A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

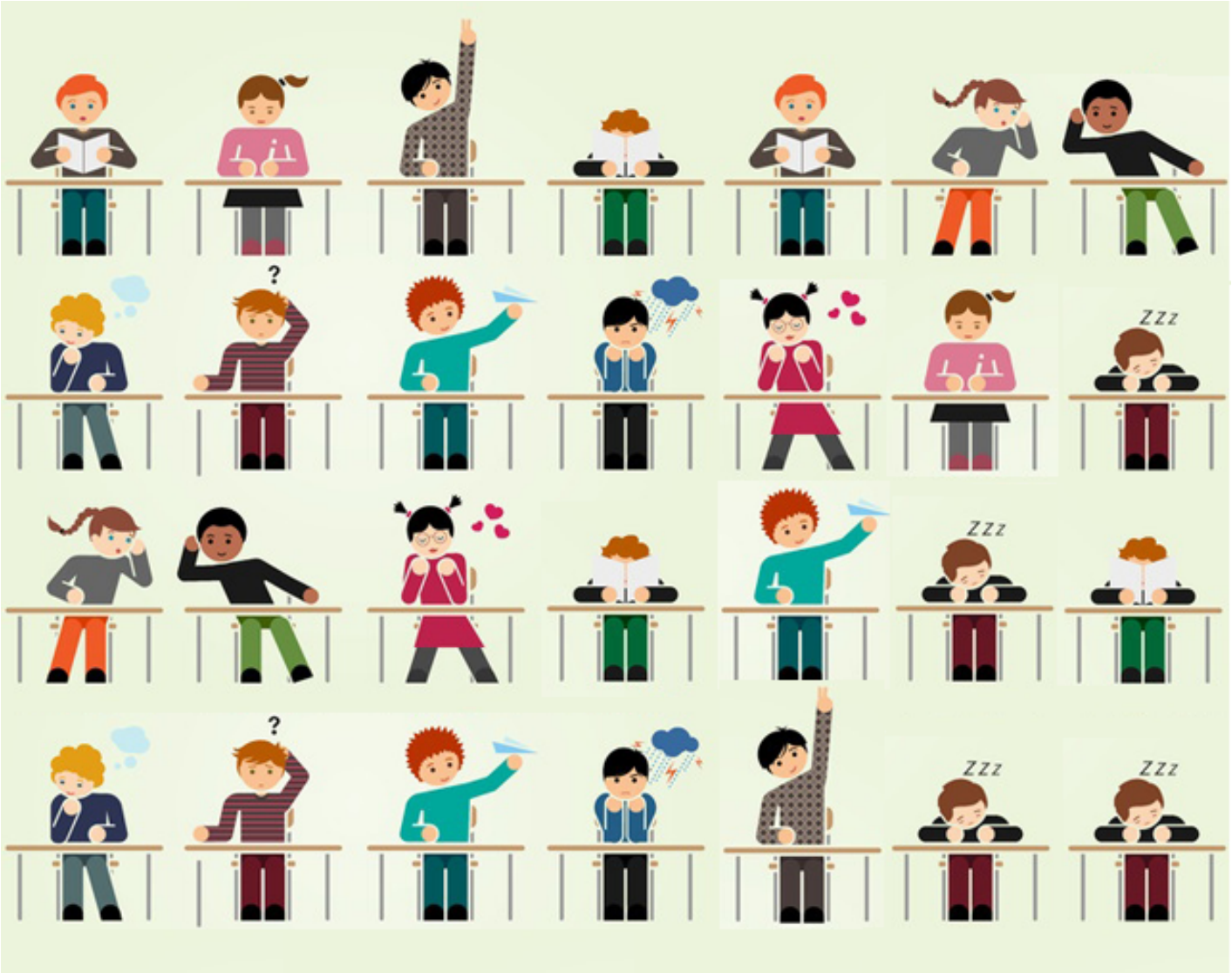
Sarah Itani, Fabian Lecron and Philippe Fortemps

