

## CFA/VISHNO 2016

**SVM pour une meilleure classification des données de monitoring par ondes guidées**

M. El Mountassir<sup>a</sup>, S. Yaacoubi<sup>a</sup>, G. Mourot<sup>b</sup> et D. Maquin<sup>b</sup>

<sup>a</sup>Institut de Soudure Association, 4 Bvd Henri Becquerel, Espace Cormontaigne, 57970 Yutz, France

<sup>b</sup>Le CRAN, UMR 7039 - Université de Lorraine, CNRS, 2, Avenue de la forêt de Haye, 54516 Nancy, France  
s.yaacoubi@isgroupe.com



LE MANS

Cet article se place dans le cadre du monitoring des structures et traite principalement la technique des ondes ultrasonores guidées. Il porte sur la classification des données de cette technique par le biais des séparateurs à vaste marge, dont la fiabilité est conditionnée par la sélection ciblée des paramètres du séparateur optimal. Différents algorithmes ont été antérieurement développés et peuvent être trouvés dans la littérature. La présente étude vise à appliquer ces algorithmes sur des données expérimentales de monitoring, résultant de l'application de la technique des ondes ultrasonores guidées, dans l'objectif d'identifier le séparateur le plus pertinent afin de réduire voire annuler complètement les fausses alarmes et fiabiliser ainsi au mieux le monitoring in-situ.

## 1 Introduction

Le monitoring est de plus en plus appelé comme renfort de la maintenance préventive des structures. Le principe de base du monitoring consiste à faire vivre la chaîne de mesure (au moins les capteurs) en permanence avec la structure à surveiller [1]. Ceux-ci y étant intrus ou attachés, offrent la possibilité d'interroger ladite structure plus fréquemment et à moindre coût comparé au contrôle non destructif classique (CND). Les données enregistrées sont alors comparées entre elles pour statuer sur l'intégrité d'une telle structure. L'évolution dans les données, même minime, peut être corrélée avec l'apparition d'un éventuel défaut, ce qui permet d'obtenir une meilleure sensibilité par rapport au CND. Malheureusement, les données peuvent évoluer à cause de changements dans les conditions opérationnelles et environnementales (COE) [2]. Ceci est susceptible de générer des fausses alarmes pouvant parfois engendrer des pertes économiques.

Des techniques statistiques peuvent être alors utilisées pour pallier ce problème [3], à l'instar des séparateurs à vaste marge (connue par son acronyme anglais SVM). Cette méthode vise à classer les données en deux lots via un séparateur dont la sélection est décisive dans le succès d'une telle classification.

Dans un contexte de monitoring, la séparation optimale entre les données de l'état sain et de l'état endommagé va permettre une meilleure gestion des fausses alarmes et des cas de non détection de défaut. Bien que cette technique ait été déjà appliquée dans ce domaine, la plupart des travaux issus de la littérature se sont focalisés sur le résultat final de classification, ainsi, le choix des paramètres optimaux du séparateur n'a pas été rigoureusement justifié. Par exemple, Hassan et al [4] ont appliqués la méthode SVM pour la classification des défauts des pipelines. Ils ont étudié l'influence du bruit sur le résultat de classification pour différents types de filtres. En revanche, aucune information sur les paramètres du classifieur n'a été divulguée. Zamani et al [5] ont procédé à une étude exhaustive qui fait intervenir non seulement les paramètres des SVMs mais aussi les caractéristiques des signaux acquis. Le résultat obtenu ne peut pas être qualifié d'optimal car certaines paramètres ont été choisis arbitrairement. De plus, d'autres facteurs d'influence n'ont pas été explorés dans cette étude. Finalement, des exemples d'applications des SVM sur des données de monitoring par ondes guidées peuvent être trouvés dans les références [6,7].

Cet article est constitué de 3 sections dont la deuxième est consacrée à un rappel concis de la technique des SVMs en mettant l'accent sur les séparateurs non-linéaires. La troisième, quant à elle, porte sur l'application de ces séparateurs sur des données expérimentales obtenues dans des conditions très proches de l'in-situ. Les résultats obtenus dont le but est de les comparer afin d'en sélectionner le plus optimal y seront discutés. Les

conclusions tirées aussi bien que les perspectives de cette étude comparative feront l'objet de la dernière section.

