

# Clustering évidentiel pour l'interprétation de trajectoire patients atteints de douleurs chroniques

V. Antoine

Université Clermont Auvergne, LIMOS, UMR CNRS 6158, France  
<https://perso.isima.fr/~viantoin>

Avril 2026

[1] A. Soubeiga, V. Antoine, A. Corteval, N. Kerckhove, S. Moreno, I. Falih, J. Phalip. *Clustering and Interpretation of time-series trajectories of chronic pain using evidential c-means*, Expert Systems With Application, 2025.



# Plan

- 1 Introduction
- 2 Prétraitements
  - Données eDOL
  - Extraction d'attributs
  - Sélection d'attributs
- 3 clustering
  - FCM
  - ECM
- 4 Résultats
- 5 Conclusion

# La douleur chronique

## Le fardeau de la douleur chronique

- 40% des Français concernés
- 10 millions souffrent des répercussions sur le quotidien
  - Fatigue, troubles du sommeil, dépression, anxiété, handicap, isolement...
- Des traitements anciens, peu efficaces au long cours, avec effets indésirables

# La douleur chronique

## Le fardeau de la douleur chronique

- 40% des Français concernés
- 10 millions souffrent des répercussions sur le quotidien
  - Fatigue, troubles du sommeil, dépression, anxiété, handicap, isolement...
- Des traitements anciens, peu efficaces au long cours, avec effets indésirables

## Difficultés de prise en charge

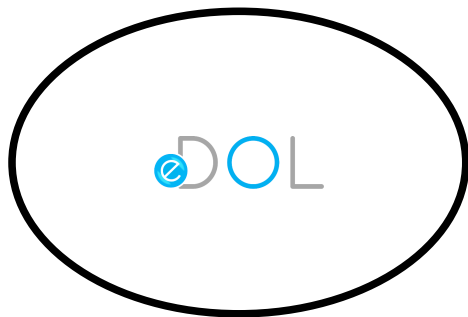
- Manque de connaissances et de temps des médecins traitants
- Surcharge des centres de la douleur
- Insuffisance de prescription des thérapies complémentaires validées
- Errance diagnostique et nomadisme médical

# Fondation Analgesia

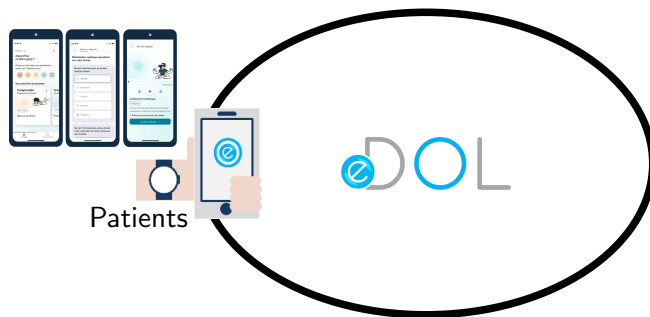
Fondation dans le développement de solutions numériques pour les patients et leurs soignants.



