

Diagnostic par Reconnaissance des Formes

Julien Marzat

Onera/DPRS



retour sur innovation

L'ONERA

Le Diagnostic des systèmes

Principe

Reconnaissance des formes

Principales Méthodes

Discrimination Paramétrique - cadre bayésien

Discrimination Paramétrique non bayésienne

Discrimination Non Paramétrique

Compléments

Discrimination Multi-classes

Adaptation du système de décision

Exemple

L'ONERA

Le Diagnostic des systèmes

Principe

Reconnaissance des formes

Principales Méthodes

Discrimination Paramétrique - cadre bayésien

Discrimination Paramétrique non bayésienne

Discrimination Non Paramétrique

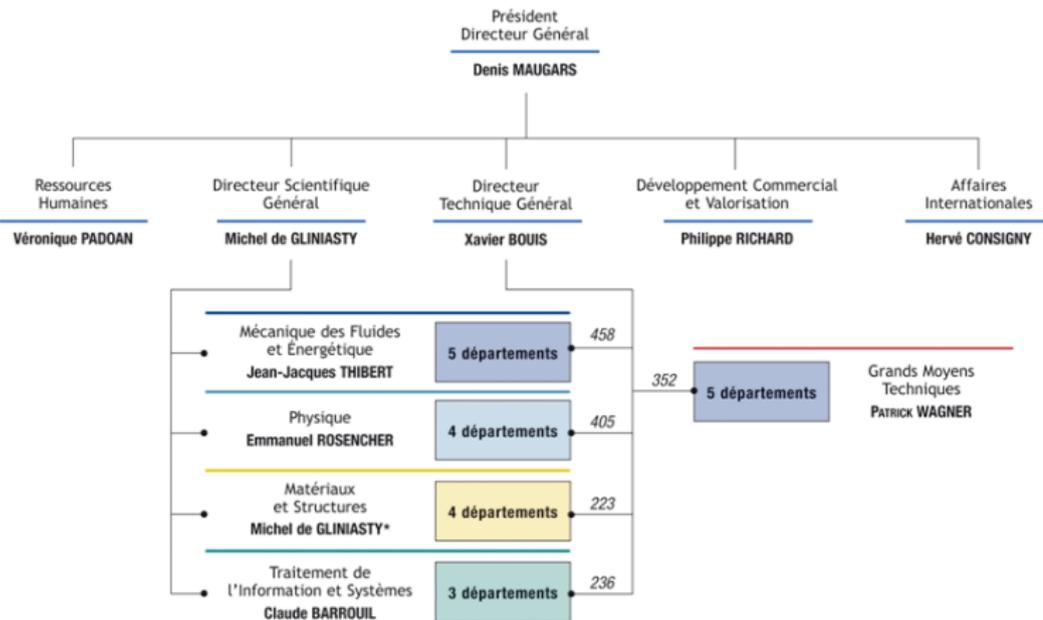
Compléments

Discrimination Multi-classes

Adaptation du système de décision

Exemple

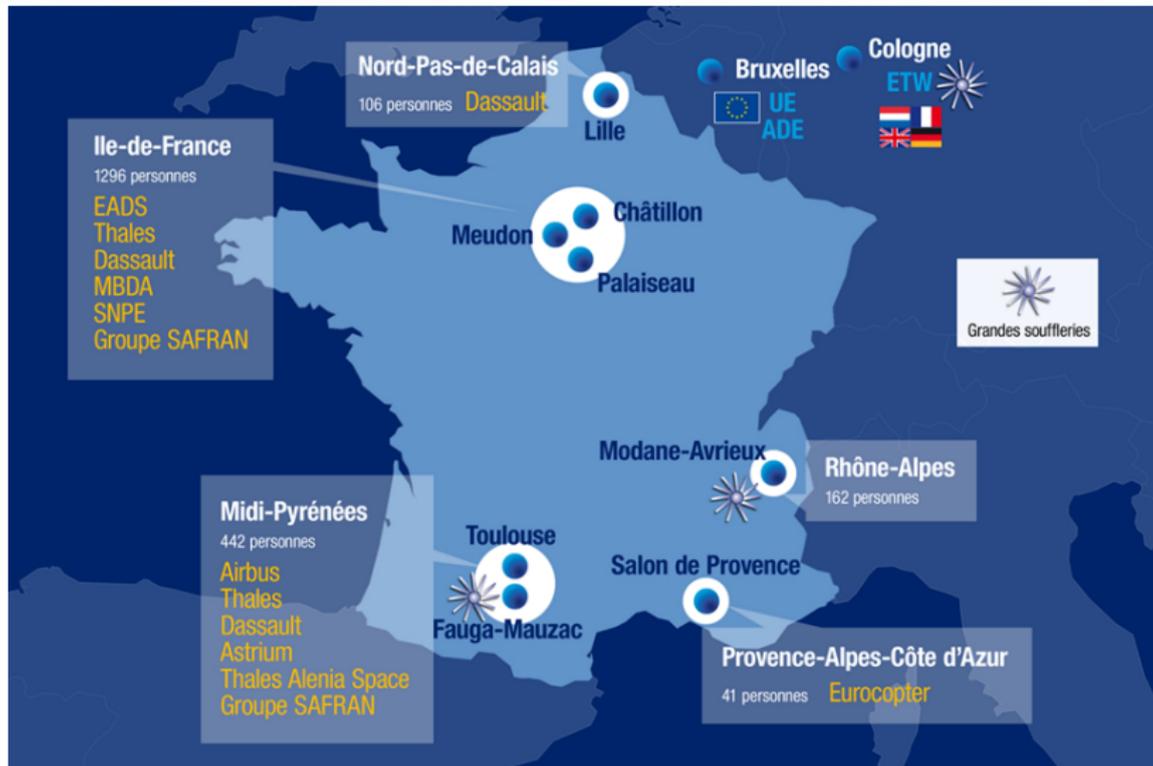
ONERA - organisation



*par intérim

- *Matériaux et structures - MAS* : Aéroélasticité, Dynamique des structures, Matériaux, Microstructures
- *Mécanique des fluides et énergétique - MFE* : Aérodynamique, Énergétique, Aéroacoustique
- *Physique - PHY* : Électromagnétisme, radar, Environnement spatial, Mesures physiques, Optique
- *Traitement de l'information et systèmes - TIS* : Commande des systèmes, Traitement de l'information et modélisation
- ⇒ <http://www.onera.fr/onera/departements>
- Chaque année : 340 stages Master 2, 80 nouveaux doctorants

