Artificial evolution on EEG signals for the classification of psychophysiological states

Pierrick Legrand

Inria, ASTRAL Team
IMB, Institut de Mathématiques de Bordeaux, UMR CNRS 5251
Université de Bordeaux

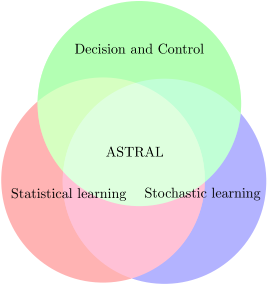
March 2023, 28

ASTRAL TEAM, Inria and Naval Group joint team
Advanced **S**tatis**T**ical infe**R**ence **A**nd contro**L**

Scientific objectives:
Development of advanced statistical and probabilistic models for the analysis and control of complex systems.

Research activities
Our project is focused on the classical triptych:

- Statistical/stochastic modeling
- Estimation/calibration
- Control/decision



Preliminary comment: Artificial Intelligence (AI): ML, NN, AE...

There is a frequent confusion in the public debate between "artificial intelligence" (AI), machine learning (ML), neural network (NN) and deep learning (DL).
However, these notions are not equivalent, but are intertwined.
And among the techniques contained in the broad spectrum of AI, we also find **artificial evolution**.

Preliminary comment: Artificial Intelligence (AI): ML, NN, AE...

There is a frequent confusion in the public debate between "artificial intelligence" (AI), machine learning (ML), neural network (NN) and deep learning (DL).
However, these notions are not equivalent, but are intertwined.
And among the techniques contained in the broad spectrum of AI, we also find **artificial evolution**.

- **Artificial intelligence**: this is a field of research that groups together all the techniques and methods that tend to understand and reproduce the functioning of an intelligence.

Preliminary comment: Artificial Intelligence (AI): ML, NN, AE...

There is a frequent confusion in the public debate between "artificial intelligence" (AI), machine learning (ML), neural network (NN) and deep learning (DL).
However, these notions are not equivalent, but are intertwined.
And among the techniques contained in the broad spectrum of AI, we also find **artificial evolution**.

- **Artificial intelligence**: this is a field of research that groups together all the techniques and methods that tend to understand and reproduce the functioning of an intelligence.
- **Machine Learning**: this is a set of techniques that give machines the ability to automatically learn a set of rules from data. Unlike programming, which consists of executing predetermined rules.

Preliminary comment: Artificial Intelligence (AI): ML, NN, AE...

There is a frequent confusion in the public debate between "artificial intelligence" (AI), machine learning (ML), neural network (NN) and deep learning (DL).
However, these notions are not equivalent, but are intertwined.
And among the techniques contained in the broad spectrum of AI, we also find **artificial evolution**.

- **Artificial intelligence**: this is a field of research that groups together all the techniques and methods that tend to understand and reproduce the functioning of an intelligence.
- **Machine Learning**: this is a set of techniques that give machines the ability to automatically learn a set of rules from data. Unlike programming, which consists of executing predetermined rules.
- **Deep Learning** : it is a machine learning technique based on the neural network model: dozens or even hundreds of layers of neurons are stacked to bring more complexity to the establishment of rules.
  Warning: very efficient but requires a large amount of data.

Preliminary comment: Artificial Intelligence (AI): ML, NN, AE...

There is a frequent confusion in the public debate between "artificial intelligence" (AI), machine learning (ML), neural network (NN) and deep learning (DL).
However, these notions are not equivalent, but are intertwined.
And among the techniques contained in the broad spectrum of AI, we also find **artificial evolution**.

- **Artificial intelligence**: this is a field of research that groups together all the techniques and methods that tend to understand and reproduce the functioning of an intelligence.
- **Machine Learning**: this is a set of techniques that give machines the ability to automatically learn a set of rules from data. Unlike programming, which consists of executing predetermined rules.
- **Deep Learning** : it is a machine learning technique based on the neural network model: dozens or even hundreds of layers of neurons are stacked to bring more complexity to the establishment of rules.
  Warning: very efficient but requires a large amount of data.
- **Artificial Evolution** : Stochastic optimization which uses mechanisms inspired by the biological evolution, such as reproduction, mutation, selection and survival of the strongest individuals.

Darwinism

"We can say, by metaphor, that natural selection seeks, at every moment and worldwide, the slightest variations; it repels those that are harmful, it preserves and accumulates those that are useful; she works in silence, imperceptibly, everywhere and always, as soon as the opportunity arises, to improve all organized beings relative to their organic and inorganic living conditions", Darwin, 1859.

