

BELIEF DECISION TREES



Patrick Vannoorenberghe

LABORATOIRE HEUDIASYC - UMR 6599 CNRS

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

B.P. 20529, 60205 COMPIEGNE CEDEX

www.hds.utc.fr/~pvannoo


RDF et Apprentissage

- Assigner \mathbf{x} à une classe dans $\Omega = \{\omega_k, k = 1, \dots, K\}$
 - $\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, n\}$: Apprentissage
 - \mathbf{x}_i : vecteur forme de dimension p
 - ω_i : étiquetage, label, vraie classe

RDF et Apprentissage

- Assigner \mathbf{x} à une classe dans $\Omega = \{\omega_k, k = 1, \dots, K\}$
 - $\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, n\}$: Apprentissage
 - \mathbf{x}_i : vecteur forme de dimension p
 - ω_i : étiquetage, label, vraie classe
- Choisir une action dans un ensemble \mathcal{A} d'actions
 - $\lambda : \mathcal{A} \times \Omega \rightarrow \mathbb{R}$: une fonction de coût
 - $\mathcal{A} = \{a_0, a_u, a_k\}$: Assigner \mathbf{x} à la classe ω_k

RDF et Apprentissage

- Assigner \mathbf{x} à une classe dans $\Omega = \{\omega_k, k = 1, \dots, K\}$
 - $\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, n\}$: Apprentissage
 - \mathbf{x}_i : vecteur forme de dimension p
 - ω_i : étiquetage, label, vraie classe
- Choisir une action dans un ensemble \mathcal{A} d'actions
 - $\lambda : \mathcal{A} \times \Omega \rightarrow \mathbb{R}$: une fonction de coût
 - $\mathcal{A} = \{a_0, a_u, a_k\}$: Assigner \mathbf{x} à la classe ω_k
- **Classifieurs crédibilistes**
 - Façon flexible de représenter l'incertitude
 - **Approche basée sur les arbres de décision** 

Points abordés

1. Théorie des fonctions de croyance

Points abordés

1. Théorie des fonctions de croyance
2. Inférence crédale
 - Le problème
 - La solution crédibiliste

Points abordés

1. Théorie des fonctions de croyance
2. Inférence crédale
 - Le problème
 - La solution crédibiliste
3. **Belief Decision Trees**
 - Rappels sur les arbres de décision
 - Principes des BDT
 - Gestion de labels incertains

Points abordés

1. Théorie des fonctions de croyance
2. Inférence crédale
 - Le problème
 - La solution crédibiliste
3. **Belief Decision Trees**
 - Rappels sur les arbres de décision
 - Principes des BDT
 - Gestion de labels incertains
4. Exemples d'application
 - Comparaison et simulations
 - Application CETIM

Théorie des fonctions de croyance

- A. Dempster (1967) : *Upper and Lower Probabilities*
- G. Shafer (1976) : *Mathematical Theory of Evidence*
- P. Smets (1994) : *Transferable Belief Model*

Théorie des fonctions de croyance

- A. Dempster (1967) : *Upper and Lower Probabilities*
- G. Shafer (1976) : *Mathematical Theory of Evidence*
- P. Smets (1994) : *Transferable Belief Model*

- Cadre de discernement : $\Omega = \{\omega_k\}$ avec $k \in [1, K]$
 - L'ensemble Ω doit être exhaustif
 - Les ω_k doivent être exclusives ($\omega_i \cap \omega_j = \emptyset$)
 - 2^Ω : le "Power Set"; les sous-ensembles $A \subseteq \Omega$

Éléments mathématiques

- Fonction de croyance : $2^\Omega \rightarrow [0, 1]$

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega$$

Éléments mathématiques

- Fonction de croyance : $2^\Omega \rightarrow [0, 1]$

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega$$

- avec m masse de croyance élémentaire (bba)

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

Éléments mathématiques

- Fonction de croyance : $2^\Omega \rightarrow [0, 1]$

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega$$

- avec m masse de croyance élémentaire (bba)

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

- Noyau N_Ω
 - Éléments focaux : masse non nulle
 - $N_\Omega = \{A \subseteq \Omega \mid m(A) > 0\}$

Pourquoi des fonctions de croyance ?

- Plusieurs interprétations
 - Upper and Lower Probability Model
 - Dempster's Model et Hint Model de Kohlas et Monney
 - **Transferable Belief Model** de Smets

Pourquoi des fonctions de croyance ?

- Plusieurs interprétations
 - Upper and Lower Probability Model
 - Dempster's Model et Hint Model de Kohlas et Monney
 - **Transferable Belief Model** de Smets
- Que représente *bel* en RdF ?
 - Dans Ω , une classe ω_0 est la vraie classe !
 - L'agent, celui qui détient la croyance ne connaît pas avec certitude la vraie classe!
 - Une source d'information donne une opinion pondérée sur la classe d'appartenance!

Extension aux labels incertains

Ensemble d'apprentissage $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$

Extension aux labels incertains

Ensemble d'apprentissage $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$

- Label certain

Élément focal de $m_i \Rightarrow$ singleton ω_i

Extension aux labels incertains

Ensemble d'apprentissage $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$

- Label certain

Élément focal de $m_i \Rightarrow$ singleton ω_i

- Label probabiliste

Éléments focaux de $m_i \Rightarrow$ singletons

Extension aux labels incertains

Ensemble d'apprentissage $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$

- Label certain
Élément focal de $m_i \Rightarrow$ singleton ω_i
- Label probabiliste
Éléments focaux de $m_i \Rightarrow$ singletons
- Label ensembliste
Élément focal $A \subseteq \Omega$ tel de $m_i(A) = 1$

Extension aux labels incertains

Ensemble d'apprentissage $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$

- Label certain
Élément focal de $m_i \Rightarrow$ singleton ω_i
- Label probabiliste
Éléments focaux de $m_i \Rightarrow$ singletons
- Label ensembliste
Élément focal $A \subseteq \Omega$ tel de $m_i(A) = 1$
- Label possibiliste
Éléments focaux de m_i emboîtés

