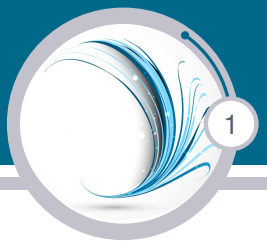
Universidade Federal do Rio de Janeiro
Palestra NCG013 “Priscila Lima”

Weightless Neural Network for Building’s energy consumption classification

Guillaume Jeusel
jeusel.guillaume@gmail.com

July 10, 2017

Content



Introduction
Dataset presentation

- Dataset from UCI :

Data Set Characteristics:	Multivariate	**Number of Instances:**	768	**Area:**	Computer
Attribute Characteristics:	Integer, Real	**Number of Attributes:**	8	**Date Donated**	2012-11-30
Associated Tasks:	Classification, Regression	**Missing Values?**	N/A	**Number of Web Hits:**	95751

Dataset Summary

- Those registers are obtained by simulating 12 different building shapes with Ecotect .
- Using building energy simulation software may provide reliable solutions to estimate the Energy Load. However this can be very time-consuming and requires user-expertise.

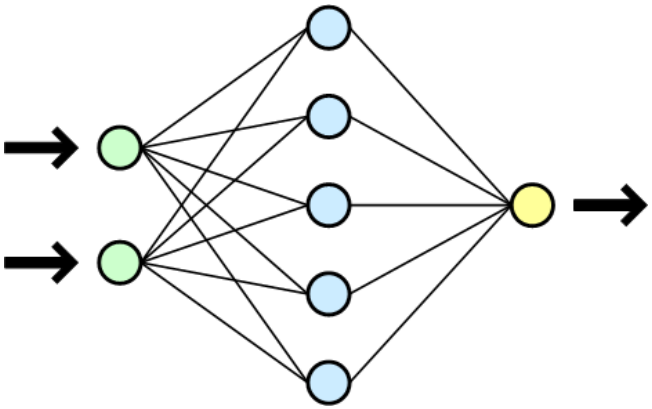
Introduction
Dataset presentation

- Hence, in practice many researchers rely on machine learning tools to study the effect of various building parameters because this is easier and faster.
- energy-efficiency.csv file contains :

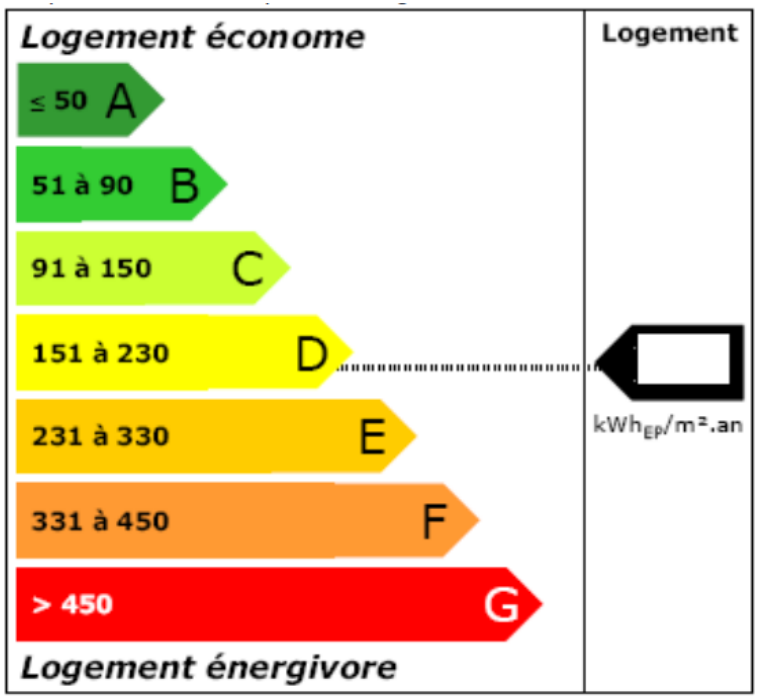
Var.	Var. meaning	N possible values	Unit
X1	Relative Compactness	12	None
X2	Surface Area	12	m^2
X3	Wall Area	7	m^2
X4	Roof Area	4	m^2
X5	Overall Height	2	m
X6	Orientation	4	Unknown
X7	Glazing Area	4	m^2
X8	Glazing Area Distribution	6	None
y	Total Load	636	Unknown

Introduction
Purpose

- **Purpose** : using machine learning WiSARD's algorithm to predict building energy classification according to their geometry.



Entries: variables converted into binaries
Output: Class of building



Classes of building

- **Different steps** :

X1	X2	X3	X4	X5	X7	X8	y
0.98	514.5	294.0	110.25	7.0	0.0	0.0	36.88
0.86	588.0	294.0	147.0	7.0	0.1	2.0	53.69
0.79	637.0	343.0	147.0	7.0	0.1	4.0	71.11

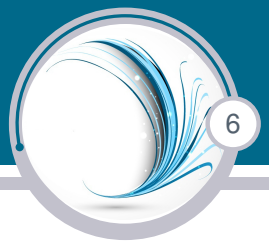
DataFrame after csv read

X1	X2	X3	X4	X5	X7	X8	y
0.98	514.5	294.0	110.25	7.0	0.0	0	B
0.86	588.0	294.0	147.0	7.0	0.1	2	C
0.79	637.0	343.0	147.0	7.0	0.1	4	C