Simple mecanisms

1. Variations, macroscopic and microscopic, within species.
2. Fight for survival.
3. Natural selection: triumph of the lineage that has a useful variation in its environment.

Artificial Darwinism / Evolutionary Algorithms / Artificial Evolution

Artificial Darwinism

Stochastic optimization which uses mechanisms inspired by the biological evolution, such as reproduction, mutation, selection and survival of the strongest individuals

A set of techniques grouped under a generic term

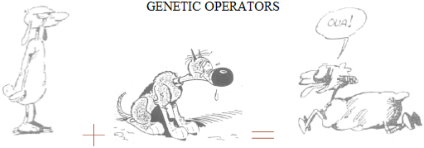
Evolutionary Algorithms	Genetic Algorithms (GA)
	Evolution Strategies (ES)
	Genetic Programming (GP)
	...

Ingredients

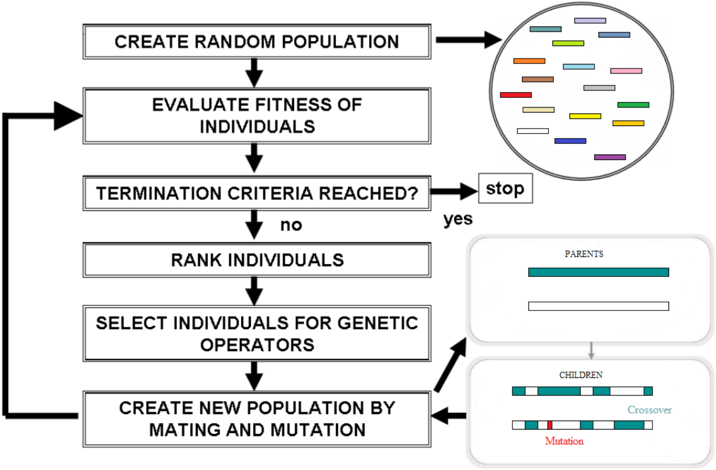
POPULATION

SELECTION

GENETIC OPERATORS

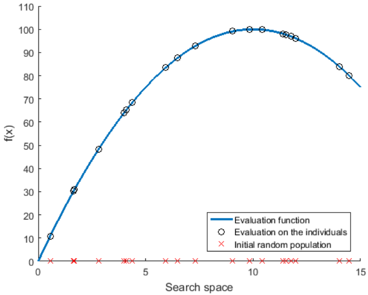


Evolutionary loop

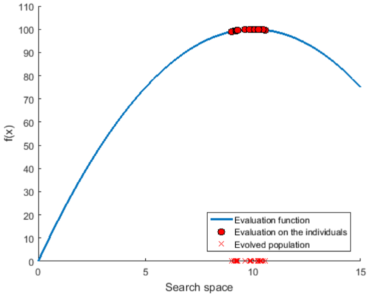


Optimize an "adaptation to the environment"

Potential Solutions = Individuals in a population



Initial random population

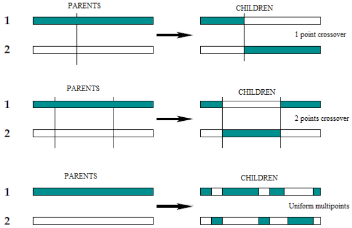


Evolved population

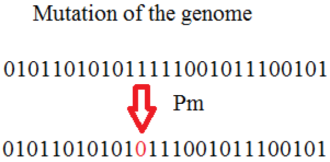
Discrete representation: Genetic Algorithms

Each individual is represented by a binary string.

John H. Holland (1960, 1975), David Goldberg (1989)



Crossover



Mutation

Continuous representation: Evolution Strategies

Each individual is a vector in R^n.

Hans-Paul Schwefel (1970)

Barycentric crossover

$\forall i \in \{1,..,n\}, x_i^{children} = \alpha x_i^{father} + (1-\alpha) x_i^{mother}$

α random value in $[-\epsilon, 1+\epsilon]$.

Gaussian mutation

$\forall i \in \{1,..,n\}, x_i^{children} = x_i^{children} + N(0,\sigma)$

Two parameters P_m and σ.