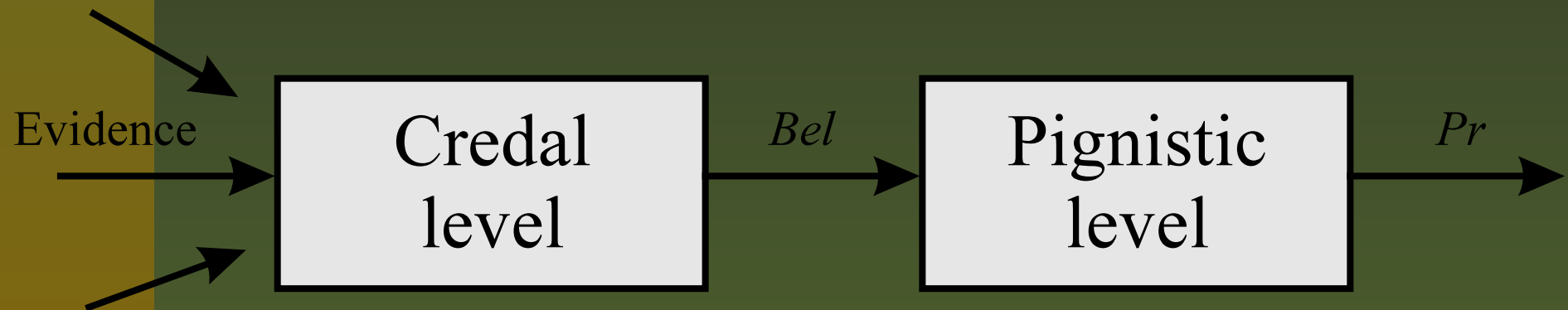
Extension aux labels incertains

Ensemble d'apprentissage $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$

- Label certain
Élément focal de $m_i \Rightarrow$ singleton ω_i
- Label probabiliste
Éléments focaux de $m_i \Rightarrow$ singletons
- Label ensembliste
Élément focal $A \subseteq \Omega$ tel de $m_i(A) = 1$
- Label possibiliste
Éléments focaux de m_i emboîtés
- Label quelconque $\Rightarrow m_i = ?$

Modèle des croyances transférables

- 2 niveaux (credal and pignistic levels)
 - la représentation et la révision des croyances (Bel)
 - la prise de décision (Pr)



Combiner des données incertaines \oplus

- Règle conjonctive
 - Sachant m_1 et m_2 sur Ω
 - Construire $m_\cap = m_1 \cap m_2$

$$m_\cap(A) = \sum_{B \cap C = A} m_1(B) m_2(C) \quad \forall A \subseteq \Omega$$

Combiner des données incertaines \oplus

- Règle conjonctive

- Sachant m_1 et m_2 sur Ω

- Construire $m_{\cap} = m_1 \cap m_2$

$$m_{\cap}(A) = \sum_{B \cap C = A} m_1(B) m_2(C) \quad \forall A \subseteq \Omega$$

- Une fonction intéressante

$$q(A) = \sum_{B \supseteq A} m(B) \quad \forall A \subseteq \Omega$$

→ on a $q(A) = q_1(A)q_2(A)$

Prise de décision

- Transformer m en une fonction de probabilité p_m

Prise de décision

- Transformer m en une fonction de probabilité p_m
- Transformation pignistique

$$p_m(\omega_k) = \sum_{A \ni \omega_k} \frac{m(A)}{|A|} \quad \forall \omega_k \in \Omega$$

Prise de décision

- Transformer m en une fonction de probabilité p_m
- Transformation pignistique

$$p_m(\omega_k) = \sum_{A \ni \omega_k} \frac{m(A)}{|A|} \quad \forall \omega_k \in \Omega$$

- $m(A)$ est distribuée de manière uniforme sur les éléments de A

Prise de décision

- Transformer m en une fonction de probabilité p_m
- Transformation pignistique

$$p_m(\omega_k) = \sum_{A \ni \omega_k} \frac{m(A)}{|A|} \quad \forall \omega_k \in \Omega$$

- $m(A)$ est distribuée de manière uniforme sur les éléments de A
- Théorie de la décision classique

Incertitude et croyance

■ Non-spécificité

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

- $N(m) = 0 \iff$ probabilité
- $N(m) = \log_2 |\Omega| \iff m(\Omega) = 1$ (vacuous function)

Incertitude et croyance

■ Non-spécificité

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

- $N(m) = 0 \iff$ probabilité
- $N(m) = \log_2 |\Omega| \iff m(\Omega) = 1$ (vacuous function)