DataFrame for classification problem

X1	y
['0', '0', '1', '1', '1', '1', '1', '1', '0', '1', '1', '1', '1', '0', '1', '0', '1', '1', '1', '0', '0', '0', '0', '1', '0', '1', '0', '0', '1', '0', '0', '0']	B
['0', '0', '1', '1', '1', '1', '1', '1', '0', '1', '0', '1', '1', '1', '0', '0', '0', '0', '1', '0', '1', '0', '0', '0', '1', '1', '1', '1', '0', '1', '1', '0']	C
['0', '0', '1', '1', '1', '1', '1', '1', '0', '1', '0', '0', '1', '0', '1', '0', '0', '0', '1', '1', '1', '1', '0', '1', '0', '1', '1', '1', '0', '0', '0', '1']	C

DataFrame binarized for classification problem



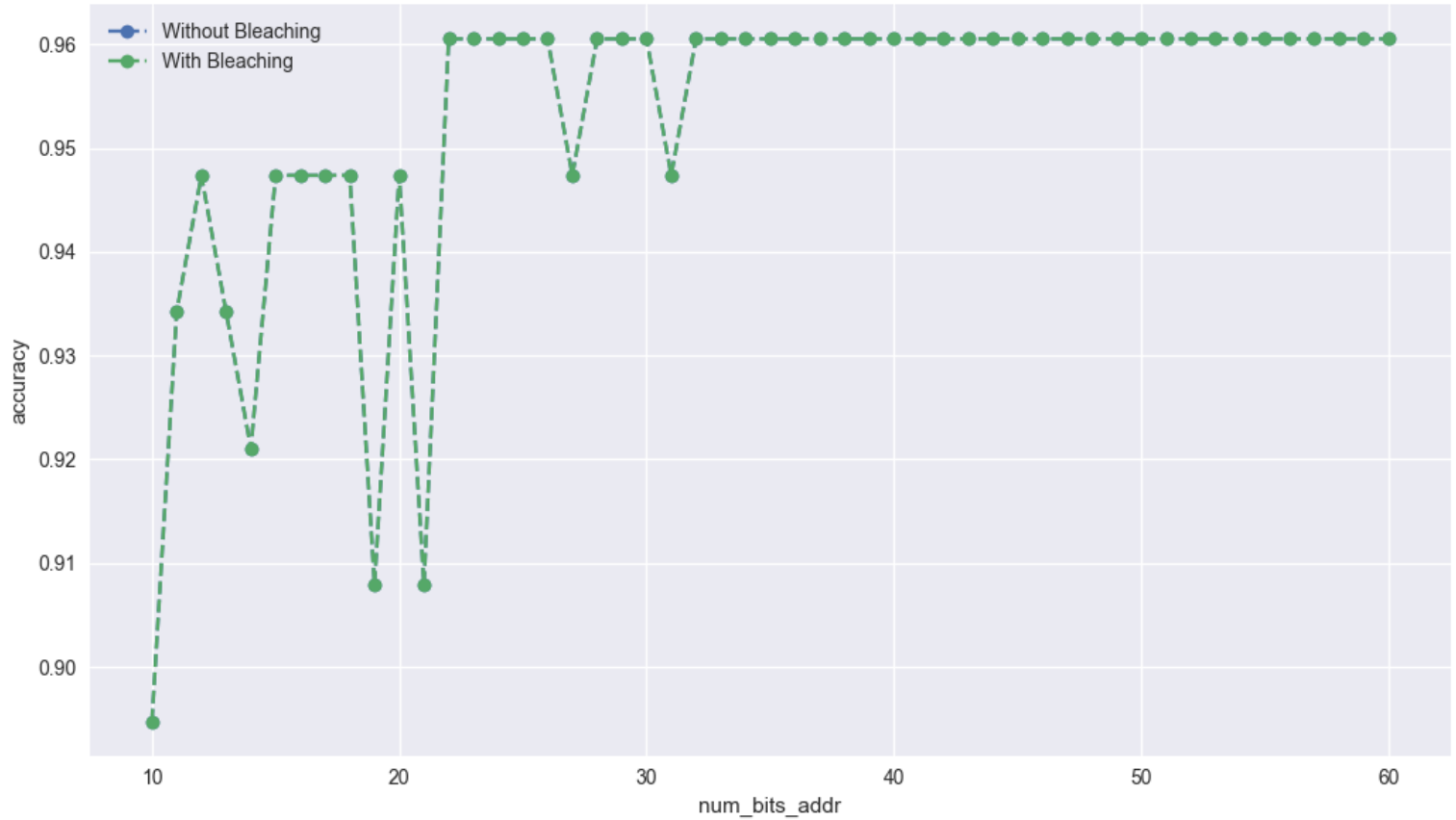
Due to a lack of registers in the dataset, the use of cross-validation with randomly choosen train and test sets where required for validation sake.

Cross-Validation parameters

- For WiSARD parameters calibration : only 20 folds with 90% for the training set and 10% for the test set.
- For **overall** scores : 40 folds with 90% for the training set and 10% for the test set.
- The **accuracy** is then computed as the mean of all accuracies obtained over the folds.

Classification Results

WiSARD parameters calibration



num_bits_addr, bleaching

Classification Results

Overall scores

Classifier	Characteristics	accuracy
GaussianProcess		0.878289473684
Support_Vector_Machine		0.836513157895
NearestNeighbors	n_neighbors=4	0.903618421053
MultiLayerPerceptron	hidden_layer_sizes : 100	0.688486842105
WiSARD	num_bits_addr = 50	0.959868421053

overall

Thank you !