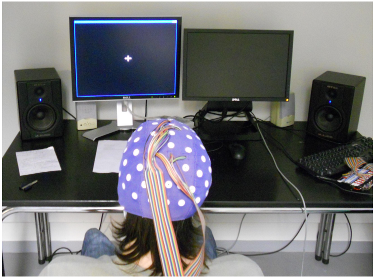
Application: Evolutionary computation for EEG classification

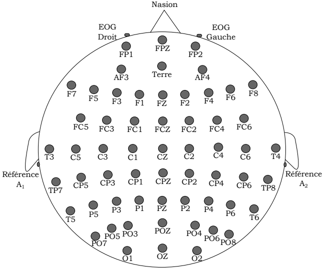
Goal

Characterize the state of alertness of a person from his electroencephalogram (EEG).

Constraints

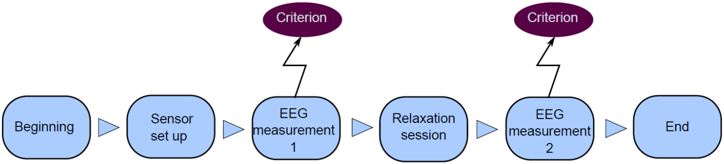
- Space-saving and easy to install system.
- Obtaining the results in real time.





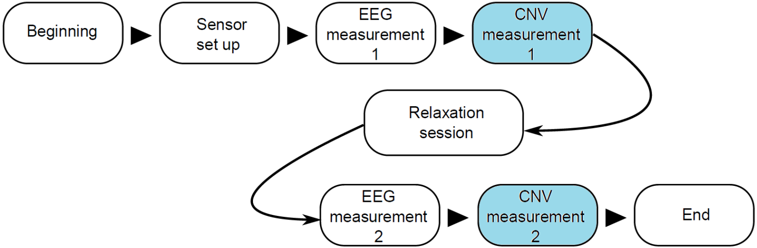
- EEG headset installation time: 45 minuts.
- Subject with open eyes.
- Sampling frequency: 256Hz.
- Recording time: 3 minuts (46000 sample points).

- First EEG recording: subject in a normal state of alertness: "normal"
- Second EEG recording: subject in a state of low vigilance: "relax"



⇒ **Is the relaxation session effective?**

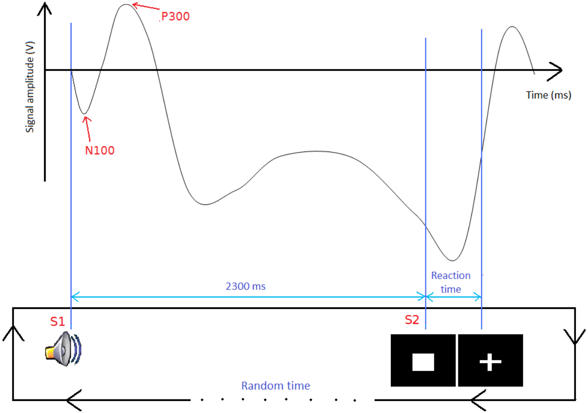
Verification of data quality



Two steps added to perform a test VCN (Variation Contingence Negative)

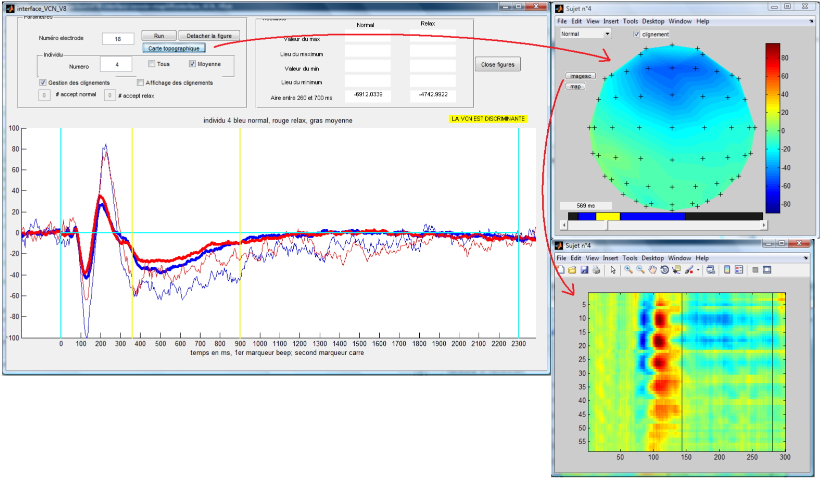
Verification of data quality - CNV measurement

50 occurrences



Acquisition Protocole

Verification of data quality - CNV graphical interface



Campaigns:

- 58 electrodes renumbered from 1 to 58
- Subjects under 35, right-handed and non-smoker
- 58 subjects $\Rightarrow$ 16 preserved