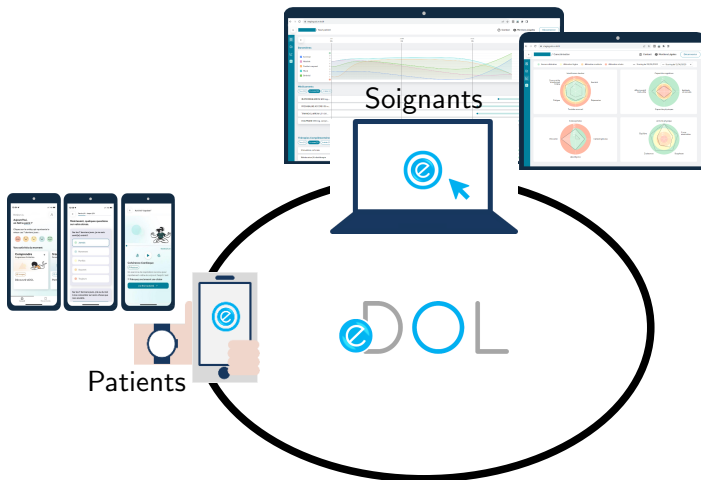
# Projet eDOL



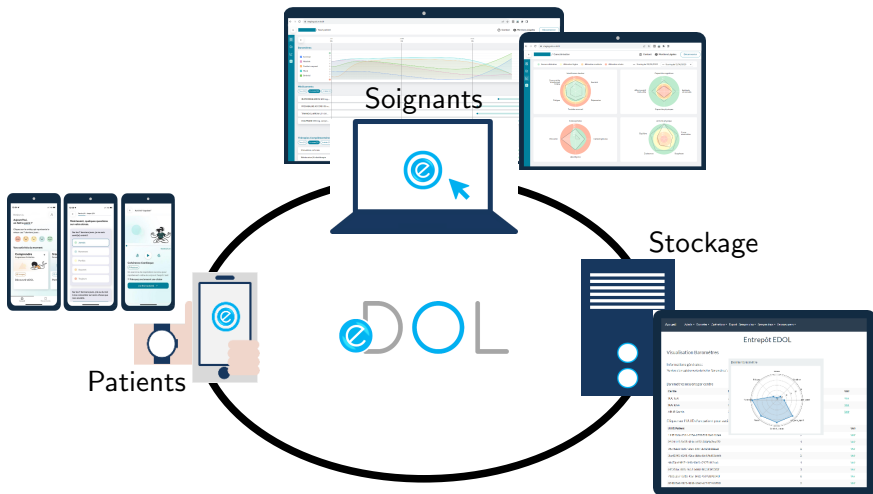
# Projet eDOL



# Projet eDOL



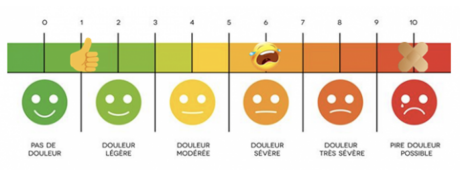
# Projet eDOL



# Objectif eDOL

## Analyse des profils de patients douloureux chroniques

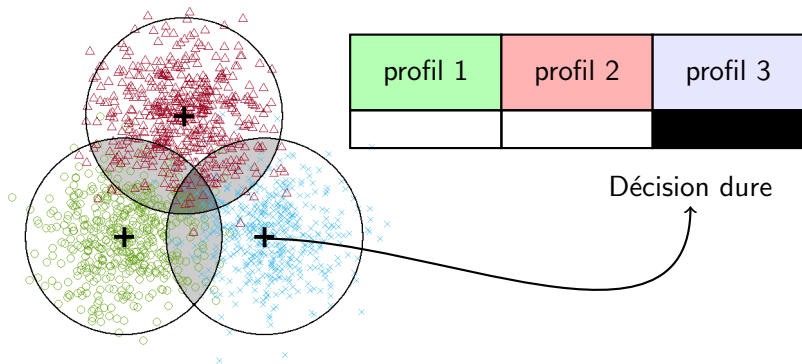
- Mieux évaluer et mieux suivre la douleur (par baromètres)
- Accompagner le généraliste auprès de ses patients
  - alerte
  - diagnostic
  - prise en charge



# Profil des patients ?

## Beaucoup d'incertitude

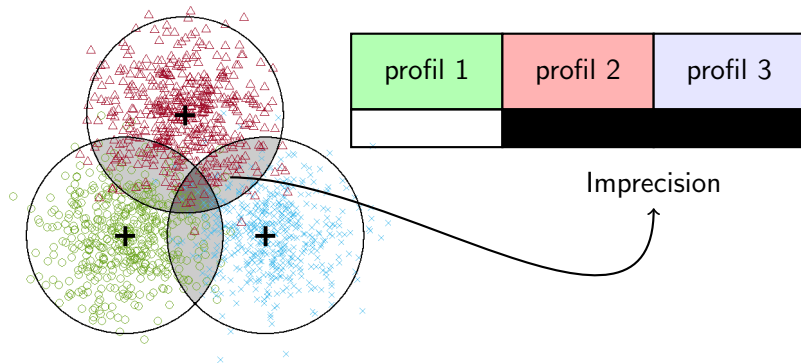
- états intermédiaires
- états atypiques



# Profil des patients ?

## Beaucoup d'incertitude

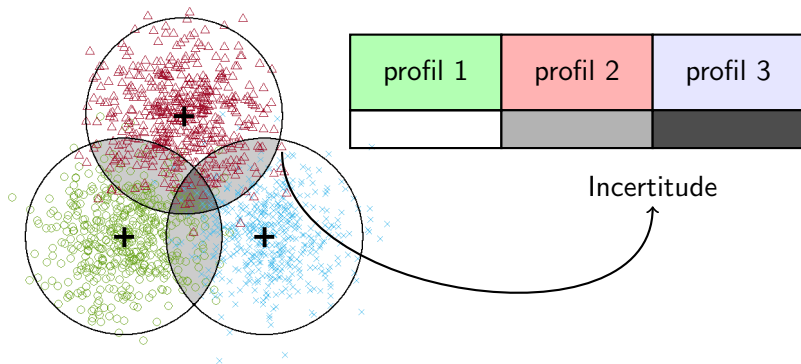
- états intermédiaires
- états atypiques



# Profil des patients ?

## Beaucoup d'incertitude

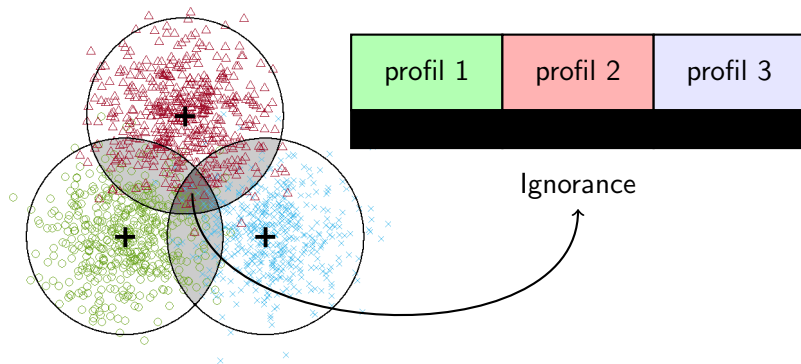
- états intermédiaires
- états atypiques



# Profil des patients ?

## Beaucoup d'incertitude

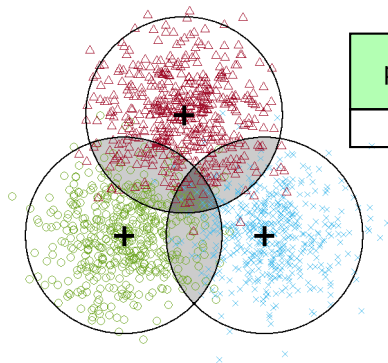
- états intermédiaires
- états atypiques



# Profil des patients ?

## Beaucoup d'incertitude

- états intermédiaires
- états atypiques



profil 1	profil 2	profil 3

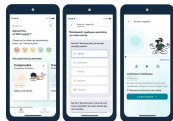
Conflit



# Motivation scientifique

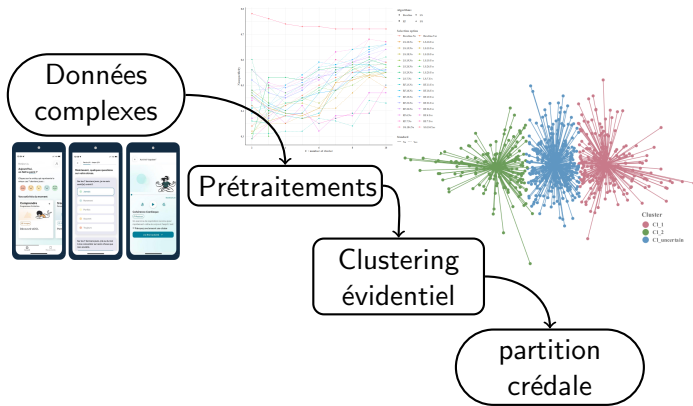
Création de clusters imprécis / incertain