■ Conflit

$$D(m) = - \sum_{A \subseteq \Omega} m(A) \log_2 p_m(A)$$

- $D(m) = \log_2 |\Omega| \iff$ probabilité uniforme

Incertitude : Une mesure totale

- Mesure totale d'incertitude

$$U(m) = (1 - \lambda)N(m) + \lambda D(m)$$

Incertitude : Une mesure totale

- Mesure totale d'incertitude

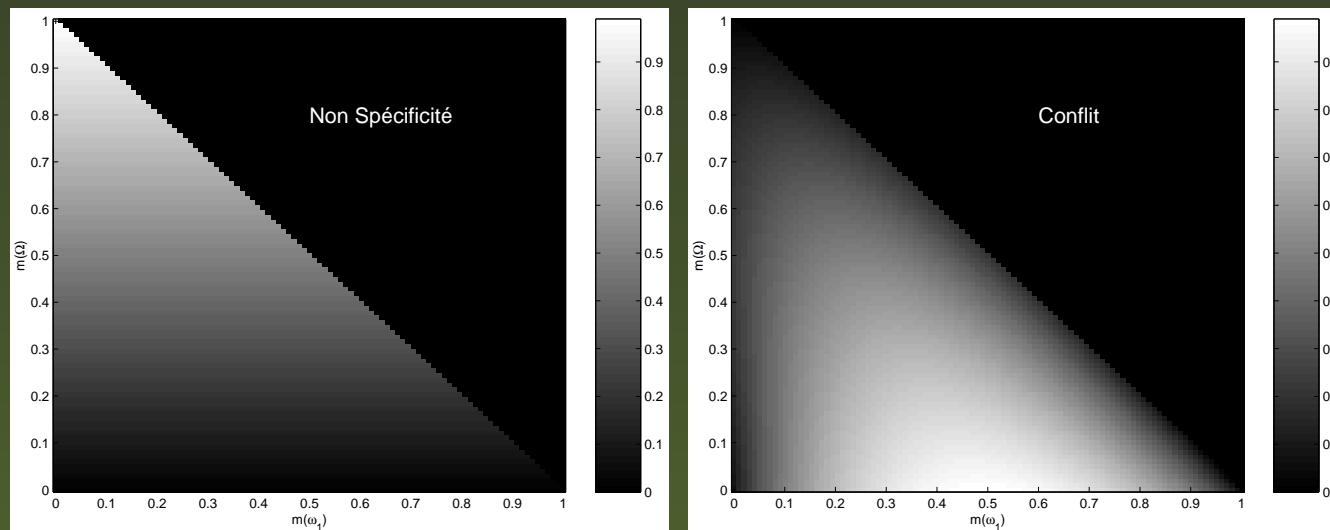
$$U(m) = (1 - \lambda)N(m) + \lambda D(m)$$

- Propriétés

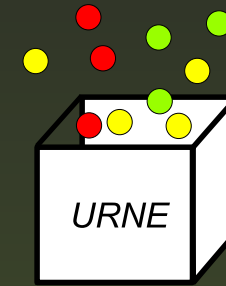
- Combinaison linéaire de N et de D
- Additivité, continuité, échelle $[0, f(|\Omega|)]$, ...
- $\lambda \in [0, 1]$, Klir $\lambda = 0.5$

Exemple à 2 classes

- $\Omega = \{\omega_1, \omega_2\}$
- Evolution des 2 mesures dans l'espace $m(\Omega), m(\{\omega_1\})$

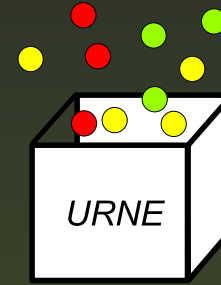


Inférence crédale



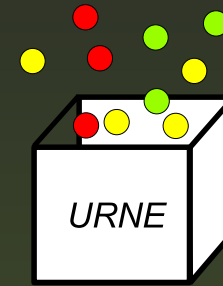
- Modèle de l'urne
 - Balles : $\mathcal{N} = \{1, 2, \dots, N\}$
 - Couleurs : $\mathcal{C} = \{c_1, c_2, \dots, c_K\}$

Inférence crédale



- Modèle de l'urne
 - Balles : $\mathcal{N} = \{1, 2, \dots, N\}$
 - Couleurs : $\mathcal{C} = \{c_1, c_2, \dots, c_K\}$
- Informations disponibles
 - Nombre de balles sélectionnées avec remplacement
 - $couleur(b_4) = c_1, couleur(b_7) \in \{c_1, c_2\}$

Inférence crédale



- Modèle de l'urne
 - Balles : $\mathcal{N} = \{1, 2, \dots, N\}$
 - Couleurs : $\mathcal{C} = \{c_1, c_2, \dots, c_K\}$
- Informations disponibles
 - Nombre de balles sélectionnées avec remplacement
 - $couleur(b_4) = c_1, couleur(b_7) \in \{c_1, c_2\}$
- Problèmes
 - Croyance sur la composition de l'urne ?
 - Croyance sur la couleur du prochain tirage ?

Le problème de l'inférence

- Essai de Bernoulli



- Exemple :

- Pile ou Face : $\Omega = \{P, F\}$
- n tirages dont n_p piles et n_f faces

La solution crédibiliste

- Ph. SMETS \Rightarrow cas où $|\Omega| = K = 2$

La solution crédibiliste

- Ph. SMETS \Rightarrow cas où $|\Omega| = K = 2$
- Fonction de croyance

$$m^{\Omega}[n_p, n_f](\{P\}) = \frac{n_p}{n + 1}$$

$$m^{\Omega}[n_p, n_f](\{F\}) = \frac{n_f}{n + 1}$$

$$m^{\Omega}[n_p, n_f](\Omega) = \frac{1}{n + 1}$$

La solution crédibiliste

- Ph. SMETS \Rightarrow cas où $|\Omega| = K = 2$
- Fonction de croyance

$$m^{\Omega}[n_p, n_f](\{P\}) = \frac{n_p}{n + 1}$$

$$m^{\Omega}[n_p, n_f](\{F\}) = \frac{n_f}{n + 1}$$

$$m^{\Omega}[n_p, n_f](\Omega) = \frac{1}{n + 1}$$

- Solution tend vers la vraie probabilité quand $n \rightarrow \infty$

La solution crédibiliste

- Ph. SMETS \Rightarrow cas où $|\Omega| = K = 2$
- Fonction de croyance

$$m^{\Omega}[n_p, n_f](\{P\}) = \frac{n_p}{n + 1}$$

$$m^{\Omega}[n_p, n_f](\{F\}) = \frac{n_f}{n + 1}$$

$$m^{\Omega}[n_p, n_f](\Omega) = \frac{1}{n + 1}$$

- Solution tend vers la vraie probabilité quand $n \rightarrow \infty$
- Incertitude notamment quand n petit

Arbres de décision crédibilistes

- Les arbres de décision
 - Rappels
 - Apprentissage

Arbres de décision crédibilistes

- Les arbres de décision
 - Rappels
 - Apprentissage
- Belief Decision Tree : Principes
 - Gestion de labels certains
 - Procédure d'élagage

Arbres de décision crédibilistes

- Les arbres de décision
 - Rappels
 - Apprentissage
- Belief Decision Tree : Principes
 - Gestion de labels certains
 - Procédure d'élagage
- Gestion de labels incertains
 - Problème à deux classes
 - Approche par dichotomie K classes

