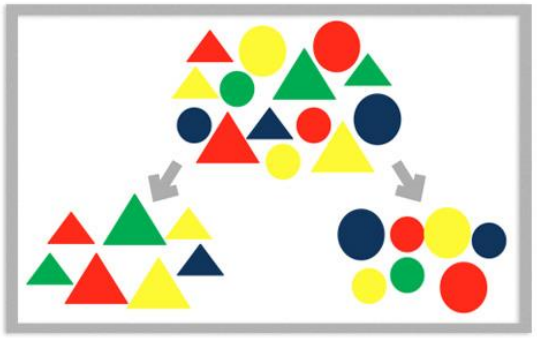
Machine Learning

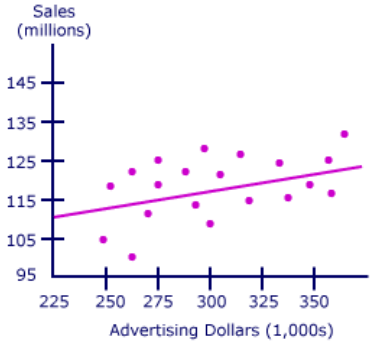
To make the machine learn by itself to solve the problems using a large quantity of data.

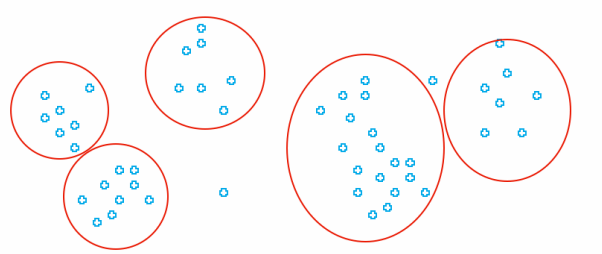
3

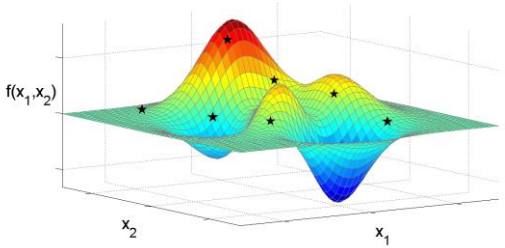
Problems

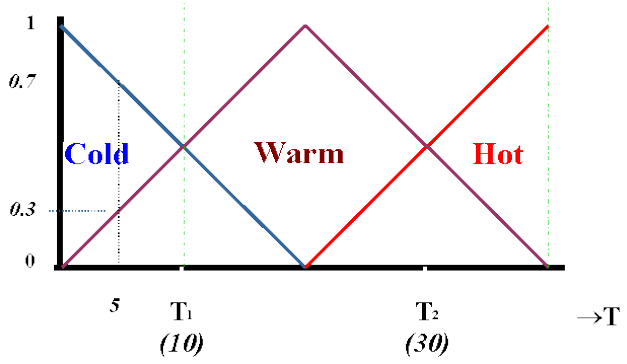
Problems that we can find:

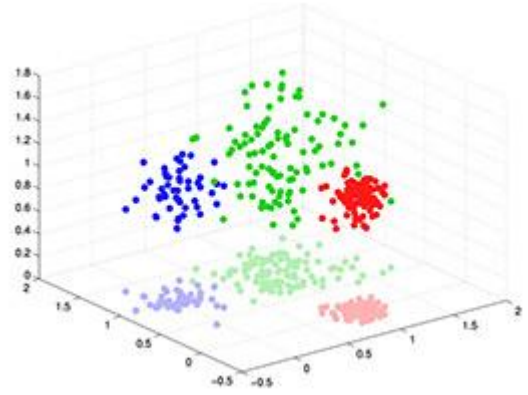












4

Approaches

Supervised

There is an expert knowledge that is desired to reproduced.

Labels

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

$$\min\left(\begin{bmatrix}y_1^*\\ y_2^*\\ \vdots\\ y_k^*\end{bmatrix}-\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_k\end{bmatrix}\right)$$

Prediction

Unsupervised

No apriori knowledge.