Management de l'Innovation Technologique

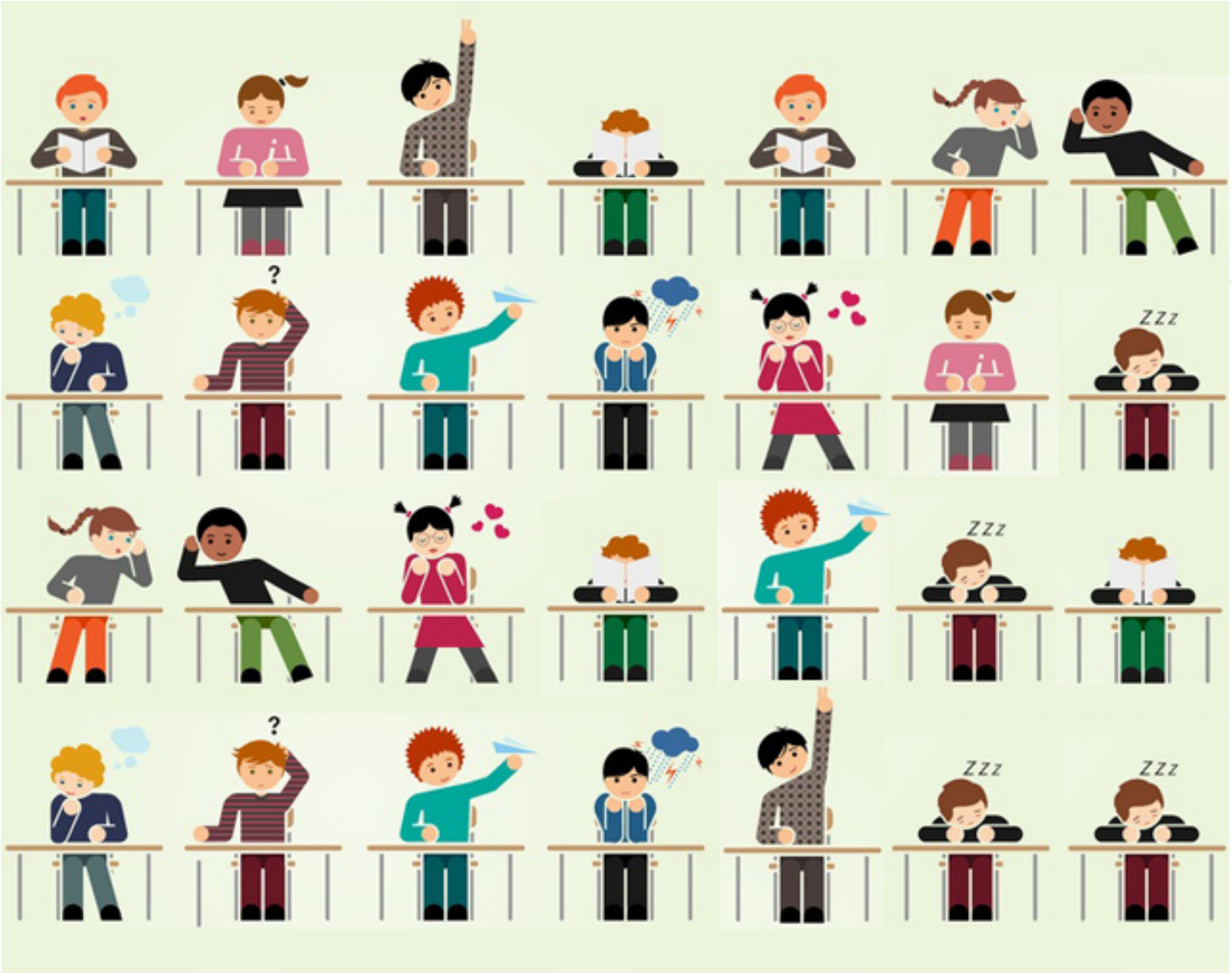
Faculté Polytechnique, Université de Mons

DA2PL'2018

Poznań University of Technology


?
Zzz
Zzz
?
Zzz
Zzz

Attention Deficit Hyperactivity Disorder concerns 5-7% of the children



Attention Deficit Hyperactivity Disorder concerns 5-7% of the children

ADHD is a mental disorder observed in children or adolescents

- problems to pay attention,
- excessive activity, or
- difficulties to control his/her behavior

in comparison of his/her age.