## 2 SVM : rappels théoriques

La méthode des SVMs fait partie des algorithmes d'apprentissage supervisé. Elle permet de résoudre le problème de la reconnaissance de formes. Elle a été mise au point par Vapnick en 1993 [8] et consiste à trouver un séparateur optimal qui maximise la marge entre deux classes de données, en utilisant un ensemble limité de séquences d'apprentissage. Pour pouvoir adapter cette méthode à la présente étude de monitoring, deux classes de données vont être considérées : une classe qui représente la structure saine (état de référence) et une autre qui caractérise la structure endommagée. La méthode des SVMs doit donc être en mesure de distinguer si le signal mesuré provient d'une structure saine ou endommagée. Afin de démystifier le fonctionnement de cette méthode, on va procéder à une description mathématique.

Soit  $A$  un ensemble composé de  $n$  paires de données/classes d'apprentissage, donné par :

$$A = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (1)$$

où  $y_i \in \{-1, 1\}$  est le label d'appartenance d'une observation à une classe, le nombre  $n$  désigne la dimension de la base de données comme il sera plus détaillé ultérieurement et  $x_i$ , le  $i^{\text{ème}}$  échantillon de cette base.

Chaque échantillon  $x_i$  possède  $p$  variables descriptives. Le vecteur  $x_i$ , peut s'exprimer comme suit :

$$x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^t \quad (2)$$

où  $t$  est le signe « transposé ».

Dans la pratique, pour pouvoir appliquer la méthode SVM, il faut tout d'abord sélectionner les paramètres descriptifs (par exemple le RMS et la variance VAR du signal). Il est à noter que le choix de ces paramètres est déterminant dans la classification des données d'ondes guidées. La deuxième étape a pour but de trouver un hyperplan optimal permettant de séparer les données d'apprentissage de sorte que tous les points d'une même classe soient du même côté de l'hyperplan. Cet hyperplan divise donc, d'un point de vue géométrique, l'espace en deux. La recherche de l'hyperplan optimal se base sur le critère de maximisation des marges, c'est-à-dire les distances entre les individus des classes d'apprentissage et l'hyperplan. Les points de distance minimale sont qualifiés de vecteurs supports. Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. Dans la plupart des problèmes réels, les classes ne sont pas linéairement séparables. Pour surmonter cet inconvénient, l'idée des SVM est d'appliquer une transformation non linéaire  $\phi$  aux données pour permettre une séparation linéaire des exemples dans

ce nouvel espace (Figure 1). En conséquence, il va y avoir un changement de dimension. Finalement, pour laisser plus de liberté au classifieur d'identifier correctement un échantillon même s'il se trouve au mauvais côté de l'hyperplan (classes non séparables), le problème d'optimisation s'écrit sous la forme suivante [9]:

$$\min_{w,b,\varepsilon} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \varepsilon_i \quad (3)$$

sous les contraintes :

$$y_i(w^t \varphi(x_i) + b) \geq 1 - \varepsilon_i \\ \varepsilon_i \geq 0$$

où  $w$  et  $b$  sont des paramètres de l'hyperplan,  $C$  est le poids donné aux échantillons se trouvant du mauvais côté de la frontière de séparation (appelé aussi contrainte de régularisation),  $\varepsilon_i$  sont des paramètres qui permettent de considérer des points mal classés.

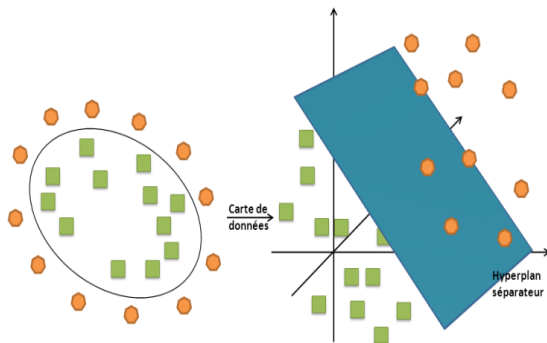


Figure 1: Changement de dimension de l'espace à fin de trouver le plan séparateur des deux exemples.