ONERA - Implantations



L'ONERA

Le Diagnostic des systèmes

Principe

Reconnaissance des formes

Principales Méthodes

Discrimination Paramétrique - cadre bayésien

Discrimination Paramétrique non bayésienne

Discrimination Non Paramétrique

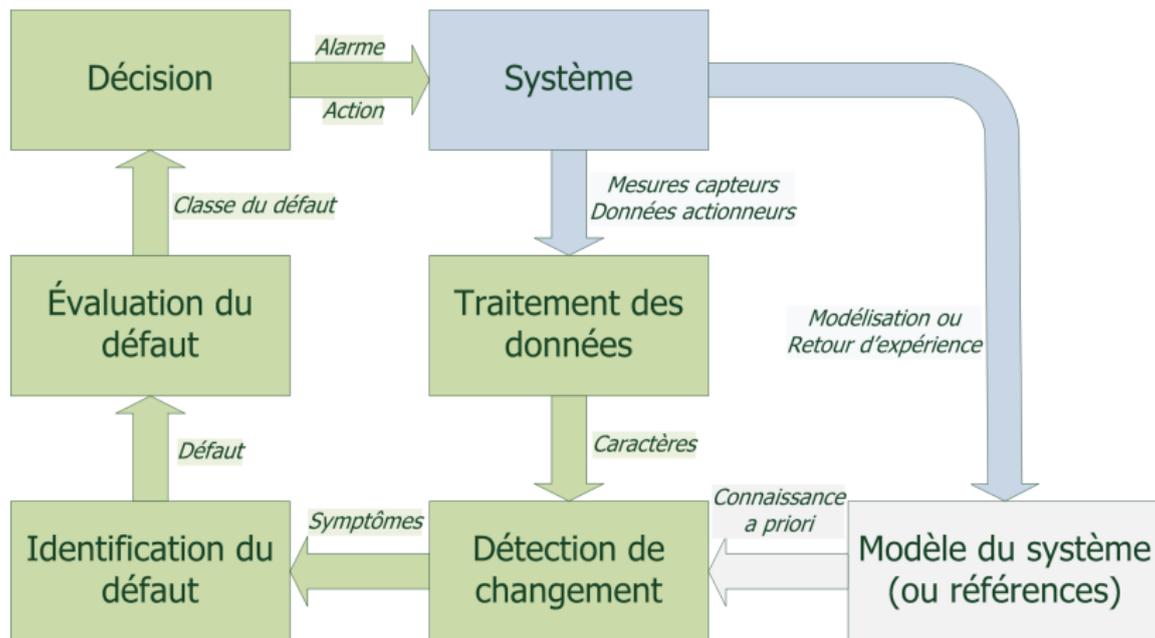
Compléments

Discrimination Multi-classes

Adaptation du système de décision

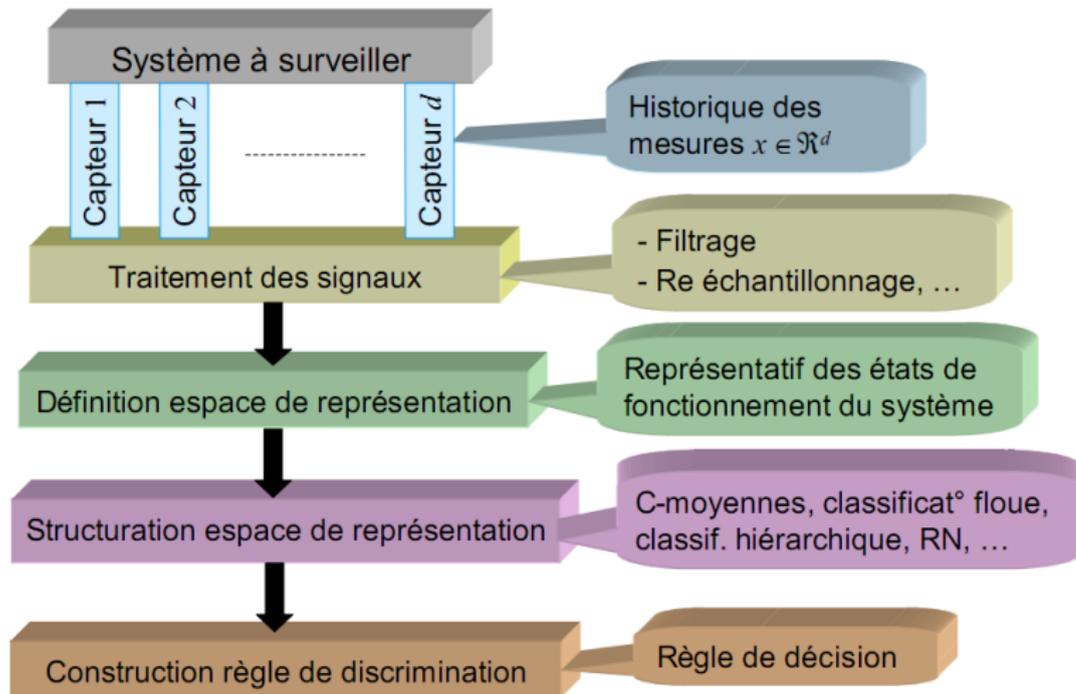
Exemple

Principe du Diagnostic

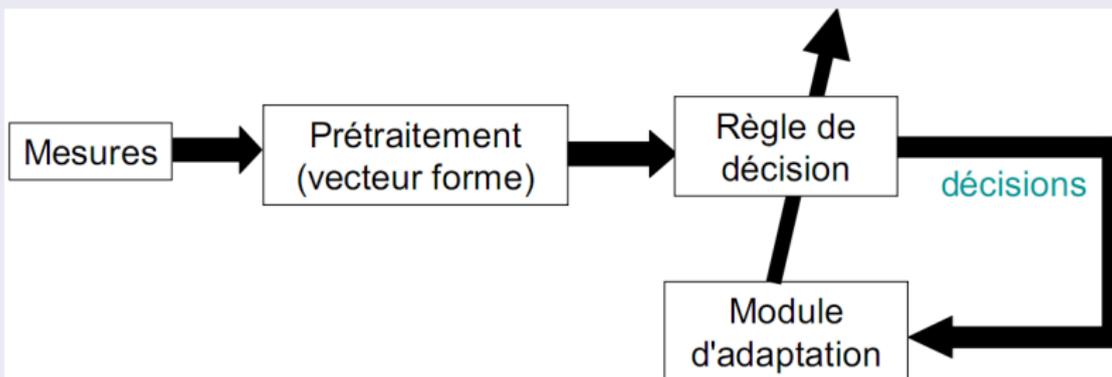


Diagnostic et RdF

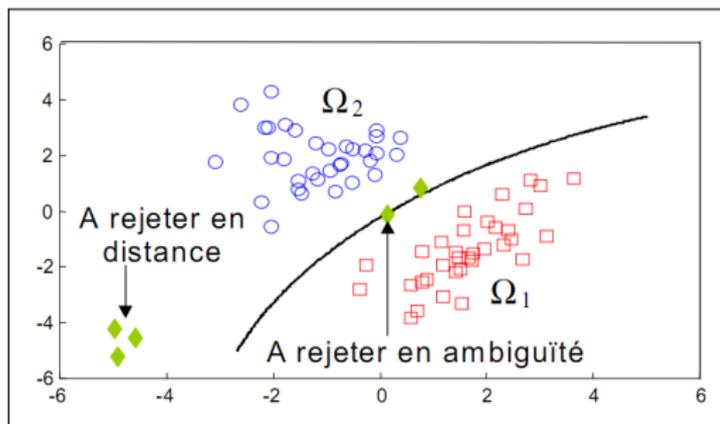
Etapes Hors-ligne



Module de diagnostic temps réel



En diagnostic, les classes connues correspondent aux modes de fonctionnement. Si un point est déclaré « trop loin » des classes existantes ou qu'il y a incertitude entre deux classes, il est rejeté (en distance ou en ambiguïté).



Réduction de la dimension

- Analyse en Composantes Principales (PCA) : projection sur les k premiers vecteurs propres
- Analyse en Composantes Indépendantes (ICA), Analyse en Composantes Curvilignes

Structuration de l'ensemble en classes : Coalescence, un algorithme

- 1 Choisir *a priori* M centres de classes
- 2 Créer une partition en affectant chaque point au centre de la classe la plus proche (selon une distance)
- 3 Calculer le centre de gravité de chaque classe