Données  
complexes



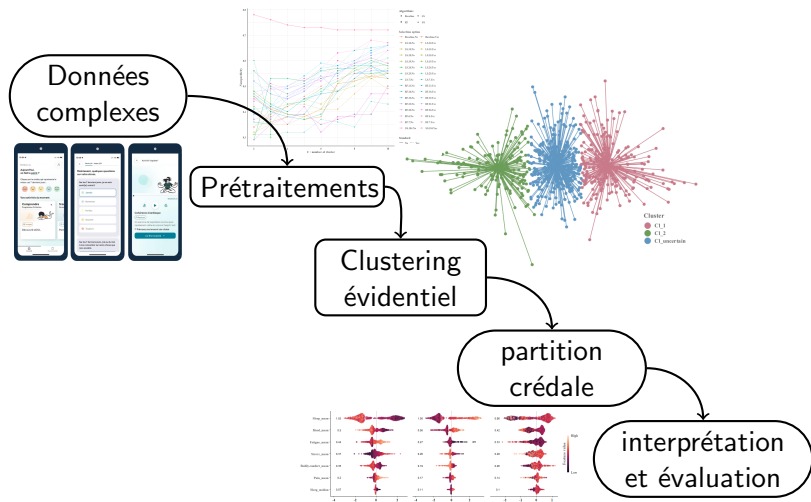
# Motivation scientifique

Création de clusters imprécis / incertain



# Motivation scientifique

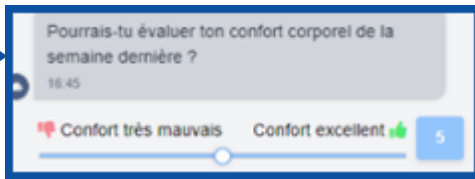
Création de clusters imprécis / incertain



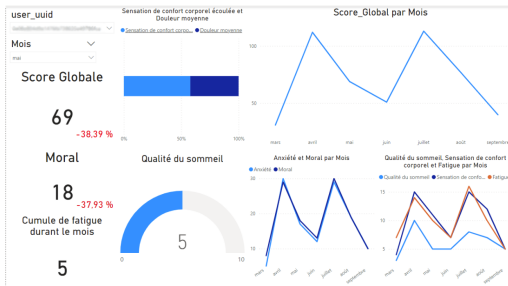
# Plan

- 1 Introduction
- 2 Prétraitements
  - Données eDOL
  - Extraction d'attributs
  - Sélection d'attributs
- 3 clustering
  - FCM
  - ECM
- 4 Résultats
- 5 Conclusion

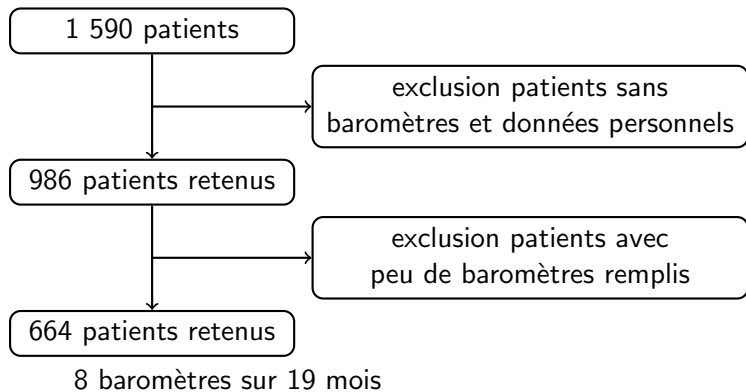
## Données eDOL



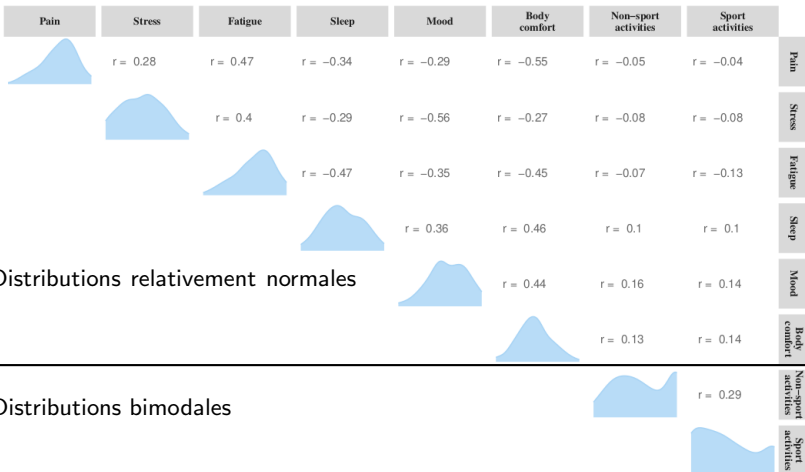
Baromètre	échelle	polarité
Sommeil	[0,10]	↗
Humeur		↗
Fatigue		↘
Confort corporel		↗
Stress		↘
Douleur		↘
Activités sportives		↗
Act. non-sportives	↗	