Relaxation session

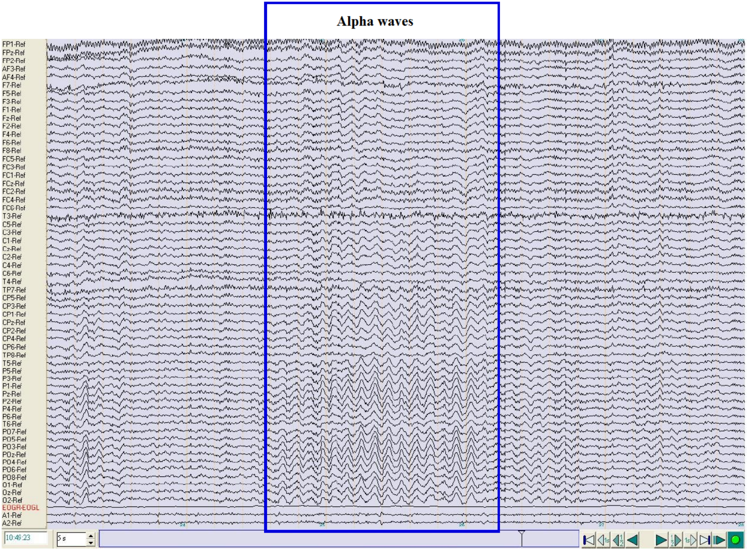
20 minutes with a recorded voice offering 3 exercises:

- Autogenic training [Schultz1958]: repetition of sentences, self-hypnosis.
- Progressive muscle relaxation [Jacobson1974].
- Mental visualization (familiar places, smells, noises).

3 minutes of EEG recording before relaxation.

Acquisition Protocole

3 minutes of EEG recording after relaxation.



Wavelets and energy

A discrete wavelet transform is used in order to calculate the energy of the signal.

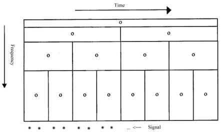
Wavelet coefficient

$C_{j,k}$ denotes the wavelet coefficient at scale j and time k.

$$C_{j,k} = 2^{-\frac{j}{2}} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t-k)dt$$

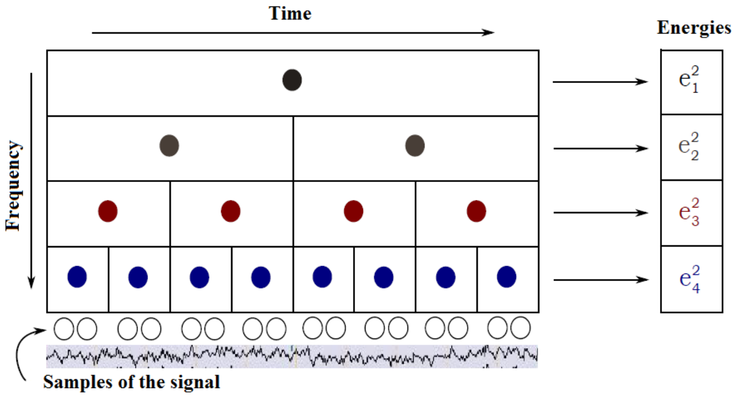
Energy at scale j

$$E_j^2 = \sum_{k=1}^{2^{j-1}} [C_{j,k}]^2$$



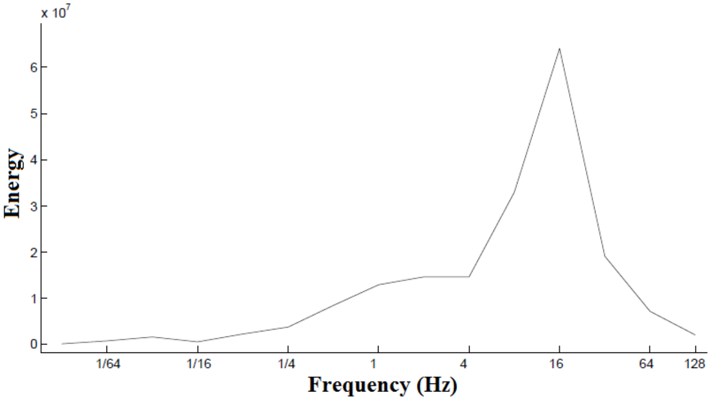
Wavelets and energy

The dyadic grid gives a spatio-frequential representation of the discrete dyadic wavelet decomposition