Les arbres de décision : Rappels

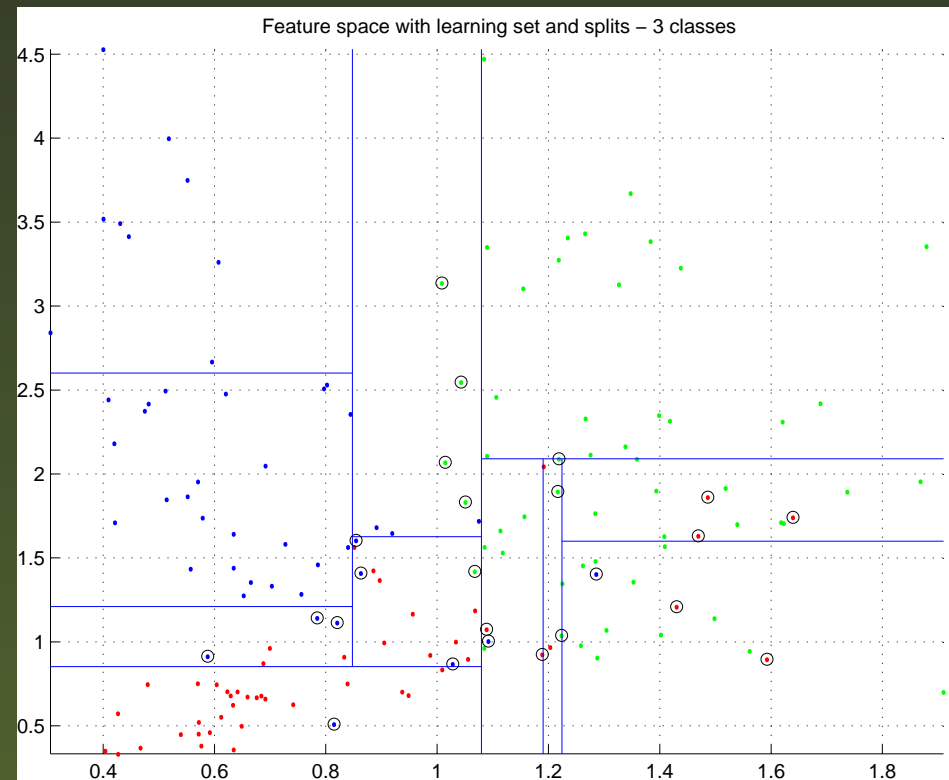
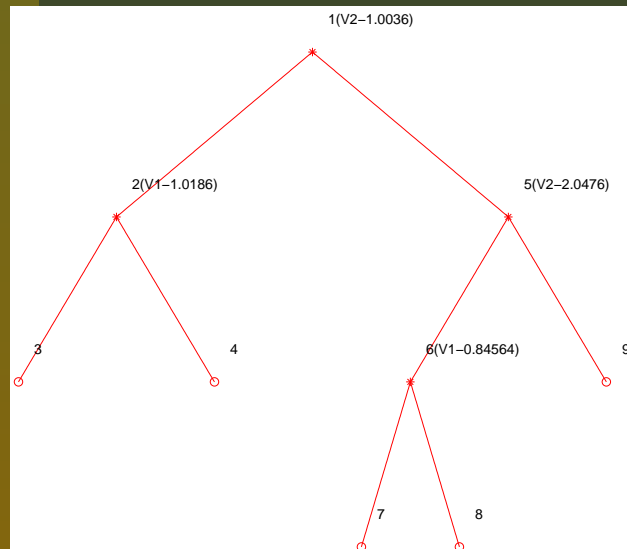
- Procédure de classification séquentielle
 - Graphe direct et acyclique composé de noeuds t
 - Noeud = { Décision, Feuille }
 - Noeud Décision = Test sur une variable + c successeurs
 - Noeud Feuille = Noeud terminal avec étiquetage

Les arbres de décision : Rappels

- Procédure de classification séquentielle
 - Graphe direct et acyclique composé de noeuds t
 - Noeud = { Décision, Feuille }
 - Noeud Décision = Test sur une variable + c successeurs
 - Noeud Feuille = Noeud terminal avec étiquetage
- Avantages
 - Représentation graphique des règles
 - Phase de test très rapide

Les arbres de décision : Rappels

- Exemples : Arbre, Espace de décision



Induction de l'arbre et classification

- Induction (Apprentissage)
 - Stratégie top-down

Induction de l'arbre et classification

- Induction (Apprentissage)
 - Stratégie top-down
 - Mesure d'impureté du style entropique

$$i(t) = - \sum_k \frac{n_k(t)}{n(t)} \log_2 \frac{n_k(t)}{n(t)}$$

Induction de l'arbre et classification

- Induction (Apprentissage)

- Stratégie top-down

- Mesure d'impureté du style entropique

$$i(t) = - \sum_k \frac{n_k(t)}{n(t)} \log_2 \frac{n_k(t)}{n(t)}$$

- Test sur les variables

$$\Delta i(s, t) = i(t) - (p_L i(t_L) + p_R i(t_R))$$

- Procédure d'arrêt

Induction de l'arbre et classification

- Induction (Apprentissage)

- Stratégie top-down

- Mesure d'impureté du style entropique

$$i(t) = - \sum_k \frac{n_k(t)}{n(t)} \log_2 \frac{n_k(t)}{n(t)}$$

- Test sur les variables

$$\Delta i(s, t) = i(t) - (p_L i(t_L) + p_R i(t_R))$$

- Procédure d'arrêt

- Classification

- Etiquetage de chaque feuille de l'arbre

BDT : Principes

- Apprentissage $\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, n\}$

BDT : Principes

- Apprentissage $\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, n\}$
- Fonction de croyance induite au noeud t

$$m^{\Omega}[t](\{\omega_k\}) = \frac{n_k(t)}{n(t) + 1}, \quad k \in \{1, 2\}$$

$$m^{\Omega}[t](\Omega) = \frac{1}{n(t) + 1}$$

- avec $n_k(t)$: nombre d'exemples de la classe ω_k

BDT : Principes

- Apprentissage $\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, n\}$
- Fonction de croyance induite au noeud t

$$m^\Omega[t](\{\omega_k\}) = \frac{n_k(t)}{n(t) + 1}, \quad k \in \{1, 2\}$$

$$m^\Omega[t](\Omega) = \frac{1}{n(t) + 1}$$

- avec $n_k(t)$: nombre d'exemples de la classe ω_k
- Solution **heuristique** valable pour $K = 2$ classes