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

~~Labels~~

Prediction

5

Techniques

List of topics

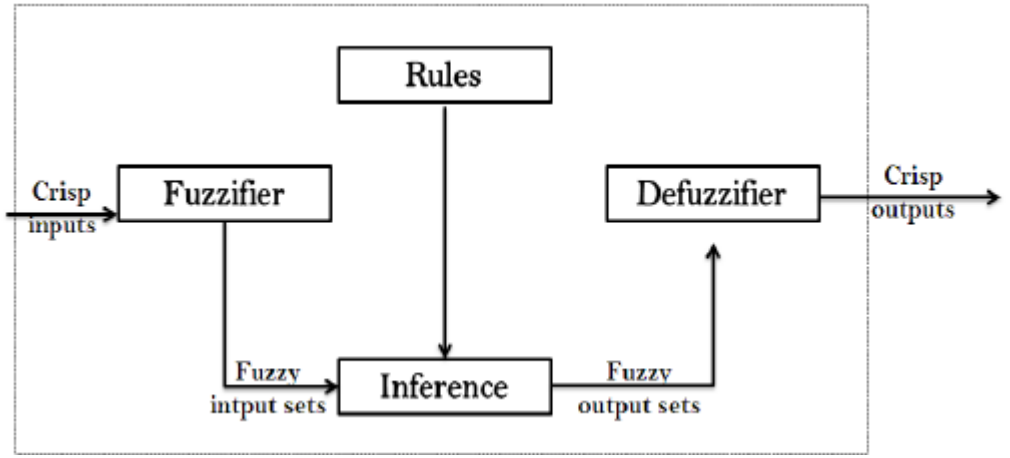
1. Fuzzy logic.
2. Regression.
3. Classification.
4. Evolutionary algorithms.
5. Clustering.

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.

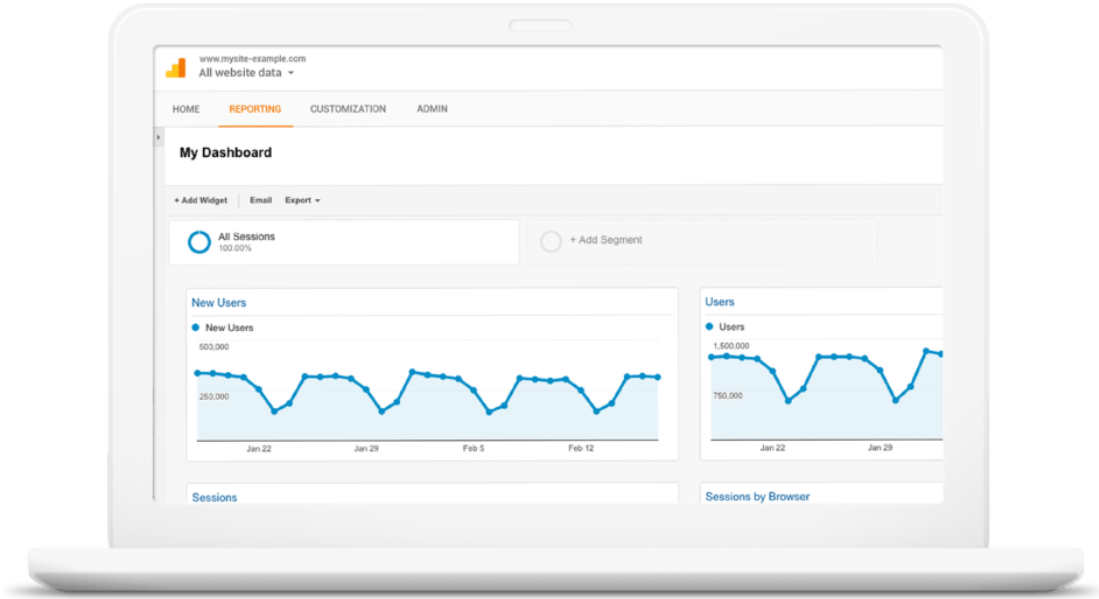
Applications video:

- https://www.youtube.com/watch?v=2d_7GqolNJg

Regression

To find relation between to or more variables.

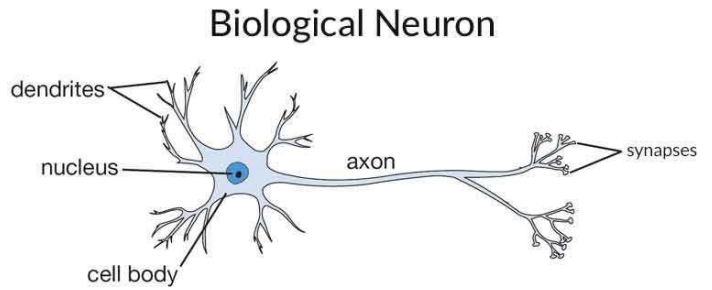
- Forecasting future opportunities.
- Predicting house costs.
- Predicting forest fires.
- Estimate web traffic.