Feature Extraction: Slope Criterion

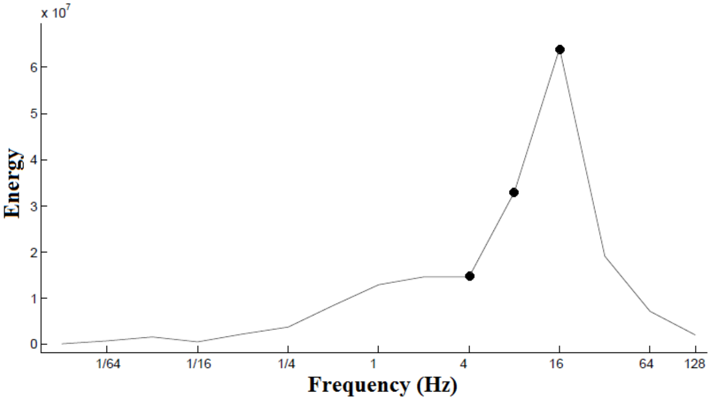
- Alpha: $8 - 12Hz$.
- Waves characteristics of a relaxed state.



Linear regression between 4 and $16Hz$.

Feature Extraction: Slope Criterion

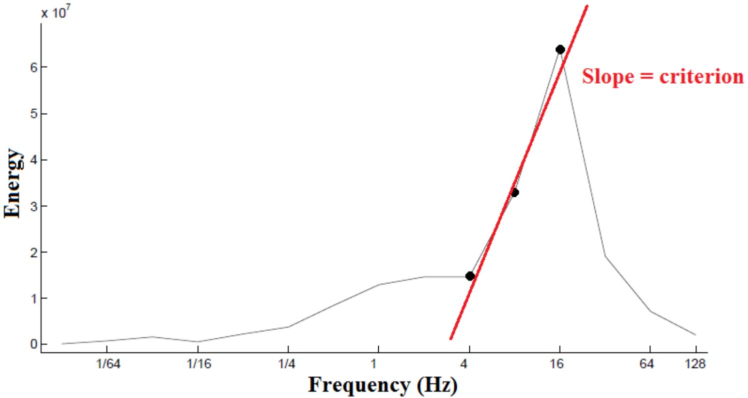
- Alpha: $8 - 12Hz$.
- Waves characteristics of a relaxed state.



Linear regression between 4 and $16Hz$.

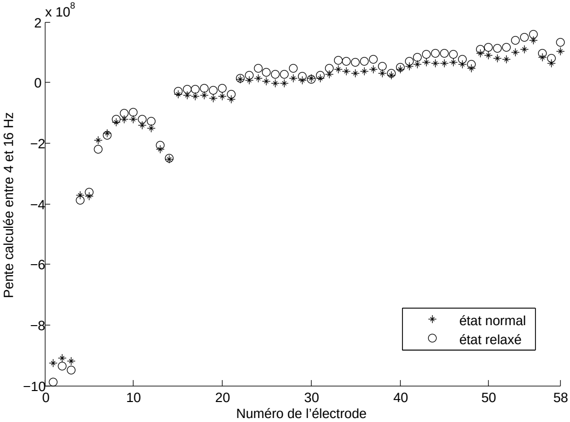
Feature Extraction: Slope Criterion

- Alpha: $8 - 12Hz$.
- Waves characteristics of a relaxed state.



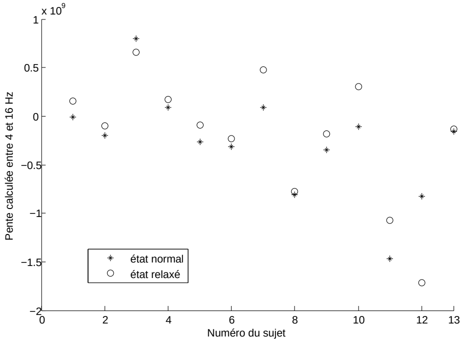
Linear regression between 4 and $16Hz$.

Feature Extraction: Slope Criterion



Slope criterion, sum on subjects for each electrode

Feature Extraction: Slope Criterion



Slope criterion, sum on electrodes for each subjects

⇒ Very strong inter-subject variability

⇒ This criterion does not allow to build a powerful classifier for different subjects.