BDT : Principes

- Mesure de non-spécificité et conflit au noeud t

$$N(m^\Omega[t]) = \frac{1}{n(t) + 1}$$

$$D(m^\Omega[t]) = - \sum_{k=1}^2 \frac{n_k(t)}{n(t) + 1} \log_2 \left(\frac{n_k(t) + 1/2}{n(t) + 1} \right)$$

BDT : Principes

- Mesure de non-spécificité et conflit au noeud t

$$N(m^\Omega[t]) = \frac{1}{n(t) + 1}$$

$$D(m^\Omega[t]) = - \sum_{k=1}^2 \frac{n_k(t)}{n(t) + 1} \log_2 \left(\frac{n_k(t) + 1/2}{n(t) + 1} \right)$$

- Mesure d'incertitude totale au noeud t

$$U(m^\Omega[t]) = (1 - \lambda)N(m^\Omega[t]) + \lambda D(m^\Omega[t])$$

- λ contrôle l'induction (donc le pré-élagage)

Induction du BDT

- Mesure de sélection d'un test à un noeud *Décision*

$$\Delta U_\lambda(s, t) = U_\lambda(t) - (p_L U_\lambda(t_L) + p_R U_\lambda(t_R))$$

Induction du BDT

- Mesure de sélection d'un test à un noeud *Décision*

$$\Delta U_\lambda(s, t) = U_\lambda(t) - (p_L U_\lambda(t_L) + p_R U_\lambda(t_R))$$

- Par rapport aux algorithmes classiques (C4.5, ID3)
 - Obtient-on les mêmes performances ?
 - Peut-on résoudre d'autres problèmes d'apprentissage ?

Gestion de labels incertains $|\Omega| = 2$

- Ensemble d'apprentissage

$$\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$$

Gestion de labels incertains $|\Omega| = 2$

- Ensemble d'apprentissage
 $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$
- Fonction de croyance $m^\Omega[t]$ dérivée des fonctions m_i
 $\Rightarrow |\Omega| = 2$

$$m^\Omega[t](\{\omega_1\}) = \sum_{j+k \leq n(t)} \alpha_{jk} \frac{j}{j+k+1}$$

Gestion de labels incertains $|\Omega| = 2$

- Ensemble d'apprentissage

$$\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$$

- Fonction de croyance $m^\Omega[t]$ dérivée des fonctions m_i
 $\Rightarrow |\Omega| = 2$

$$m^\Omega[t](\{\omega_1\}) = \sum_{j+k \leq n(t)} \alpha_{jk} \frac{j}{j+k+1}$$

- Expressions similaires pour $m^\Omega[t](\{\omega_2\})$ et $m^\Omega[t](\Omega)$

Approche par dichotomie

- $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$ avec m_i quelconque

Approche par dichotomie

- $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$ avec m_i quelconque
 - 💡 Utiliser les résultats avec $|\Omega| = 2$
 - Construction de K arbres par dichotomie

Approche par dichotomie

- $\mathcal{L} = \{(\mathbf{x}_i, m_i), i = 1, \dots, n\}$ avec m_i quelconque
 - 💡 Utiliser les résultats avec $|\Omega| = 2$
 - Construction de K arbres par dichotomie
- Fonction sur $\Theta_k = \{\theta_{k,1}, \theta_{k,2}\} \Rightarrow$ coarsening

$$m_i^{\Theta_k}(\{\theta_{k,1}\}) = m_i^{\Omega}(\{\omega_k\})$$

$$m_i^{\Theta_k}(\{\theta_{k,2}\}) = \sum_{\{A \subseteq \Omega \mid \omega_k \notin A\}} m_i^{\Omega}(A)$$

$$m_i^{\Theta_k}(\Theta_k) = \sum_{\{A \subseteq \Omega \mid |A| > 1, \omega_k \in A\}} m_i^{\Omega}(A)$$

En phase de test

- Extension des fonctions \hat{m}^{Θ_k} sur Ω pour l'exemple \mathbf{x}

$$\hat{m}_k^\Omega(\{\omega_k\}) = \hat{m}^{\Theta_k}(\{\theta_{k,1}\})$$

$$\hat{m}_k^\Omega(\Omega \setminus \{\omega_k\}) = \hat{m}^{\Theta_k}(\{\theta_{k,2}\})$$

$$\hat{m}_k^\Omega(\Omega) = \hat{m}^{\Theta_k}(\Theta_k)$$

En phase de test

- Extension des fonctions \hat{m}^{Θ_k} sur Ω pour l'exemple \mathbf{x}

$$\hat{m}_k^{\Omega}(\{\omega_k\}) = \hat{m}^{\Theta_k}(\{\theta_{k,1}\})$$

$$\hat{m}_k^{\Omega}(\Omega \setminus \{\omega_k\}) = \hat{m}^{\Theta_k}(\{\theta_{k,2}\})$$

$$\hat{m}_k^{\Omega}(\Omega) = \hat{m}^{\Theta_k}(\Theta_k)$$

- Aggrégation des fonctions de croyance $m_{\mathbf{x},k}^{\Omega}$

En phase de test

- Extension des fonctions \hat{m}^{Θ_k} sur Ω pour l'exemple \mathbf{x}

$$\hat{m}_k^\Omega(\{\omega_k\}) = \hat{m}^{\Theta_k}(\{\theta_{k,1}\})$$

$$\hat{m}_k^\Omega(\Omega \setminus \{\omega_k\}) = \hat{m}^{\Theta_k}(\{\theta_{k,2}\})$$