Grâce à l'astuce du noyau, l'équation 3 prend la forme duale suivante [9] :

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (4)$$

sous les contraintes :

$$0 \leq \alpha_i \leq C \\ \sum_{i=1}^n \alpha_i y_i = 0$$

où  $\alpha_i$  sont les multiplicateurs de Lagrange et  $K(.,.)$  représente la fonction noyau.

En pratique, quelques familles de fonctions noyau paramétrables sont couramment utilisées et il revient à l'utilisateur d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. Dans la littérature [10], on trouve les fonctions noyaux suivantes :

1. Linéaire.
2. Polynomiale
3. Quadratique
4. Perceptrons multicouches.
5. Fonction à base radiale (RBF)

La solution optimale à ce problème permet de déterminer la fonction de décision donnée à l'équation 5. Cette fonction est nécessaire à la classification d'un nouvel échantillon :

$$f(x) = \text{signe}(\sum_{i=1}^M \alpha_i y_i K(x_i, x) + b) \quad (5)$$

où  $x_i$  et  $y_i$  sont respectivement les vecteurs de support et leurs classes d'appartenance,

Dans ce qui suit, cette méthode sera appliquée pour la détection de défauts géométriques dans une structure tubulaire, dans un contexte de monitoring.

## 3 Base de données expérimentales, résultats et discussions

### 3.1 Construction de la base de données expérimentale

La procédure qui a été adoptée dans la présente étude pour construire la base de données consiste à :

1. effectuer une acquisition d'un signal expérimental par le biais des ondes ultrasonores guidées se propageant dans un tube, sain au début et puis, possédant un défaut artificiel dont la taille a été augmentée en quatre étapes.
2. polluer ce signal en lui ajoutant un bruit blanc gaussien stochastique 199 fois. La base de données est constituée donc de 200 échantillons pour chacun des 5 cas (sain, 1 défaut, 2 défauts, 3 défauts et 4 défauts). Le défaut a été simulé artificiellement en ajoutant des aimants sur la surface de la structure, comme le montre la Figure 2.



Figure 2: Défaut à extension circumférentielle, effectué en quatre étapes augmentant ainsi sa taille et modifiant sa géométrie

Le bruit a été créé au moyen d'un simulateur de bruit stochastique, générant des valeurs aléatoires suivant une distribution normale avec une moyenne nulle et un écart-type égal à l'unité. La fonction bruit a été multipliée par un facteur atténuateur qui permet de contrôler l'amplitude du bruit.

La Figure 3 montre l'exemple d'un signal type acquis dans cette campagne d'essais. Le 1<sup>er</sup> écho (bleu) correspond au signal d'excitation tandis que le 2<sup>ème</sup> (vert) est dû à la fin du tube. Après plusieurs tests, il s'est avéré plus judicieux de travailler sur une portion du signal et non sur sa totalité. En effet, l'écho d'excitation, qui n'est pas vraiment dû qu'à l'excitation mais également à un bruit électronique varie considérablement d'une acquisition à une autre, comme un bruit aléatoire. L'onde interagit avec le défaut, une partie est réfléchiée et l'autre, transmise ; celles-ci varient en fonction des caractéristiques du défaut. Ainsi, l'écho de fin du tube (qui peut correspondre à une soudure en réalité) est tributaire des caractéristiques du défaut. On peut donc d'ores et déjà conclure que pour la suite de l'étude, ainsi que les futures procédures qui découleront de la poursuite de la présente étude, qu'il est nécessaire de sélectionner la partie utile du signal.