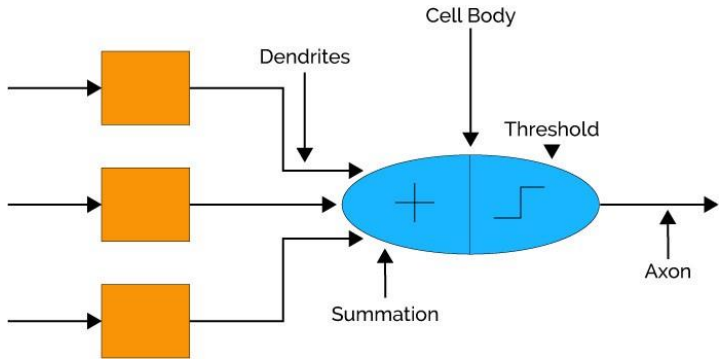


Classification

From a set of features the algorithms discriminate between classes.

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





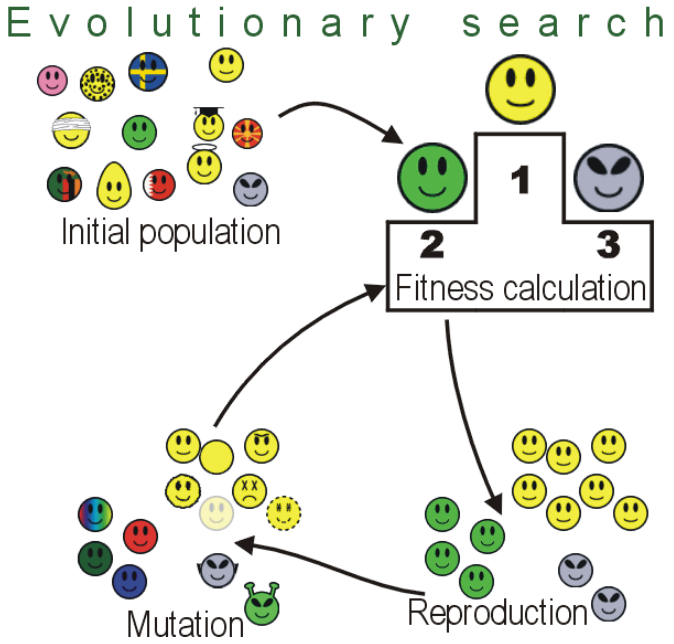
Applications video:

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

H

Evolutionary algorithms

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



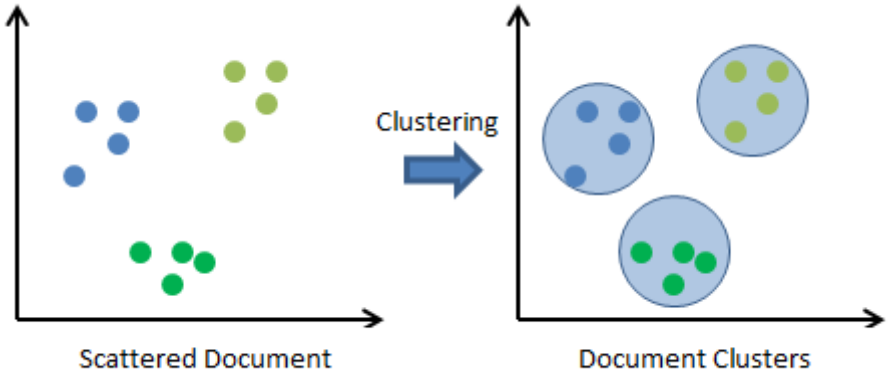
- Automatic design
- Robotics
- Optimization (Connection routing, traffic)
- Computer games
- Strategies

Applications video:

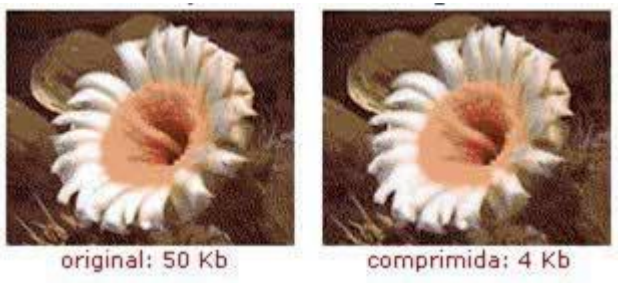
- https://www.youtube.com/watch?v=yQTurXpXd1M

Clustering

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



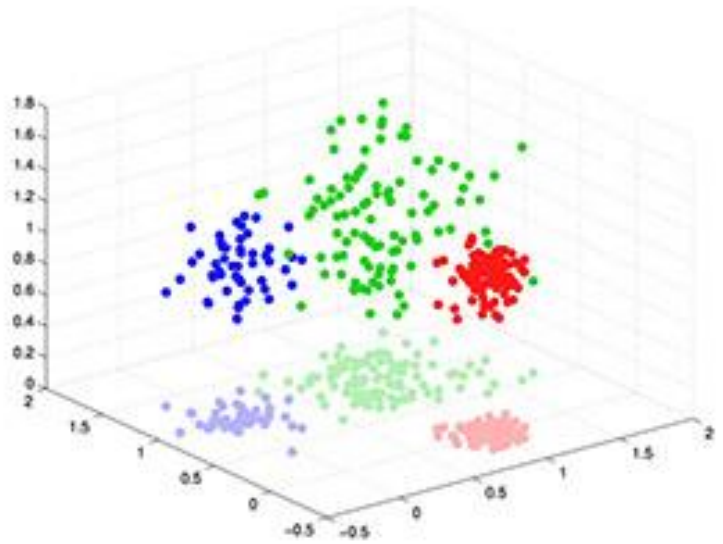
- Search results
- Customer segmentation
- Discovering patterns
- Compression of information



Dimensionality reduction

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

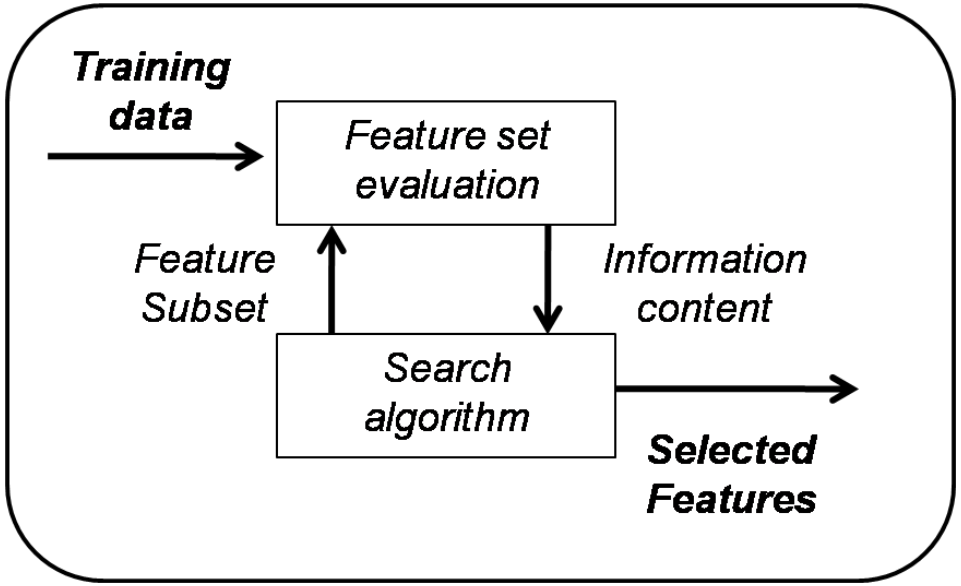
- Creating indicators.
- Visualize the information.



Feature selection

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

- Improve performance of the algorithms.
- Hypothesis generation.



6

Libraries

Javascript libraries

Machine learning tools in JavaScript

https://github.com/mljs/ml

Features:

Clustering, Dimensionality reduction, Neural networks, Regression, Matrix operations, Mathematics

Tensorflow JS

https://js.tensorflow.org/

Features:

Neural networks, Deep learning, Mathematics

BrainJS

http://brain.js.org

Features:

Async Training, Network serialize with JSON, Feedforward Neural Network with backpropagation

Python libraries

SciKit-Learn

http://scikit-learn.org

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