$$\hat{m}_k^\Omega(\Omega) = \hat{m}^{\Theta_k}(\Theta_k)$$

- Aggrégation des fonctions de croyance $m_{\mathbf{x},k}^\Omega$
 - Opérateur idempotent, symétrique et linéaire

$$\hat{m}^\Omega = \frac{1}{K} \sum_{k=1}^K \hat{m}_k^\Omega$$

Définition des coûts

- $L : \mathcal{A} \times \Omega \mapsto \mathbb{R}$ + Extension au labels incertains

$$L(\hat{m}, \omega) = 1 - p_{\hat{m}}(\omega)$$

$$L(\hat{m}, m) = 1 - \sum_{A \subseteq \Omega} m(A) p_{\hat{m}}(A)$$

Définition des coûts

- $L : \mathcal{A} \times \Omega \mapsto \mathbb{R} +$ Extension au labels incertains

$$L(\hat{m}, \omega) = 1 - p_{\hat{m}}(\omega)$$

$$L(\hat{m}, m) = 1 - \sum_{A \subseteq \Omega} m(A) p_{\hat{m}}(A)$$

- Deux fonctions de coût \Rightarrow évaluation

$$C_{1ou2} = 1 - \frac{1}{n'} \sum_{i=1}^{n'} \sum_{\omega \in \Omega} p_{\hat{m}_i}(\omega) \cdot q_i(\{\omega\}) \text{ ou } p_{m_i}(\{\omega\})$$

Illustrations

- Comparaison avec approche classique

Illustrations

- Comparaison avec approche classique
- Simulations
 - Analyse qualitative
 - Analyse quantitative

Illustrations

- Comparaison avec approche classique
- Simulations
 - Analyse qualitative
 - Analyse quantitative
- Application aux données CETIM
 - Position du problème
 - Résultats

Comparaison - Approche classique

Jeux de données	Méthodes	
	C4.5	BDT
Iris	0.056	0.048
Diabetes	0.253	0.271
Glass	0.343	0.331
Wine	0.120	0.086
Thyroid	0.109	0.056
Ionosphere	0.182	0.138

Simulations

- \mathcal{L} : Ensemble d'apprentissage
 - 3 classes contenant 50 vecteurs bidimensionnels

Simulations

- \mathcal{L} : Ensemble d'apprentissage
 - 3 classes contenant 50 vecteurs bidimensionnels
 - Un vecteur gaussien \mathbf{z} : $f(\mathbf{z}|\omega_q) \sim \mathcal{N}(\mu_q, \Sigma_q)$

Simulations

- \mathcal{L} : Ensemble d'apprentissage
 - 3 classes contenant 50 vecteurs bidimensionnels
 - Un vecteur gaussien \mathbf{z} : $f(\mathbf{z}|\omega_q) \sim \mathcal{N}(\mu_q, \Sigma_q)$
 - Suivie d'une transformation non-linéaire

$$\mathbf{z} \mapsto \mathbf{x} = \exp(0.3 \mathbf{z})$$

$$\mu_1 = (-1, -1)', \mu_2 = (1, 2)', \mu_3 = (-1.5, 2)'$$

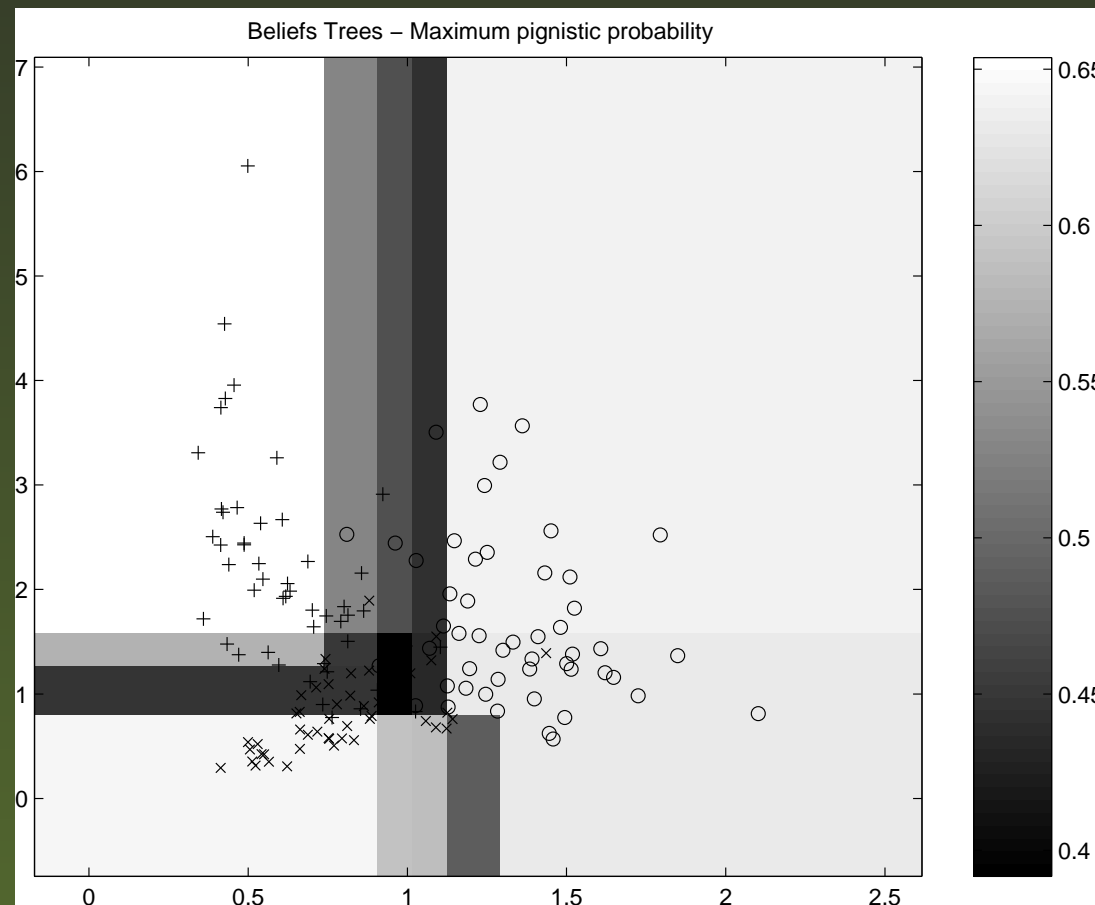
$$\Sigma_q = D_q A D_q' \text{ with}$$

$$A = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{3}/3 \end{pmatrix} \quad D_q = \begin{pmatrix} \cos \theta_q & -\sin \theta_q \\ \sin \theta_q & \cos \theta_q \end{pmatrix}$$