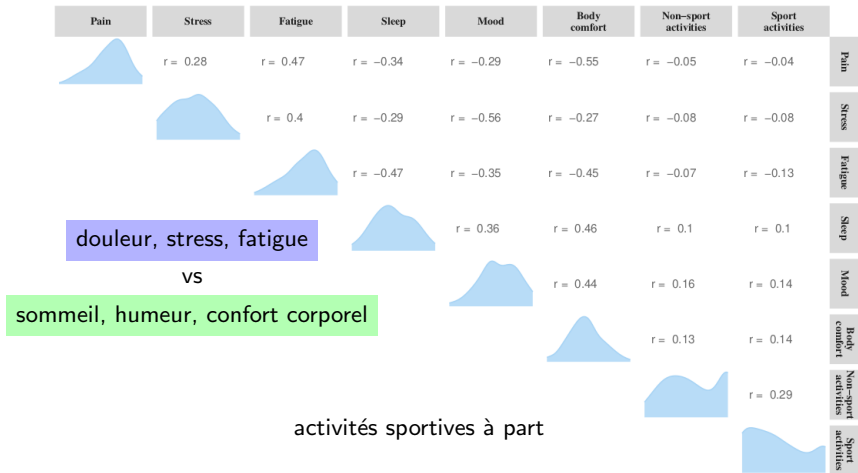
# Données eDOL



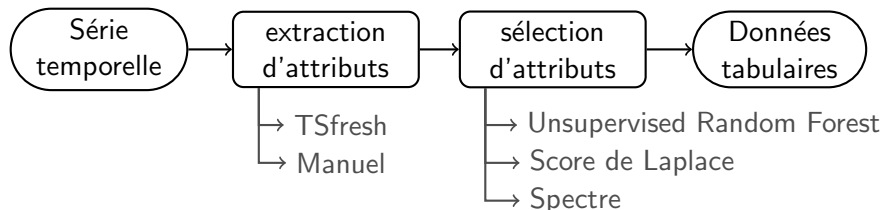
## Données eDOL



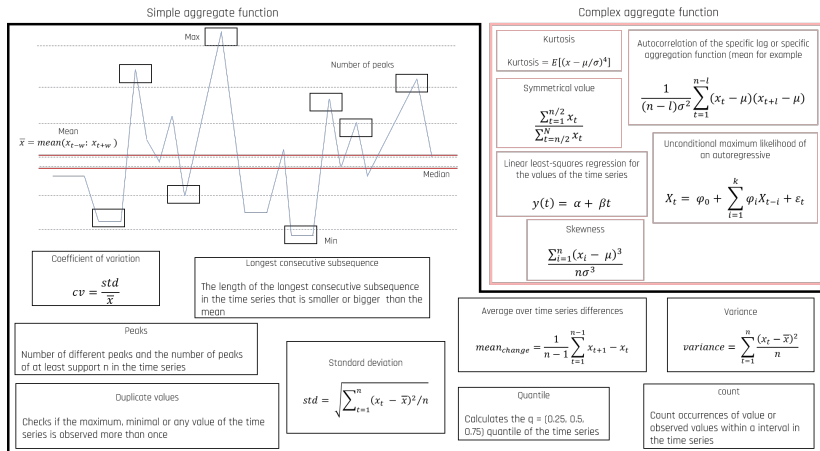
## Données eDOL



# Prétraitements

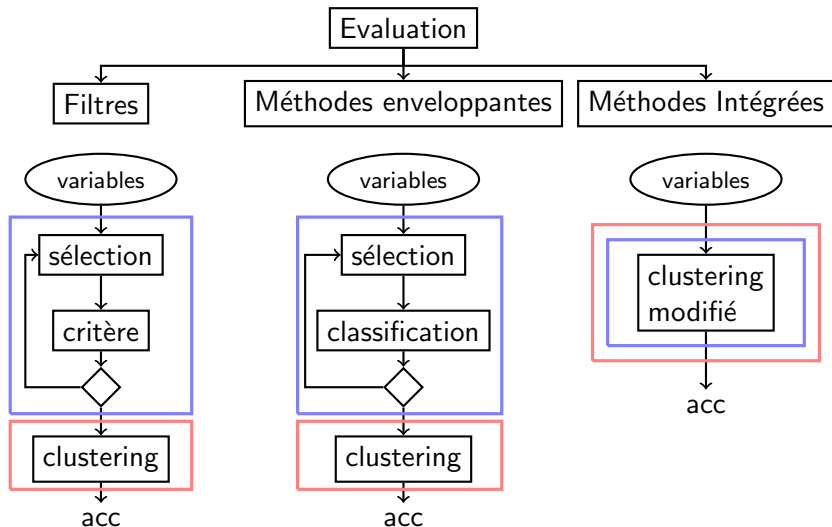


# Extraction d'attributs



23 formules  $\times$  8 baromètres = 184 attributs

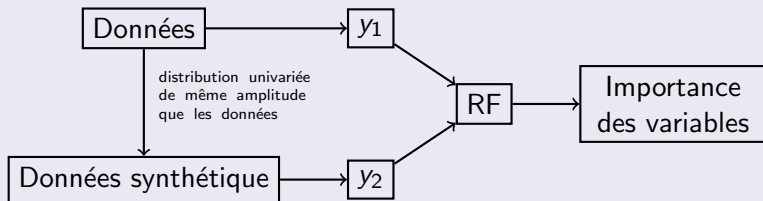
# Sélection d'attributs par des méthodes non supervisée



# Sélection d'attributs par des méthodes non supervisée

Algorithmes les plus efficaces dans le domaine de la santé [1,2]

## Méthode enveloppante : unsupervised Random Forest (RF)



## Méthode par filtre

- Score de Laplace : évalue la préservation de la structure locale
- Spectre : évalue la préservation de la structure globale

[1] S. Solorio-Fernández, J. Carrasco-Ochoa, and J. F. Martínez-Trinidad. A review of unsupervised feature selection methods. *Artificial Intelligence Review*, 2020.

[2] B. Remeseiro and V. Bolon-Canedo. A review of feature selection methods in medical applications. *Computers in Biology and Medicine*, 2019.