Without a consensus of the physiological bases of the trouble

- diagnosis is mainly based on parents'report
- diagnosis is often dependent on the physician ($\kappa = 61\%$)
- the risk of false positive is rather high

One needs better and earlier diagnosis, but also meaningful insights on the disorder

To avoid the risk of a wrong (and masking) medication,
the diagnosis is often postponed up to the age of 7 to 12 years.

For both the parents and the children,
this is a painful delay, since there are impacts on

- emotions
- relationships
- academic results

The diagnosis aiding tool has to be

- objective, relying on physiological indicators
- efficient, providing confident answers
- interpretable, able to give understanding

ADHD-200, a Data Mining competition, launched in 2012, with a performance mindset

The challenge

- Reach the highest prediction rate for ADHD diagnosis
- About 1000 patients, from different hospitals
- The most challenging site was NYU, with 210+41 patients
- The best prediction rate was 61% on the test set (but only 37% on NYU)

The data

- Phenotype: age, gender, handedness, IQ
- Magnetic Resonance Images (MRI): resting state brain functional activity
- fMRI signals computed on a brain atlas of 116 regions of interest (ROI)

Can MCDA tools bring new insights in such Data Mining context ?

We are looking for a model

- efficient (prediction rate)
- compact (number of ROI)
- meaningful (readability)

able to cope with ADHD-200 dataset.

We'll focus on NYU sample.

A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

The MR-Sort model makes sense for medical diagnosis

MR-Sort, a simplified version of the Electre TRI procedure [Yu, 1992]

For sorting alternatives evaluated on m criteria
to p ordered classes C^h for $h = 1, \ldots, p$, one needs

- A set of separating profiles of performances b^h for $h = 1, \ldots, p-1$
- m criteria weights w_j for $j = 1, \ldots, m$
- A majority threshold λ

An alternative is assigned above the highest profile it outranks

$$a \in C^h \iff \sum_{j : a_j \geq b_j^{h-1}} w_j \geq \lambda \text{ and } \sum_{j : a_j \geq b_j^h} w_j < \lambda$$

The MR-Sort model makes sense for medical diagnosis

MR-Sort, a simplified version of the Electre TRI procedure [Yu, 1992]

For sorting alternatives evaluated on m criteria
to p ordered classes C^h for $h = 1, \ldots, p$, one needs

- A set of separating profiles of performances b^h for $h = 1, \ldots, p-1$
- m criteria weights w_j for $j = 1, \ldots, m$
- A majority threshold λ

An alternative is assigned above the highest profile it outranks

$$a \in C^h \iff \sum_{j : a_j \geq b_j^{h-1}} w_j \geq \lambda \text{ and } \sum_{j : a_j \geq b_j^h} w_j < \lambda$$

This kind of assignment rule is usual in medical diagnosis:

if you enjoy sufficient relevant symptoms,
than you can be diagnosed with a given disease...

The MR-Sort model is quite easy to learn from data

Learning procedures exist for a MR-Sort model

Among others,

- Linear Program or Mixed Integer Program [Leroy et al, 2011], to learn the best weights and threshold for given profiles
- Metaheuristic [Sobrie et al, 2013], to learn good profiles
- SAT approach [Belahcene et al, 2018], to learn completely such a kind of models

The MR-Sort model is quite easy to learn from data

Learning procedures exist for a MR-Sort model

Among others,

- Linear Program or Mixed Integer Program [Leroy et al, 2011], to learn the best weights and threshold for given profiles
- Metaheuristic [Sobrie et al, 2013], to learn good profiles
- SAT approach [Belahcene et al, 2018], to learn completely such a kind of models

In our case, there are only two classes: healthy or pathological

- C_1: healthy children (TD : typical development)
- C_2: ADHD children

There is only one profile, a set of weights and a majority threshold.

Learning the weights and the majority threshold is easy

A linear program can help

Given a profile b to separate between healthy (TD) and ADHD children:

$$\text{minimize} \sum_{a_i \in A} y_i$$

$$\sum_{j : a_{i,j} \geq b_j} w_j + y_i \geq \lambda \qquad \forall a_i \in A_2(ADHD)$$

$$\sum_{j : a_{i,j} \geq b_j} w_j - y_i \leq \lambda \qquad \forall a_i \in A_1(TD)$$

Minimize the sum of slack variables, where the slack associated to a child is the difference between the threshold and the coalition in favor of diagnosing the child as ADHD-affected.