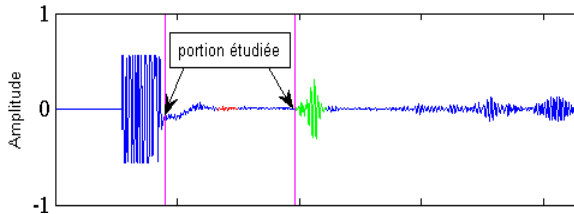


Figure 3: Exemple d'un signal expérimental montrant la portion étudiée : en vert, l'écho de fin du tube qui peut représenter une soudure en réalité et en rouge, l'écho résultant du défaut

Chaque échantillon  $x_i$ , (voir équation 2) peut être caractérisé par un ou plusieurs paramètres ( $x_{ij}$ ). Dans la présente étude, le choix a été porté sur le RMS et la variance, mais d'autres paramètres peuvent également être testés. Ceux-ci vont servir pour le calcul d'indice d'endommagement qui doit renseigner sur l'état et la criticité d'un défaut dans une structure.

L'ensemble des données a été divisé en trois ensembles. Les données de l'ensemble d'apprentissage vont permettre l'identification des caractéristiques du séparateur optimal à savoir les vecteurs-supports et les poids associés. Pour limiter le risque de surapprentissage, le deuxième ensemble sera considéré comme données de validation pour déterminer les paramètres optimaux du noyau utilisé ainsi que la valeur adéquate de la contrainte de régularisation. Le dernier ensemble sera dédié au test du SVM développé.

### 3.2 Résultats et discussions

Afin d'assurer une homogénéité entre les deux classes de données (saine et endommagée), la première classe contient les données correspondantes à l'état de référence (aucun défaut) et la deuxième contient les données relatives aux quatre types de défauts créés auparavant. De plus, les données sont normalisées pour que les valeurs des deux paramètres (RMS et Variance) utilisés soient du même ordre de variation. La Figure 4 illustre les deux classes utilisées dans l'espace des paramètres choisis.

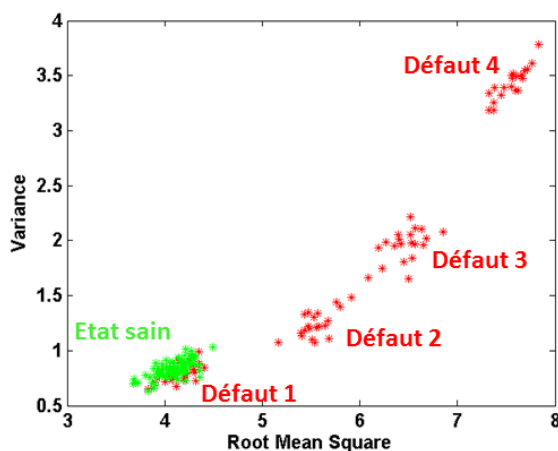


Figure 4: Représentation des données de la structure saine et endommagée dans l'espace des caractéristiques

Une étape préliminaire à la recherche des paramètres du séparateur optimal est de trouver la fonction noyau qui donne le meilleur taux de bonne reconnaissance des différents états de la structure. Dans cette étude, le taux de

classification a été évalué pour chaque type de noyau, le résultat est illustré dans le tableau suivant :

Tableau 1: Taux de classification pour différents noyaux.

Fonction noyau	Taux de classification	
	Etat sain	Défaut 1
Linéaire	<b>100%</b>	<b>0%</b>
Quadratique	<b>100%</b>	<b>0%</b>
Polynomiale	<b>100%</b>	<b>0%</b>
Perceptron	<b>100%</b>	<b>0%</b>
RBF (Sigma=0.2)	<b>95%</b>	<b>32.5%</b>

Le résultat montre qu'aucun échantillon de défaut 1 n'a été correctement classifié pour la plupart des fonctions sauf la RBF qui affiche un taux de classification non nul pour ce type défaut. Il est à noter que les autres types de défauts ont été détectés par toutes les fonctions noyaux (Le taux de classification est de 100 %) et par conséquent, ils ne seront pas considérés dans la suite de l'article. Les résultats obtenus par les autres types de noyaux s'explique par la nature de la frontière de décision qui divise le plan de données en deux comme il peut être montré dans l'exemple de la fonction quadratique illustré dans la Figure 5. Les conséquences directes d'une telle séparation seraient :

- augmentation de la probabilité de non détection, compte tenu de l'énorme espace qu'occupe l'état sain.
- réduction de la sensibilité de système car le défaut le plus petit n'a pas été détecté.