- 4 Répéter 2. et 3. jusqu'à extremum pour un critère défini
- 5 Revoir le nombre de classes en regroupant les plus petites

Réduction de la dimension

- Analyse en Composantes Principales (PCA) : projection sur les k premiers vecteurs propres
- Analyse en Composantes Indépendantes (ICA), Analyse en Composantes Curvilignes

Structuration de l'ensemble en classes : Coalescence, un algorithme

- 1 Choisir *a priori* M centres de classes
- 2 Créer une partition en affectant chaque point au centre de la classe la plus proche (selon une distance)
- 3 Calculer le centre de gravité de chaque classe
- 4 Répéter 2. et 3. jusqu'à extremum pour un critère défini
- 5 Revoir le nombre de classes en regroupant les plus petites

L'ONERA

Le Diagnostic des systèmes

Principe

Reconnaissance des formes

Principales Méthodes

Discrimination Paramétrique - cadre bayésien

Discrimination Paramétrique non bayésienne

Discrimination Non Paramétrique

Compléments

Discrimination Multi-classes

Adaptation du système de décision

Exemple

Hypothèses

- M classes $\Omega_i, i = 1..M$
- Les probabilités *a priori* $P(\Omega_i)$ sont connues
- La loi multidimensionnelle de \mathbf{x} , $P(\mathbf{x} | \Omega_i)$ est connue

Ce que l'on veut