Learning the profile(s) is more difficult

Some metaheuristics have been proposed

The main idea is to

- generate several random or adhoc profiles
- (randomly) optimize these profiles locally

Very nice and tricky tools have been proposed by [Sobrie et al., 2013].

A crucial point is to provide a good start

For a known dataset, it may be possible to incorporate knowledge in both the profile generation and its optimization.

$\Longrightarrow$ Domain-inspired Data Mining

A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

The ADHD-200 dataset provides phenotype and rs-fMRI signals over 90 brain areas

The initial dataset consists of

- MRI signals measuring the time course of each brain region of interest
- Brain is parcelled into 116 ROI (atlas AAL)

A first preprocessing leads to

- for each region, compute the (log-)variance of its signal
- infer some information about the intensity of the ROI activity

We have thus 210 training examples and 41 test examples, described by 116 signal variances and 4 phenotype attributes.

ADHD is associated to high signal variances

Local preference

From neuropsychology knowledge

- Phenotype (age, IQ, handedness) shouldn't be useful, except Gender
  Indeed, ADHD is more prevalent in boys than in girls.
  In the dataset, 68% of the boys and 32% of the girls have ADHD.
- Behavioral hyperactivity may be linked to neuronal hyperactivty.
  In other words, high activity in brain should be an indication of ADHD.

We can enjoy from monotonic attributes

- Higher the brain signal variance, higher the possibility to have ADHD
- Being a boy, higher the possibility to have ADHD

All these attributes are positively related to ADHD

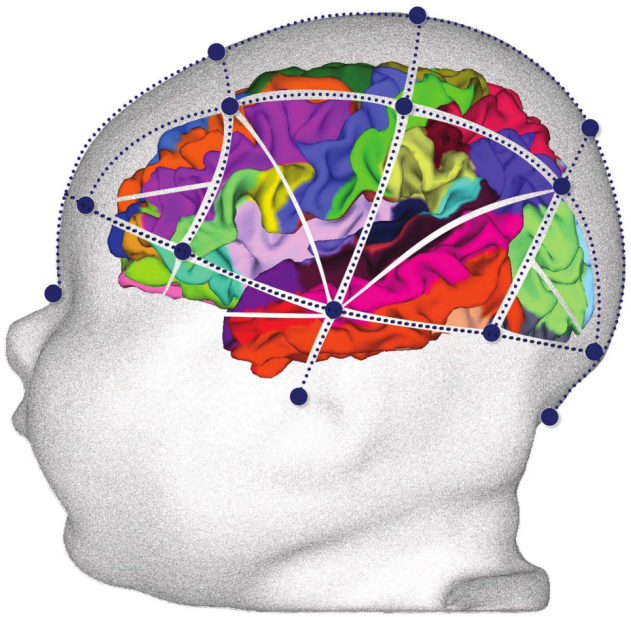
Among the brain areas, the limbic system has proved to be relevant

Several "theories" explain ADHD

No real help to focus on specific parts of the brain

From previous studies [Itani et al, 2018]

- With Gender, the limbic system is sufficient to explain ADHD
- This is related to one of the neuro-psychology "theories"



We managed to reduce the set of considered ROI, to a set of 26 meaningful brain areas (plus Gender)

Our first approach of ADHD by MR-Sort...

The procedure

1. Generate 100 profiles on the 27 attributes,
   - either at random
   - or ad hoc (i.e., between the medians)
2. Locally optimize each profile, with the best weights (LP)
3. Keep the overall best results

The measures of quality

1. Training accuracy (on the training set)
2. Prediction accuracy (on the test set)
3. Number of ROI in the model

The first results are convincing about the approach

Accuracies in comparison with other models

Sample	MR-Sort(1)	C4.5	ADHD-200	Colby 2012
Training	71%	73%	-	
Test	56%	61%	35.2%	37%

The MR-Sort model is compact

Due to the LP, the MR-SORT model contains 11 positive weights.
11 attributes are necessary: Gender and 10 ROI.