# Plan

- 1 Introduction
- 2 Prétraitements
  - Données eDOL
  - Extraction d'attributs
  - Sélection d'attributs
- 3 clustering**
  - FCM
  - ECM
- 4 Résultats
- 5 Conclusion

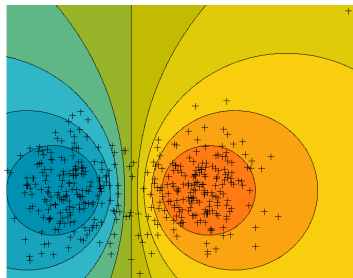
# Partition floue

- Chaque objet a un degré d'appartenance à chaque cluster
- $\mathbf{U} = (u_{ik})$  tel que  $u_{ik} \in [0, 1]$ ,  $\sum_{k=1}^c u_{ik} = 1$

## Exemple

- $\omega_1$  la classe des carrés
- $\omega_2$  la classe des cercles

	$p_{i1}$	$p_{i2}$
□	0	1
○	1	0
◻	0.9	0.1
◉	0.5	0.5



# Fuzzy c-means (FCM)

## Modèle géométrique

- Chaque cluster  $\omega_k$  est représenté par un centre  $\mathbf{v}_k$
- Distance Euclidienne  $d_{ik}^2 = (\mathbf{x}_i - \mathbf{v}_k)^T (\mathbf{x}_i - \mathbf{v}_k)$

## Fonction objectif

$$J_{FCM}(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^n \sum_{k=1}^c u_{ik}^\beta d_{ik}^2$$

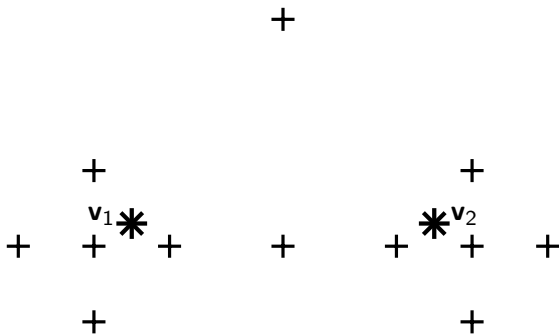
## tel que

$$\sum_{k=1}^c u_{ik} = 1 \text{ et } u_{ik} \geq 0 \quad \forall i, k$$

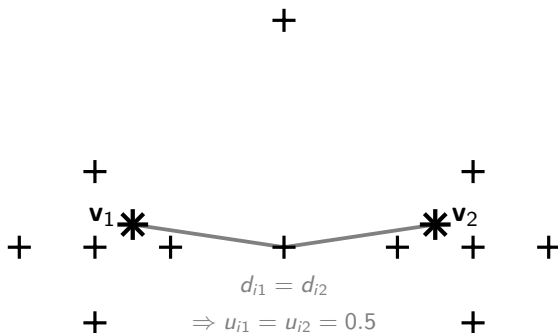
## Méthode d'optimisation de type Gauss-Seidel

$$\min_{\mathbf{U}} J_{FCM} \rightarrow \min_{\mathbf{V}} J_{FCM} \rightarrow \dots$$

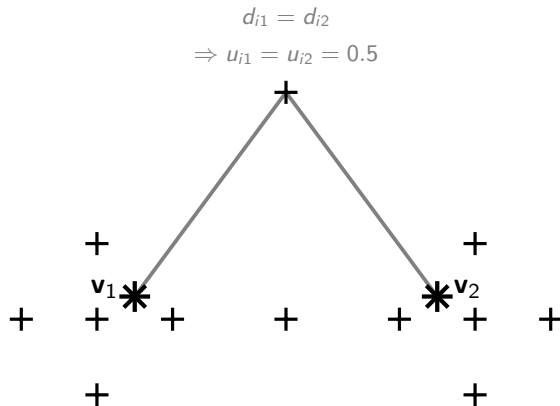
# Problématique : affectations imprécises et objets atypiques



# Problématique : affectations imprécises et objets atypiques



# Problématique : affectations imprécises et objets atypiques



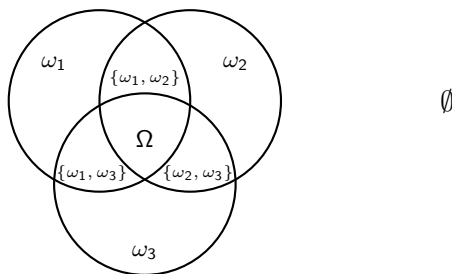
# Théorie des fonctions de croyance

Soit  $Y$  une variable prenant des valeurs dans l'ensemble fini  $\Omega$

Fonction de masse  $m : 2^\Omega \rightarrow [0, 1]$

$$\sum_{A \subseteq \Omega} m(A) = 1$$

- $m(A)$  : degré de croyances spécifique à  $Y \in A$
- Si  $m(A) > 0$  alors  $A$  est un ensemble focal

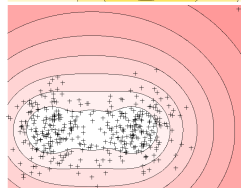
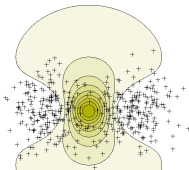
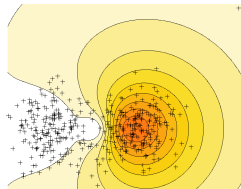
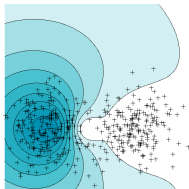


# Partition crédale

- Chaque objet a un degré de croyance pour chaque sous-ensemble  $A_j \subseteq \Omega$
- $\mathbf{M} = (m_{ij})$  tel que  $m_{ij} \in [0, 1]$ ,  $\sum_{A_j \subseteq \Omega} m_{ij} = 1$

## Exemple

	$m_{i\emptyset}$	$m_{i\omega_1}$	$m_{i\omega_2}$	$m_{i\Omega}$
○	0	0	1	0
□	0	1	0	0
◻	0	0.9	0.1	0
◻	0	0	0	1
☆	1	0	0	0



# Transformation crédale

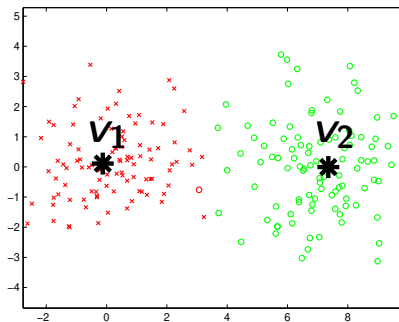
Transformation pignistique pour une prise de décision

$$BetP(\omega) = \frac{1}{1 - m(\emptyset)} \sum_{\{A \subseteq \Omega | \omega \in A\}} \frac{m(A)}{|A|}$$