$$\text{Règle de décision : } d := \begin{cases} d(\mathbf{x}) = i, & i = 1..M, \text{ classe } i \\ d(\mathbf{x}) = 0, & \text{rejet distance} \\ d(\mathbf{x}) = -1, & \text{rejet ambiguïté} \end{cases}$$

Hypothèses

- M classes $\Omega_i, i = 1..M$
- Les probabilités *a priori* $P(\Omega_i)$ sont connues
- La loi multidimensionnelle de \mathbf{x} , $P(\mathbf{x} | \Omega_i)$ est connue

Ce que l'on veut

$$\text{Règle de décision : } d := \begin{cases} d(\mathbf{x}) = i, & i = 1..M, \text{ classe } i \\ d(\mathbf{x}) = 0, & \text{rejet distance} \\ d(\mathbf{x}) = -1, & \text{rejet ambiguïté} \end{cases}$$

Les outils dont on dispose

- Probabilité *a posteriori* (Bayes) :
$$P(\Omega_i | \mathbf{x}) = \frac{P(\mathbf{x}|\Omega_i) \cdot P(\Omega_i)}{\sum_{i=1}^M P(\mathbf{x}|\Omega_i) \cdot P(\Omega_i)}$$

- Coût $C := \begin{cases} C(i, \Omega_j) = 1, & \text{si } i \neq j \\ C(i, \Omega_j) = 0, & \text{si } i = j \\ C(-1, \Omega_j) = a, & j = 0..M \end{cases} \quad i, j = 0..M$

- Risque de Bayes :
$$R(\mathbf{x}) = \sum_{j=0}^M C(d(\mathbf{x}), \Omega_j) \cdot P(\Omega_j | \mathbf{x})$$

Règle de décision

$$P(\Omega_i | \mathbf{x}) = \max_j P(\Omega_j | \mathbf{x}) \text{ et } \begin{cases} d(\mathbf{x}) = i & \text{si } P(\Omega_i | \mathbf{x}) \geq 1 - a \\ d(\mathbf{x}) = -1 & \text{si } P(\Omega_i | \mathbf{x}) < 1 - a \end{cases}$$