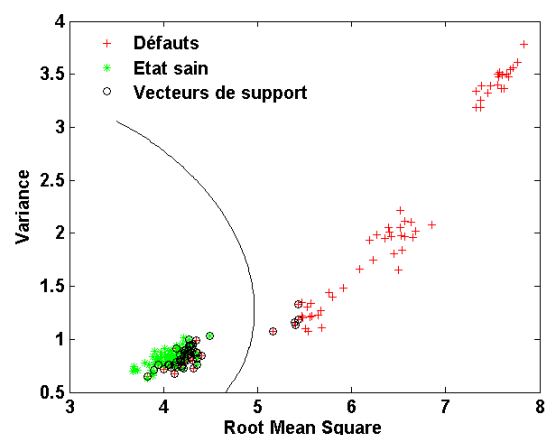


Figure 5: Exemple d'une séparation avec une fonction noyau quadratique

La RBF ne fonctionne pas de la même manière car elle tente de limiter les frontières de décision de l'état sain. Par ailleurs, le résultat de classification est conditionné par le choix de paramètre de réglage ( $\sigma$ ) de ce noyau qui représente la portée de la fonction noyau. Un autre paramètre qui pourrait influencer le résultat final est le paramètre de régularisation ( $C$ ) qui contrôle le pourcentage des échantillons qui peuvent se trouver au mauvais côté du séparateur. En d'autres termes, le taux de classification maximale va dépendre à la fois de  $\sigma$  et  $C$ . Pour bien

expliciter l'influence de la portée de noyau, le taux de classification a été calculé pour différentes valeurs de  $\sigma$  et pour la même valeur du paramètre de régularisation.

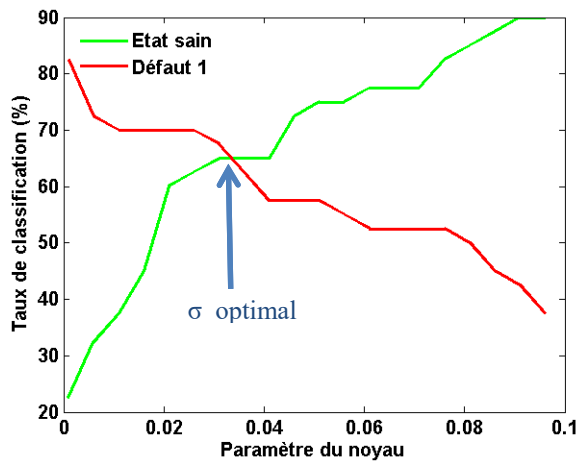


Figure 6. Taux de classification pour différentes valeurs de  $\sigma$

La Figure 6 montre que le taux de classification des données de l'état sain est inversement proportionnel à celui des cas d'un défaut. En effet, quand la portée du noyau augmente, le nombre d'échantillons de l'état sain bien classifiés augmente alors que ce nombre diminue pour les échantillons en défaut. Ce résultat se justifie par le fait que lorsque la portée du noyau est trop petite, chaque échantillon de la base de données d'apprentissage crée une frontière autour de lui favorisant ainsi le sur-apprentissage, contrairement au cas d'une portée trop grande qui permet d'englober les échantillons du défaut 1 dans les frontières de l'état sain. Finalement, le compromis pour la valeur de la portée peut facilement être identifié par l'intersection des deux courbes.

Afin de tenir en compte de l'influence du paramètre de régularisation  $C$ , cette étape a été reconduite pour différentes valeurs de  $C$ . Pour chaque valeur de  $C$ , la portée optimale du noyau ainsi que son taux de classification (commun pour l'état sain et défaut 1) ont été retenus. Le but final est trouver le couple optimal  $(\sigma, C)$  qui permet d'avoir le taux de classification maximal. Le résultat est donné dans la Figure 7.