	Partition crédale					Partition floue	
	$m_{i\emptyset}$	$m_{i\omega_1}$	$m_{i\omega_2}$	$m_{i\Omega}$		$u_{i\omega_1}$	$u_{i\omega_2}$
○	0	0	1	0	transformation → pignistique	0	1
□	0	1	0	0		1	0
◻	0	0.9	0.1	0		0.9	0.1
◐	0	0	0	1		0.5	0.5
☆	1	0	0	0		0.5	0.5

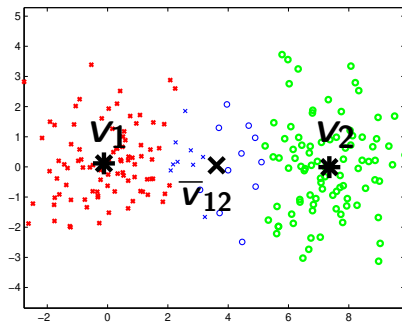
# Evidential c-means (ECM)

- Chaque cluster  $\omega_k$  est représenté par un centre  $\mathbf{v}_k$
- Centre  $\bar{\mathbf{v}}_j$  : barycentre des centres associés aux classes composant  $A_j \subseteq \Omega$
- Distance  $d_{ij}^2$  entre  $\mathbf{x}_i$  et  $\bar{\mathbf{v}}_j$



# Evidential c-means (ECM)

- Chaque cluster  $\omega_k$  est représenté par un centre  $\mathbf{v}_k$
- Centre  $\bar{\mathbf{v}}_j$  : barycentre des centres associés aux classes composant  $A_j \subseteq \Omega$
- Distance  $d_{ij}^2$  entre  $\mathbf{x}_i$  et  $\bar{\mathbf{v}}_j$



# Evidential c-means (ECM)

## Fonction objectif

$$J_{ECM}(\mathbf{M}, \mathbf{V}) = \sum_{i=1}^n \sum_{A_j \subseteq \Omega, A_j \neq \emptyset} |A_j|^\alpha m_{ij}^\beta d_{ij}^2 + \sum_{i=1}^n \delta^2 m_{i\emptyset}^\beta$$

## Tel que

$$\sum_{A_j \subseteq \Omega, A_j \neq \emptyset} m_{ij} + m_i(\emptyset) = 1, m_i(A_j) \geq 0 \quad \forall i, j$$

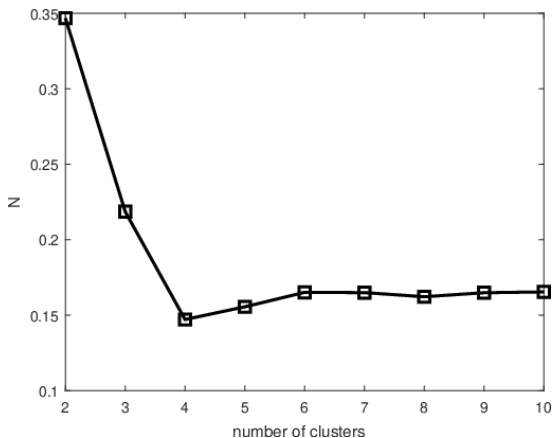
## Méthode d'optimisation de type Gauss-Seidel

$$\text{opt}(\mathbf{M}) \rightarrow \text{opt}(\mathbf{V}) \rightarrow \dots$$

# Détermination du nombre de clusters

Non spécificité : quantification de l'incertitude globale

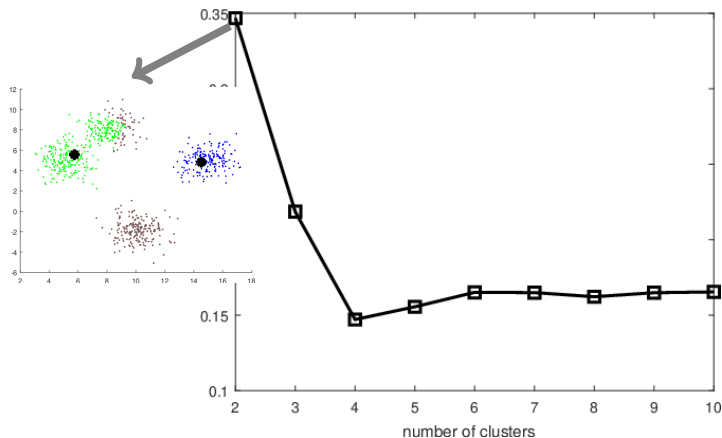
$$N(m) = \sum_{A \subseteq \Omega} m(A) \log(|A|) + m(\emptyset) \log_2(|\Omega|)$$



# Détermination du nombre de clusters

Non spécificité : quantification de l'incertitude globale

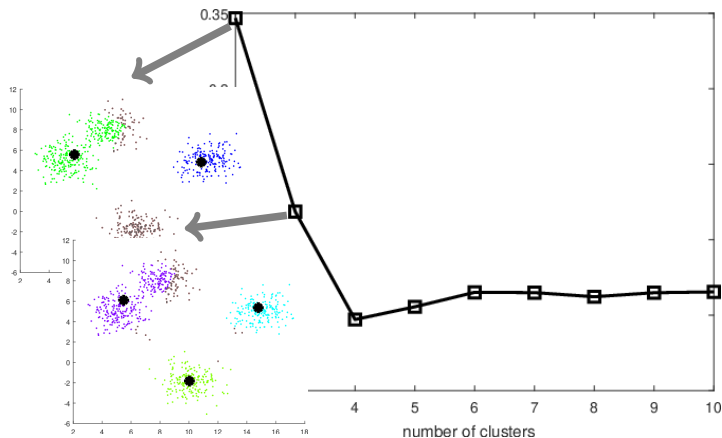
$$N(m) = \sum_{A \subseteq \Omega} m(A) \log(|A|) + m(\emptyset) \log_2(|\Omega|)$$