A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

A gender-differentiated MR-Sort model is possible on the basis of the weights

A gender-based differentiation should improve the performances

- Gender appears as a main attribute in the previous MR-Sort model
- Boys and Girls do not enjoy the same risk in front of the trouble, neither in the dataset nor in the reality
- The prevalence is different with the Gender, but maybe also the concerned brain areas

A gender-differentiated MR-Sort model is possible on the basis of the weights

A gender-based differentiation should improve the performances

- Gender appears as a main attribute in the previous MR-Sort model
- Boys and Girls do not enjoy the same risk in front of the trouble, neither in the dataset nor in the reality
- The prevalence is different with the Gender, but maybe also the concerned brain areas

The differentiation can be made on the weights

- Two profiles would mean two different protocols for the examination
- A single profile with two weight vectors means a single examination with two different diagnosis mechanisms

Learning the double weights and the majority threshold is as easy as previously

A similar linear program can help

Given a profile b to separate between healthy (TD) and ADHD children:

$$\text{minimize} \sum_{a_i \in A} y_i$$

$$\sum_{j : a_{i,j} \geq b_j} w_j^{Gender} + y_i \geq \lambda \qquad \forall a_i \in A_2(ADHD)$$

$$\sum_{j : a_{i,j} \geq b_j} w_j^{Gender} - y_i \leq \lambda \qquad \forall a_i \in A_1(TD)$$

Where *Gender* is either *Boy* or *Girl*.

The second results are even more interesting

Accuracies in comparison with other models

Sample	MR-Sort(2)	MR-Sort(1)	C4.5	ADHD-200	Colby 2012
Training	75%	71%	73%	-	
Test	61%	56%	61%	35.2%	37%

The MR-Sort model with 2 weight vectors is more compact

Due to the LP, the MR-SORT model contains 7 positive weights.
7 attributes are necessary: Gender(isBoy, isGirl) and 5 ROI.

This efficient model allows also for interpretation

The LP step determines $\lambda = 1$ and the following weight vectors

Attribute	W^{Girl}	W^{Boy}
15	–	0.002
32	0.001	–
40	–	0.002
41	0.001	–
5	0.001	–
isBoy	–	0.998
isGirl	0.997	–

The double weight vectors can be interpreted as

- a boy needs a high variance signal on either zone 15 or zone 40
- a girl needs a high variance signal on the three zones 21, 41 and 5

to be diagnosed as ADHD-affected.

A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

MCDA models are interesting in medical data mining, in a domain-expert guided approach

A priori, we used the domain-expert knowledge

To implement the MR-Sort model learning,

- Monotony in the relevant attributes (signals and gender) with the classes ordering
- Building of a "gender-differentiated MR-Sort", as a single profile, with two different weights

A posteriori, we can enrich the domain-expert knowledge

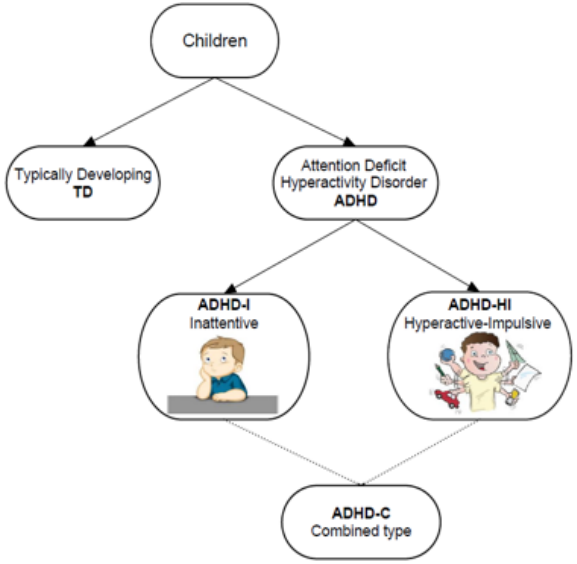
The MR-Sort model allows for interpretation as

- it is simple to read: *being above a profile*
- it is compact (in the numbers of ROI)