Les outils dont on dispose

- Probabilité *a posteriori* (Bayes) :
$$P(\Omega_i | \mathbf{x}) = \frac{P(\mathbf{x}|\Omega_i) \cdot P(\Omega_i)}{\sum_{i=1}^M P(\mathbf{x}|\Omega_i) \cdot P(\Omega_i)}$$

- Coût $C := \begin{cases} C(i, \Omega_j) = 1, & \text{si } i \neq j \\ C(i, \Omega_j) = 0, & \text{si } i = j \\ C(-1, \Omega_j) = a, & j = 0..M \end{cases} \quad i, j = 0..M$

- Risque de Bayes :
$$R(\mathbf{x}) = \sum_{j=0}^M C(d(\mathbf{x}), \Omega_j) \cdot P(\Omega_j | \mathbf{x})$$

Règle de décision

$$P(\Omega_i | \mathbf{x}) = \max_j P(\Omega_j | \mathbf{x}) \text{ et } \begin{cases} d(\mathbf{x}) = i & \text{si } P(\Omega_i | \mathbf{x}) \geq 1 - a \\ d(\mathbf{x}) = -1 & \text{si } P(\Omega_i | \mathbf{x}) < 1 - a \end{cases}$$

En pratique

- On estime (maximum de vraisemblance...) les densités de probabilités à partir de l'historique et des avis d'expert.
Inconvénient : hypothèses sur la loi (gaussienne...)
- Valeur de la constante a , Contrainte : $0 \leq a \leq \frac{M-1}{M}$...
- Méthode puissante si l'on a une très bonne connaissance du processus et un système *statique* (typiquement, un processus industriel)

Enoncé du problème : Classification linéaire binaire

- Soit $y = [x_1, \dots, x_d, 1]^T$, et Ω_1, Ω_2 les deux classes du système
- Trouver v tel que $v^T \cdot z \geq 0$ avec $z = \begin{cases} y & \text{si } x \in \Omega_1 \\ -y & \text{si } x \in \Omega_2 \end{cases}$

Détermination de v

- approche gradient : $v_{k+1} = v_k - a_k \nabla J(v_k) \Rightarrow$ choisir a_k et J
- Perceptron : $J_p = \sum_{z \in Z} -v^T \cdot z$, avec Z l'ensemble des vecteurs « mal classés » $Z : \{v^T \cdot z < 0\}$
- Variantes : relaxation, coût quadratique, Ho-Kashyap

Enoncé du problème : Classification linéaire binaire

- Soit $y = [x_1, \dots, x_d, 1]^T$, et Ω_1, Ω_2 les deux classes du système
- Trouver v tel que $v^T \cdot z \geq 0$ avec $z = \begin{cases} y & \text{si } x \in \Omega_1 \\ -y & \text{si } x \in \Omega_2 \end{cases}$

Détermination de v

- approche gradient : $v_{k+1} = v_k - a_k \nabla J(v_k) \Rightarrow$ choisir a_k et J
- Perceptron : $J_p = \sum_{z \in Z} -v^T \cdot z$, avec Z l'ensemble des vecteurs « mal classés » $Z : \{v^T \cdot z < 0\}$
- Variantes : relaxation, coût quadratique, Ho-Kashyap

Discrimination paramétrique Non Bayésienne

Séparatrices non linéaires

Séparatrices non linéaires

- Autres fonctions (Fonctions quadratiques,...)
- Les Réseaux de Neurones
- Les Support Vector Machines (SVM)

Réseaux de neurones

- $y = h\left(\sum_{j=1}^d w_j \cdot x_j\right)$ avec $\begin{cases} w_j \text{ poids synaptique} \\ h \text{ fonction d'activation} \end{cases}$

- choix de h : seuil, linéaire, sigmoïde...
- Réglage ? Rétropropagation du gradient - MLP

$$\text{Minimiser } J(w) = \sum_{i=1}^{nN} \|x_{ik}^N - d_{ik}\|^2$$

- *Réglage et choix des paramètres délicats*

Discrimination paramétrique Non Bayésienne

Séparatrices non linéaires

Séparatrices non linéaires

- Autres fonctions (Fonctions quadratiques,...)
- Les Réseaux de Neurones
- Les Support Vector Machines (SVM)

Réseaux de neurones

- $y = h\left(\sum_{j=1}^d w_j \cdot x_j\right)$ avec $\begin{cases} w_j \text{ poids synaptique} \\ h \text{ fonction d'activation} \end{cases}$