# Détermination du nombre de clusters

Non spécificité : quantification de l'incertitude globale

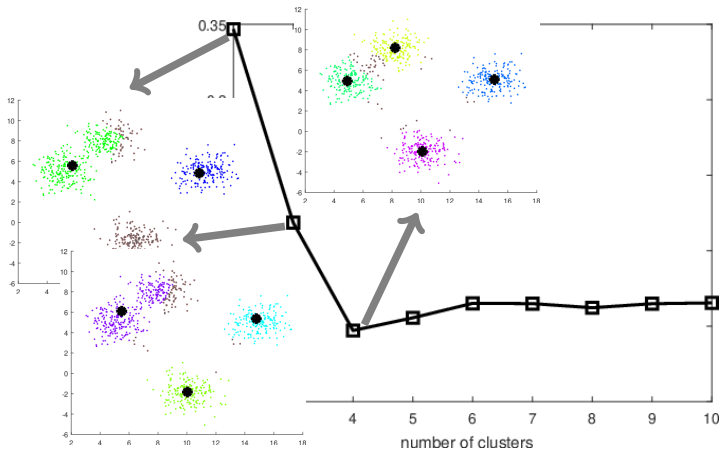
$$N(m) = \sum_{A \subseteq \Omega} m(A) \log(|A|) + m(\emptyset) \log_2(|\Omega|)$$



# Détermination du nombre de clusters

Non spécificité : quantification de l'incertitude globale

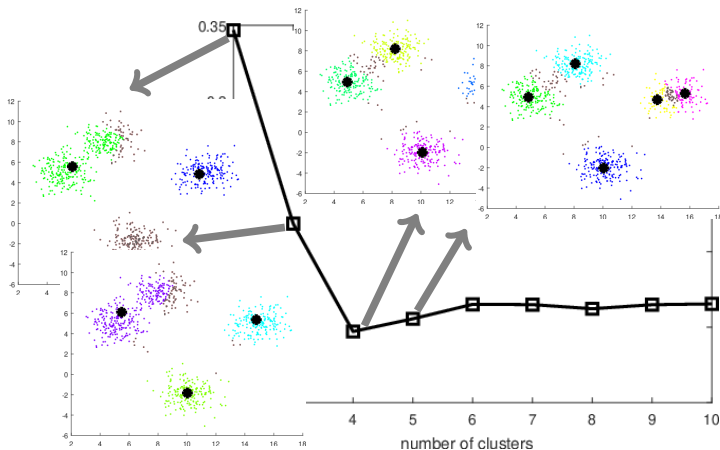
$$N(m) = \sum_{A \subseteq \Omega} m(A) \log(|A|) + m(\emptyset) \log_2(|\Omega|)$$



# Détermination du nombre de clusters

Non spécificité : quantification de l'incertitude globale

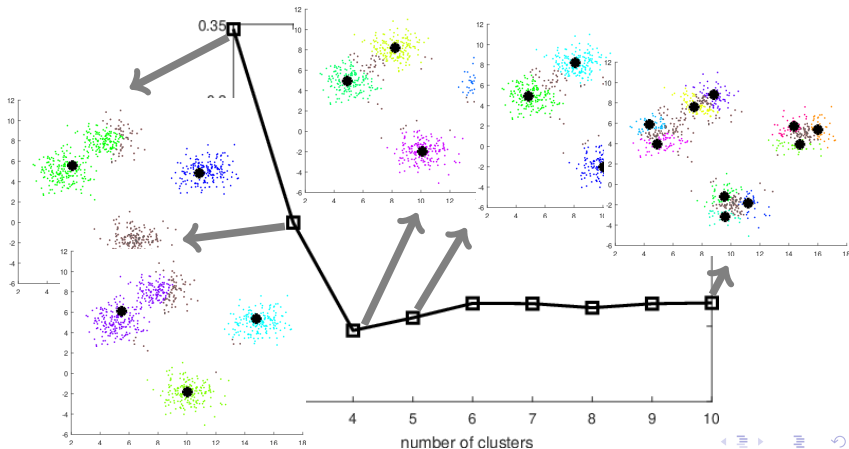
$$N(m) = \sum_{A \subseteq \Omega} m(A) \log(|A|) + m(\emptyset) \log_2(|\Omega|)$$



# Détermination du nombre de clusters

Non spécificité : quantification de l'incertitude globale

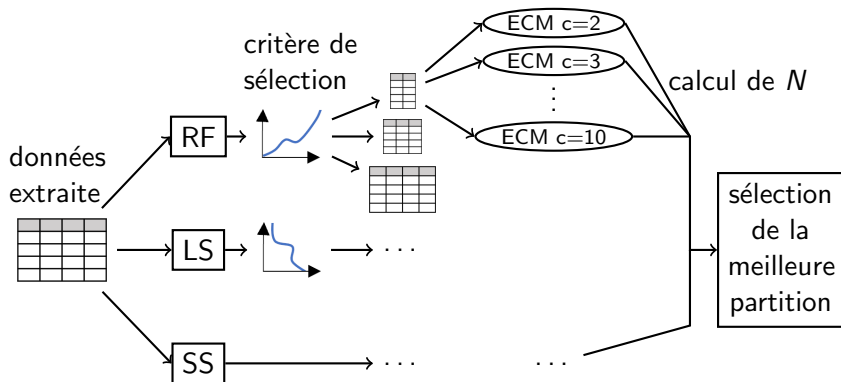
$$N(m) = \sum_{A \subseteq \Omega} m(A) \log(|A|) + m(\emptyset) \log_2(|\Omega|)$$