Perspectives: to go further on both performances and interpretability

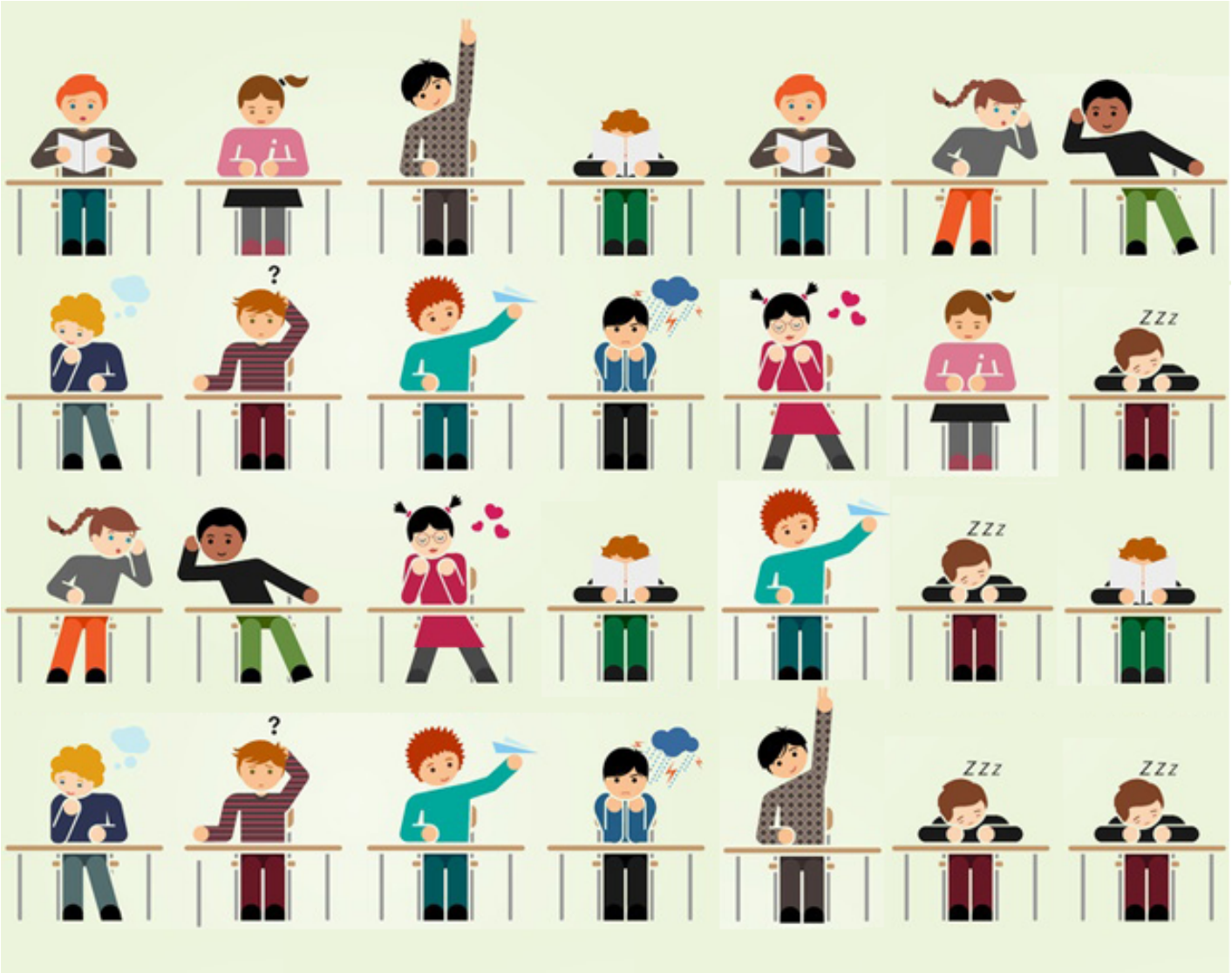
We have study on one site on the presence of ADHD

- We will study all the sites of the dataset
- We will also consider the (not linearly ordered !) levels of ADHD