- choix de h : seuil, linéaire, sigmoïde...
- Réglage ? Rétropropagation du gradient - MLP

$$\text{Minimiser } J(w) = \sum_{i=1}^{nN} \|x_{ik}^N - d_{ik}\|^2$$

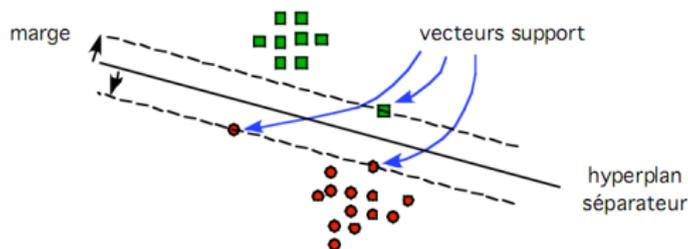
- *Réglage et choix des paramètres délicats*

Discrimination paramétrique Non Bayésienne

Support Vector Machines (SVM)

Idée : Séparatrice linéaire entre 2 classes qui maximise la *marge* entre l'hyperplan $\{\omega \cdot x + b = 0\}$ et les individus des deux classes.

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 \\ \forall i, y_i(\omega \cdot x_i + b) \geq 1 \end{cases} \quad \text{avec } y_i = \begin{cases} -1 & \text{si } x_i \in \Omega_0 \\ 1 & \text{si } x_i \in \Omega_1 \end{cases}$$

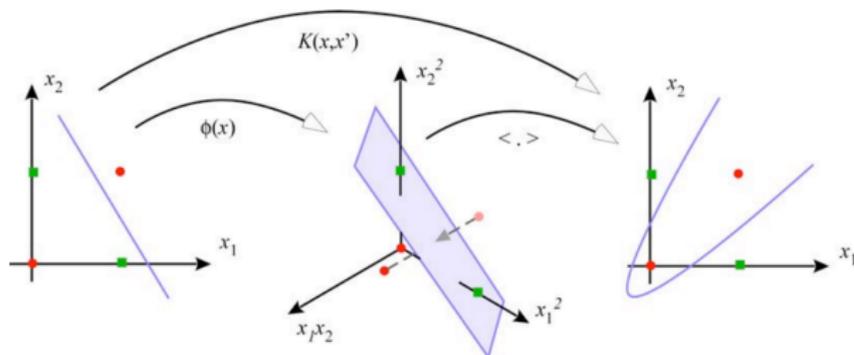


Résolution du problème dual \Rightarrow expression de ω^* sous forme d'une pondération des points « les plus proches » de l'hyperplan solution : les vecteurs de support.

Discrimination paramétrique Non Bayésienne

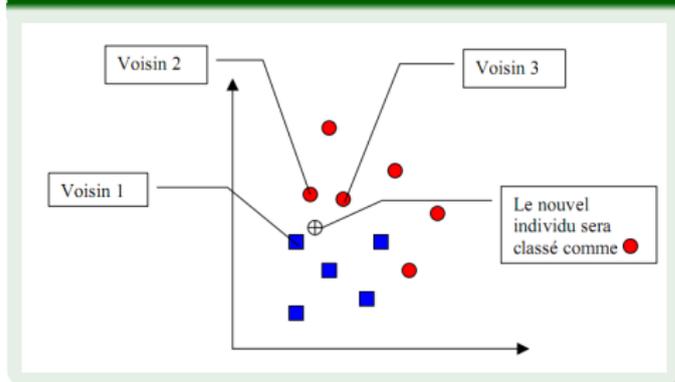
Support Vector Machines (SVM)

- Traitement des erreurs : variables 'ressort' :
$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + C \cdot \sum_{i=1}^n \xi_i \\ \forall i, y_i(\omega \cdot x + b) \geq 1 - \xi_i \end{cases}$$
- Séparation non linéaire : utilisation du « kernel trick » et recherche d'une séparatrice linéaire dans le nouvel espace des données. Fonction noyau K (polynôme, RBF, MLP, splines...)



- Paramètres à régler : noyau, constantes 'ressort'

Règle des 3-ppv



- *Rejet d'ambiguïté* : il faut plus que la majorité pour décider : seuil à fixer pour chaque classe
- *Rejet de distance* : si la distance moyenne aux k-ppv est supérieure à un seuil T fixé, le point est rejeté.
- *Réglages* : nb de voisins, choix de la distance, seuils de décision et de rejet

- Analyse en composantes principales : tests statistiques (T^2 , SPE...) dans l'espace réduit et l'espace résiduel
- Processus Gaussiens / Krigeage : méthodes à noyaux
- Détection de changement statistiques dans les signaux et systèmes (à base de modèles) : cf Basseville et Nikiforov, *Detection of Abrupt Changes : Theory and Application*, Prentice-Hall. 1993.