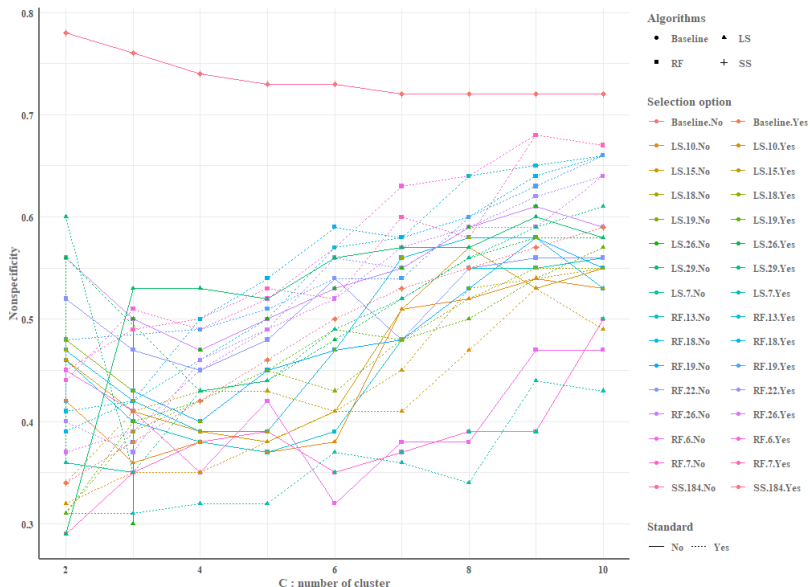
# Plan

- 1 Introduction
- 2 Prétraitements
  - Données eDOL
  - Extraction d'attributs
  - Sélection d'attributs
- 3 clustering
  - FCM
  - ECM
- 4 Résultats
- 5 Conclusion

# Protocole expérimental



# Recherche du nombre de clusters

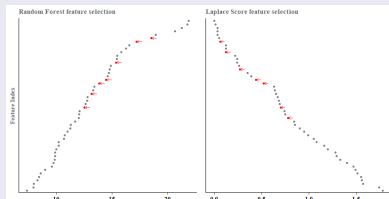


# Recherche du meilleur algorithme de sélection d'attributs

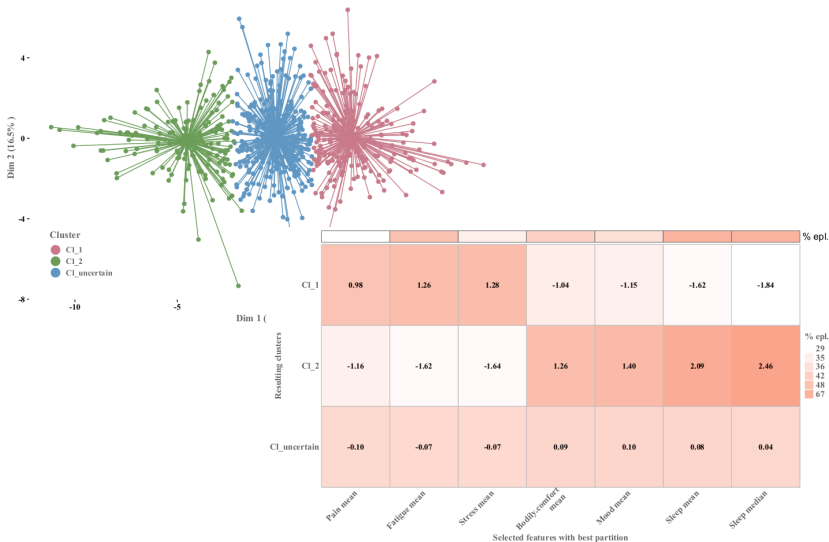
Algorithms	# feat.	C	$N^*$	Time (s)
Laplacian Score	7	2	0.31	1.1
<b>Random Forest</b>	<b>7</b>	<b>2</b>	<b>0.29</b>	<b>2.7</b>
Spectral Score	184	2	0.34	3.7
Baseline	184	2	0.34	3.7

## Interprétations

- Spectral Score : importance équitable à toutes les variables  
 ⇒ inadapté pour des variables très corrélées comme ici
- Random Forest et Laplacian Score : possibilité de sélectionner selon un seuil



# Interprétation des résultats de clustering



# Interprétation des résultats de clustering

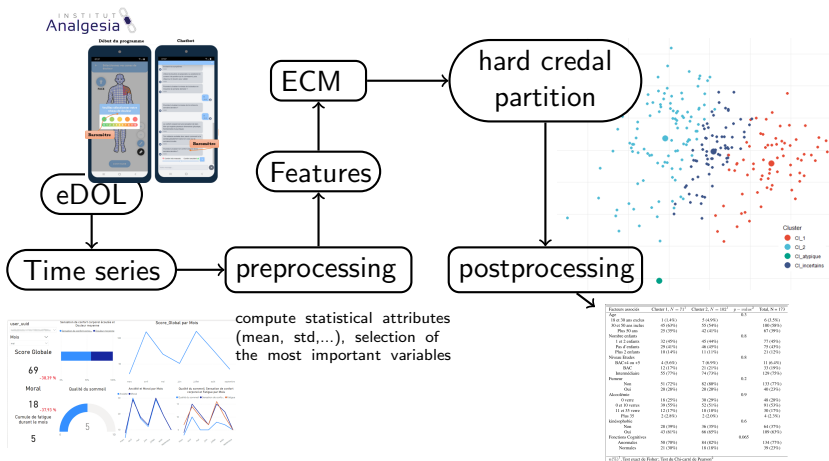
## Identification de deux groupes

- Patients avec des douleurs liées à la fatigue et au stress
- Patients avec des douleurs liées au sommeil, au confort physique, et à l'humeur

# Plan

- 1 Introduction
- 2 Prétraitements
  - Données eDOL
  - Extraction d'attributs
  - Sélection d'attributs
- 3 clustering
  - FCM
  - ECM
- 4 Résultats
- 5 Conclusion

## Conclusion ...



[1] A. Soubeiga, V. Antoine, A. Corteval, N. Kerckhove, S. Moreno, I. Falih, J. Phalip. *Clustering and Interpretation of time-series trajectories of chronic pain using evidential c-means*, Expert Systems With Application, 2025

# ... et perspectives

## Clustering

- création d'un algorithme relationnel évidentiel
- création d'un algorithme évidentiel qui intègre la sélection d'attributs

## Santé

- découper les séries temporelles pour prendre en compte une suite d'état chez le patient
- chaîner les données avec le SNDS