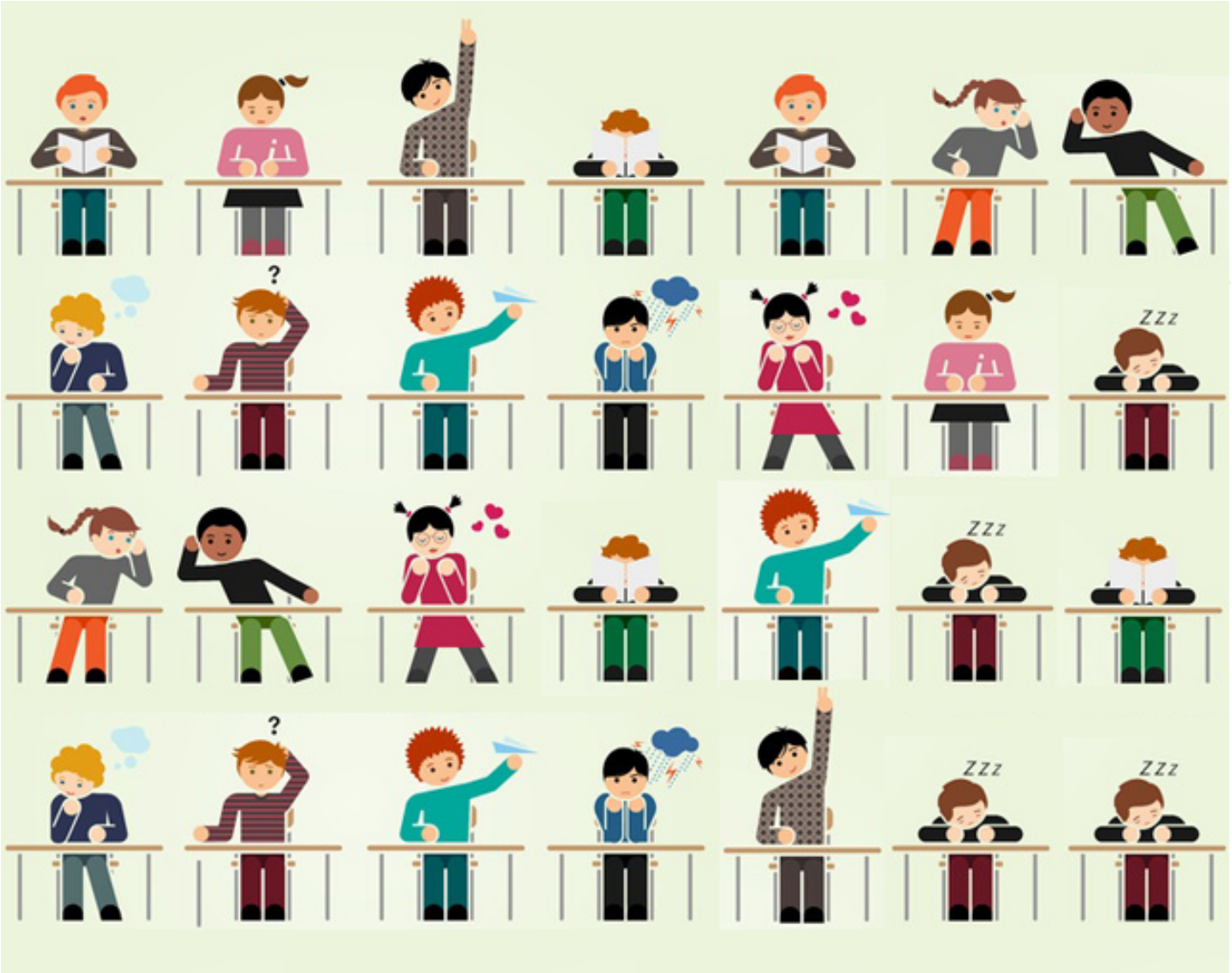


One can look for several separating profiles between the two classes

- We should extend a bagging-like approach,
  - either, by means of a vote between several MR-Sort models
  - or, by incorporating the different profiles in a single MR-Sort model.


?
Zzz
Zzz
?
Zzz
Zzz

Have you ever seen such a class in front of you ?
Don't worry, it is probably the case of each of us !



A gender-differentiated MR-Sort model for diagnosis aid of Attention Deficit Hyperactivity Disorder

Sarah Itani, Fabian Lecron and Philippe Fortemps

Management de l'Innovation Technologique

Faculté Polytechnique, Université de Mons

DA2PL'2018

Poznań University of Technology