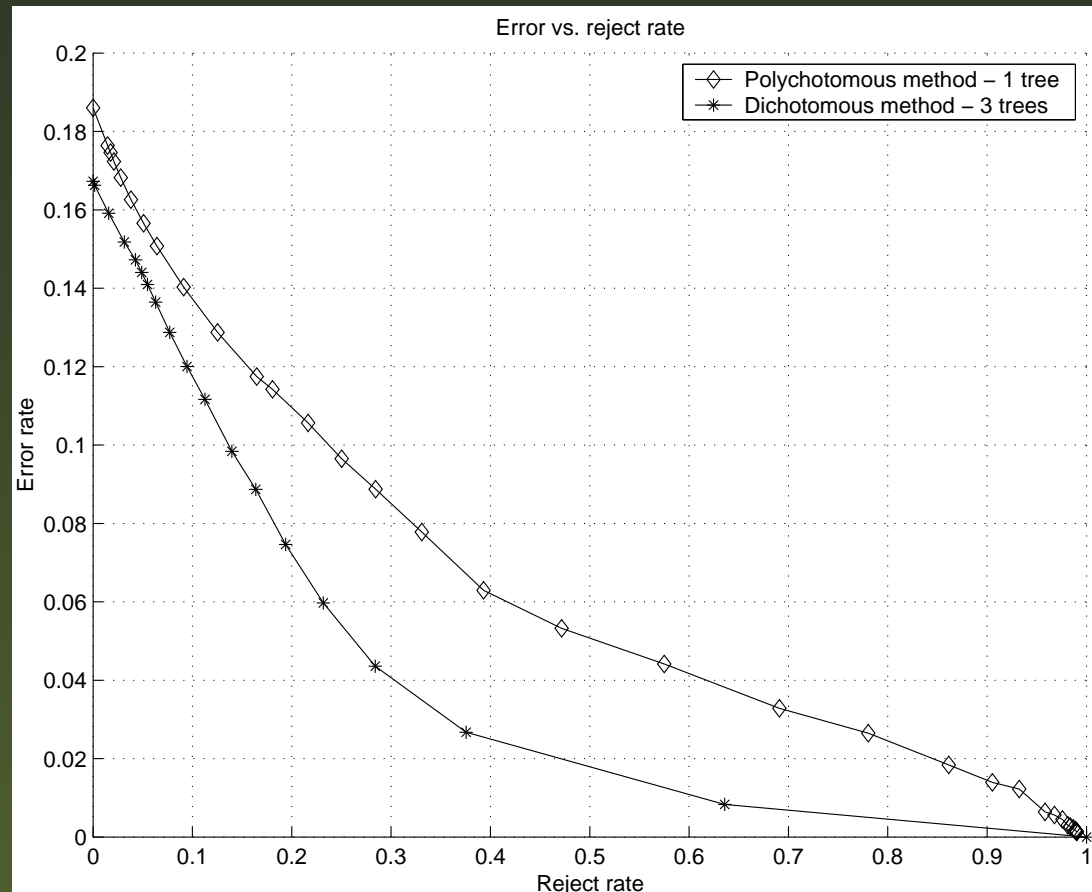
$$\text{and } \theta_1 = \pi/3, \theta_2 = \pi/2, \theta_3 = -\pi/3.$$