L'ONERA

Le Diagnostic des systèmes

Principe

Reconnaissance des formes

Principales Méthodes

Discrimination Paramétrique - cadre bayésien

Discrimination Paramétrique non bayésienne

Discrimination Non Paramétrique

Compléments

Discrimination Multi-classes

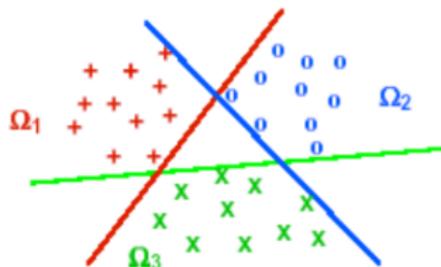
Adaptation du système de décision

Exemple

Discrimination Multi-Classes

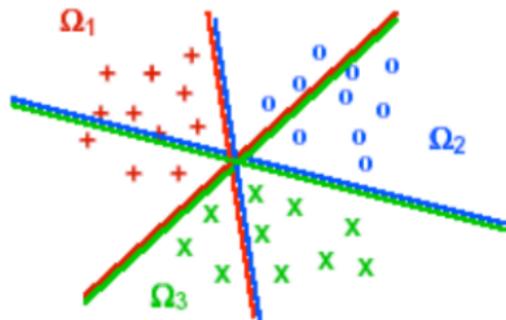
Deux possibilités

1 parmi M



M sous-classifieurs

2 à 2



$M(M-1)$ sous-classifieurs

Adaptation des classes

- En cas de variation (paramètres de la loi / CdG des classes), on redéfinit d'une part les contours des classes existantes et on applique à nouveau une méthode de *coalescence* aux points rejetés pour déterminer l'apparition ou non de nouvelles classes.

Apport de la logique floue

- Attribution à chaque nouveau point d'un degré d'appartenance ($0 \leq \mu_i \leq 1$) à chaque classe Ω_i
- Analyse des points rejetés : création d'une nouvelle classe et analyse des degrés d'appartenance des points à celle-ci puis en cas de rejet création d'une classe supplémentaire etc...

Adaptation des classes

- En cas de variation (paramètres de la loi / CdG des classes), on redéfinit d'une part les contours des classes existantes et on applique à nouveau une méthode de *coalescence* aux points rejetés pour déterminer l'apparition ou non de nouvelles classes.

Apport de la logique floue

- Attribution à chaque nouveau point d'un degré d'appartenance ($0 \leq \mu_i \leq 1$) à chaque classe Ω_i
- Analyse des points rejetés : création d'une nouvelle classe et analyse des degrés d'appartenance des points à celle-ci puis en cas de rejet création d'une classe supplémentaire etc...

Exemple

« Diagnostic de défaillances par Reconnaissance des Formes », R.Bensaadi et H.Mouss

Données

- Trois variables : $\begin{cases} x_1 & \text{pression au point } A_1 \\ x_2 & \text{pression au point } A_2 \\ x_3 & \text{frequence bruit sonore} \end{cases}$
- Un mode nominal, deux modes de défaillances connues
- 100 éléments dans l'échantillon d'apprentissage

Méthode

- Réduction de la dimension par ACP (dim 3 \Rightarrow dim 2)
- Séparatrice non linéaire : Réseau de neurones (RBF)
- coalescence floue : fonctions d'appartenance à chaque classe
- Détection de changement de classe par test statistique sur les fonctions d'appartenance

Exemple

« Diagnostic de défaillances par Reconnaissance des Formes », R.Bensaadi et H.Mouss