Classification: Usual methods

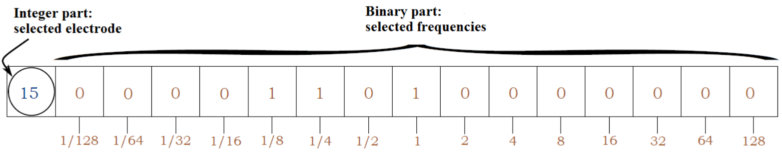
	K nearest neighbors	Binary decision trees	Random forests	Discriminant PLS	Sparse Discriminant PLS
Mean	37.28	33.98	32.03	40.63	36.25
Standard Deviation	10.47	5.15	6.46	8.55	7.96

Mean and standard deviations of Correct Classification Rates for different classification methods applied on slope criterion.

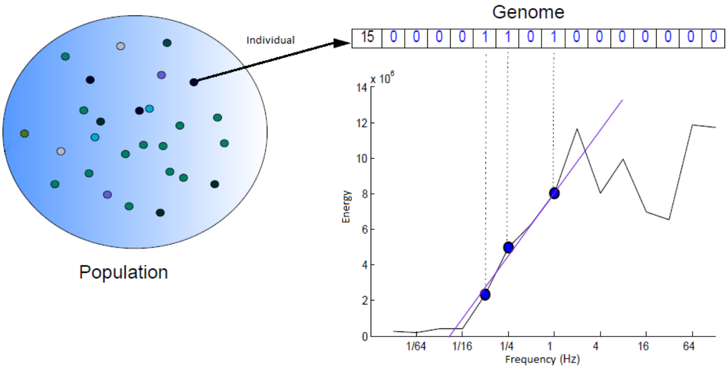
This approach is not efficient

Our contribution: Design a relevant **evolutionary algorithm** to solve this task of classification.
$\Rightarrow$ **Find the relevant electrodes.**
$\Rightarrow$ **Find the relevant frequencies for the calculation of the slope criterion.**

Example of a genome in the evolutionary algorithm

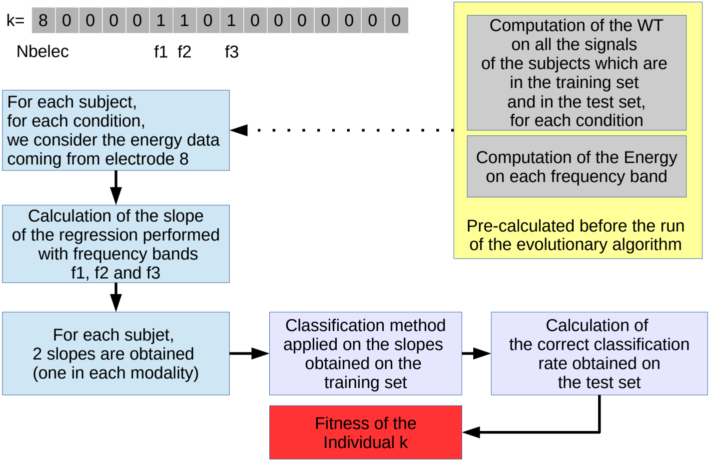


Relationship between the genome and the calculation of the slope criterion



Evaluation of a individual in the evolutionary algorithm

Evaluation of the individual k



Average correct classification rate

Evaluation Method	CCR	
	Mean	Standard deviation
CART	86.68	1.87
SVC	83.49	2.37

Average and standard deviations of the correct classification rates obtained for the 100 runs of the evolutionary algorithm and for two methods of evaluation.

Best genome

Evaluation method	BEST genome		
	Selected electrode	Selected frequency (Hz)	Correct classification rates
CART	F4	$1/8$, $1/4$, 2, 4 et 64	$89,33\%$
SVC	F2	$1/32$, $1/16$, 2, 4, 8, 64 et 128	$89,33\%$

Table summarizing the two best genomes found during the 100 runs of the genetic algorithm with two methods of evaluation.

One electrode selected => Quick installation of the acquisition system.
Results obtained in less than a second => Can be used on real application.

Our contribution

Design of a relevant **evolutionary algorithm** to solve this task of classification.
$\Rightarrow$ **Find the relevant electrodes.**
$\Rightarrow$ **Find the relevant frequencies for the calculation of the slope criterion.**

Application on similar real applications

For instance: sleeping non sleeping, alertness tunneling of a pilot, etc.
What is needed:

- Data on 2 (or more) conditions
- Use of the same protocol

Warning:

This method cannot be used when the subject is making pronounced movements. Indeed, the motor activity overwrites the other information contained in the EEG signals.
$\Rightarrow$ **One can use other physiological data (eye tracking, EDA, ECG, etc.) and try the same protocol.**