Simulations : Analyse qualitative

- Simulation : Etiquetage ensembliste



Simulations : Analyse quantitative



Application CETIM

- Centre d'Etudes Techniques des Industries Mécaniques

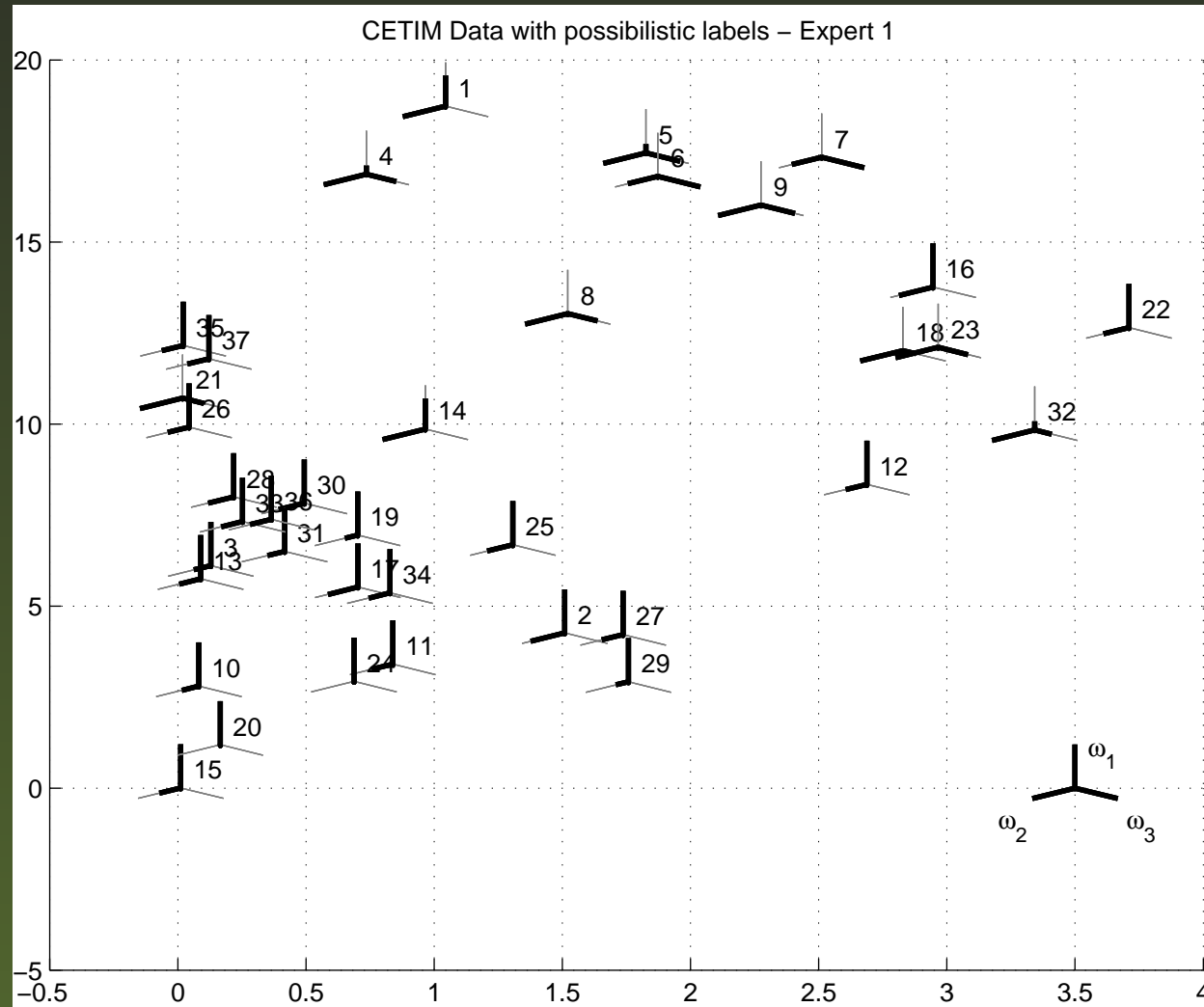
Application CETIM

- Centre d'Etudes Techniques des Industries Mécaniques
- Contrôle qualité d'appareils sous pression
 - $K = 3$ classes
 - $p = 27$ variables
 - $n = 37$ échantillons

Application CETIM

- Centre d'Etudes Techniques des Industries Mécaniques
- Contrôle qualité d'appareils sous pression
 - $K = 3$ classes
 - $p = 27$ variables
 - $n = 37$ échantillons
- Etiquetage possibiliste
 - 2 experts
 - Combinaison

Eiquetage possibiliste - Expert 1



Application CETIM - Expérience 1

- Apprentissage avec labels certains

Data set	E	C_1	C_2
expert 1	0.32	0.38	0.58
expert 2	0.35	0.40	0.59
1+2 (labels)	0.32	0.43	0.59
1+2 (decisions)	0.33	0.42	0.56

Application CETIM - Expérience 1

- Apprentissage avec labels certains

Data set	E	C_1	C_2
expert 1	0.32	0.38	0.58
expert 2	0.35	0.40	0.59
1+2 (labels)	0.32	0.43	0.59
1+2 (decisions)	0.33	0.42	0.56

- \Rightarrow Fusionner les informations des experts

Application CETIM - Expérience 2

- Apprentissage avec labels possibilistes

Data set	E	C_1	C_2
expert 1	0.32	0.39	0.59
expert 2	0.29	0.41	0.60
1+2 (labels)	0.29	0.45	0.59
1+2 (decisions)	0.30	0.37	0.57

Application CETIM - Expérience 2

- Apprentissage avec labels possibilistes

Data set	E	C_1	C_2
expert 1	0.32	0.39	0.59
expert 2	0.29	0.41	0.60
1+2 (labels)	0.29	0.45	0.59
1+2 (decisions)	0.30	0.37	0.57

- \Rightarrow Apprentissage en étiquetage possibiliste

Conclusions

- Méthode d'induction d'arbres basée sur les fonctions de croyance

Conclusions

- Méthode d'induction d'arbres basée sur les fonctions de croyance
 - Avantages et inconvénients des arbres de décision
 - Procédure d'élagage intégrée
 - Différents types d'étiquetage

Conclusions

- Méthode d'induction d'arbres basée sur les fonctions de croyance
 - Avantages et inconvénients des arbres de décision
 - Procédure d'élagage intégrée
 - Différents types d'étiquetage
- Evaluation
 - Taux d'erreur comparables
 - Modéliser l'apprentissage au niveau étiquetage
 - Fusionner les informations d'experts

Travaux futurs

- Comparaison avec d'autres approches (+ variables qualitatives)

Travaux futurs

- Comparaison avec d'autres approches (+ variables qualitatives)
- Codes correcteurs d'erreur

Travaux futurs

- Comparaison avec d'autres approches (+ variables qualitatives)
- Codes correcteurs d'erreur
- Gestion des variables manquantes

Travaux futurs

- Comparaison avec d'autres approches (+ variables qualitatives)
- Codes correcteurs d'erreur
- Gestion des variables manquantes
- Versions "Baggées" et "Randomisées"
 - Opérateurs de combinaison d'évidence \neq moyenne
 - Réglage des hyper-paramètres
 - Interprétation sous forme d'arbre agrégé
 - Mixer Bagging et Randomization

Travaux futurs

- Comparaison avec d'autres approches (+ variables qualitatives)
- Codes correcteurs d'erreur
- Gestion des variables manquantes
- Versions "Baggées" et "Randomisées"
 - Opérateurs de combinaison d'évidence \neq moyenne
 - Réglage des hyper-paramètres
 - Interprétation sous forme d'arbre agrégé
 - Mixer Bagging et Randomization
- Pages WEB : Références et Programmes
 - <http://www.hds.utc.fr/~pvannoo> ou [tdenoeux](http://www.hds.utc.fr/~tdenoeux)