Données

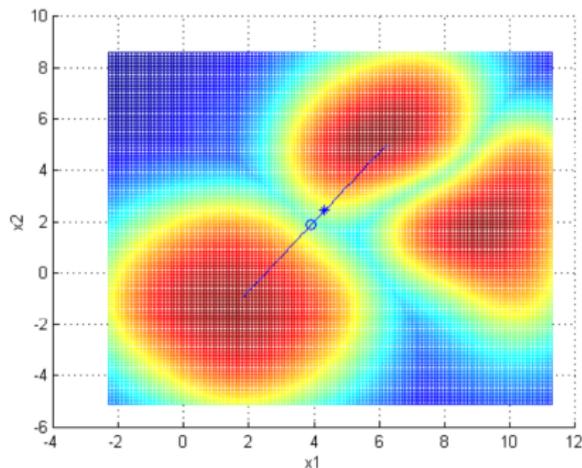
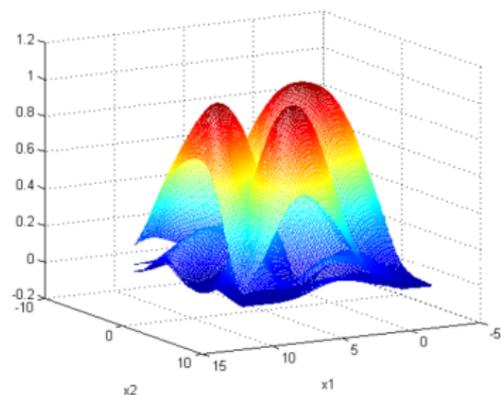
- Trois variables : $\begin{cases} x_1 & \text{pression au point } A_1 \\ x_2 & \text{pression au point } A_2 \\ x_3 & \text{frequence bruit sonore} \end{cases}$
- Un mode nominal, deux modes de défaillances connues
- 100 éléments dans l'échantillon d'apprentissage

Méthode

- Réduction de la dimension par ACP (dim 3 \Rightarrow dim 2)
- Séparatrice non linéaire : Réseau de neurones (RBF)
- coalescence floue : fonctions d'appartenance à chaque classe
- Détection de changement de classe par test statistique sur les fonctions d'appartenance

Exemple

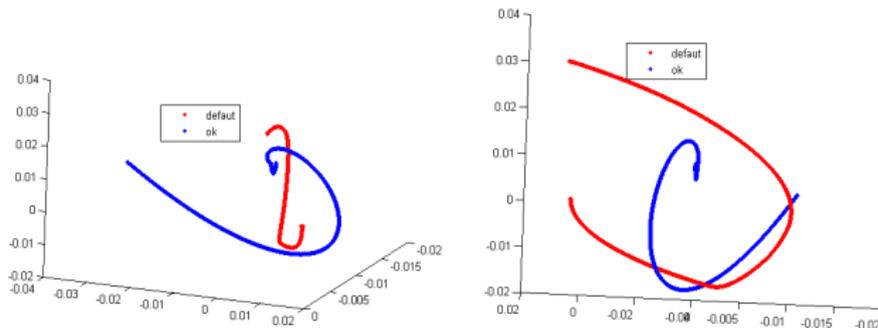
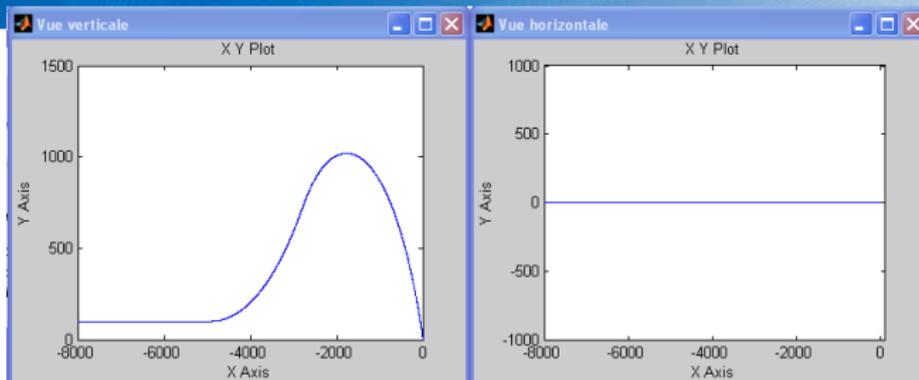
Surveillance d'un procédé industriel fictif - résultats



Fonctions d'appartenance ——— ACP dimension 2 et résultat

« Méthodologie pour la sécurisation en vol d'un aéronef sans ajout de capteur »

- Cas d'étude : missile de croisière
- Variables : accélération, vitesse, position, vitesse angulaire, angles d'attitude.
- Commandes : gouvernes, taux de poussée
- Défauts pouvant intervenir : perte de poussée, blocage gouverne, perte de mesures inertielles, paramètres statistiques capteurs erronés
- Exemple pour la reconnaissance des formes : trajectoires après projection ACP en dimension 3 des variables $\{z, v_x, v_z, q, \theta\}$



Projection ACP 3D des trajectoires avec et sans défaut



B.Dubuisson

Diagnostic et Reconnaissance des formes

Hermes, 1990



P.Aknin

De la mesure à sa représentation et sa discrimination.

Application au diagnostic des infrastructures ferroviaires

Habilitation à Diriger des Recherches, 2008



R.Bensaadi, H.Mouss

Diagnostic de défaillances par Reconnaissance de Formes

SETIT 2005