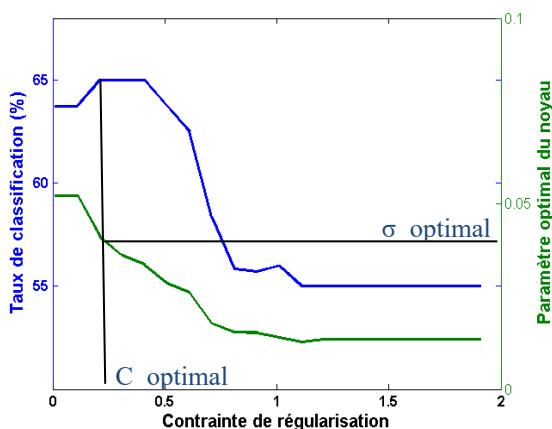


Figure 7. Taux de classification optimal en fonction du paramètre de régularisation et de la portée du noyau

Le taux de classification maximale pour les deux états de la structure a été estimé à 65% dans cette étude. Ce point permet l'identification de couple optimal  $(\sigma, C)$  à partir des données de validation. L'allure de la courbe de la portée du noyau est en cohérence avec la variation du paramètre de régularisation. En effet, plus la portée optimale diminue plus le poids des échantillons, qui peuvent être mal positionnés par rapport à la frontière de décision, est important.

Une fois le couple optimal identifié, la dernière étape de l'élaboration du séparateur est de tester ce modèle sur d'autres types de données qui n'ont servi ni pour l'apprentissage ni pour la validation. Le résultat montré dans le tableau suivant est très satisfaisant et concorde avec celui obtenu lors de la validation

Tableau 2: Taux de classification optimale pour les données de validation et de test

Données	$C = 0.21 ; \sigma = 0.041$	
	Etat sain	Défaut 1
Validation	<b>65%</b>	<b>65%</b>
Test	<b>67.5%</b>	<b>65%</b>

## 4 Conclusion

Cette étude a eu pour objectif le développement de la méthode SVM (Séparateur à vaste marge) pour la détection et la classification des données issues du monitoring par onde guidées des structures tubulaires. La base de données utilisée lors de la phase d'apprentissage de cette méthode a été construite à partir des signaux expérimentaux des différents états de la structure (sain et endommagé) et grâce à l'ajout d'un bruit qui permet de simuler la variation des conditions environnementales et opérationnelles (température, humidité, pluie etc). Compte tenu de la représentation des échantillons dans l'espace de données, le choix s'est porté sur un séparateur non linéaire assuré via une fonction noyau. Les résultats obtenus montrent que la fonction RBF permet de donner de meilleure performance en termes de détection et de pourcentage de fausse alarme. Il a été également montré que la sélection optimale de la portée du noyau et du paramètre de régularisation permet d'améliorer le résultat de la classification et par conséquent d'augmenter la sensibilité à la détection de la méthode.

## Références

- [1] C. R. Farrar and K. Worden, "An introduction to structural health monitoring.," Philos. Trans. A. Math. Phys. Eng. Sci., vol. 365, no. 1851, pp. 303–15, Feb. 2007.
- [2] H. Sohn, "Effects of environmental and operational variability on structural health monitoring.," Philos.

- Trans. A. Math. Phys. Eng. Sci., vol. 365, no. 1851, pp. 539–60, Feb. 2007.
- [3] H. Sohn, C. R. Farrar, N. F. Hunter, and K. Worden, “Structural Health Monitoring Using Statistical Pattern Recognition Techniques,” *J. Dyn. Syst. Meas. Control*, vol. 123, no. 4, p. 706, 2001.
- [4] M. Hassan, R. Rajkumar, D. Isa, and R. Arelhi, “Pipeline Defect Classification by Using Non-Destructive Testing and Improved Support Vector Machine Classification,” *International journal of engineering and innovative technology*, vol. 2, no. 7, pp. 85–93, 2013.
- [5] H. Zamani HosseinAbadi, R. Amirfattahi, B. Nazari, H. R. Mirdamadi, and S. A. Atashipour, “GUW-based structural damage detection using WPT statistical features and multiclass SVM,” *Appl. Acoust.*, vol. 86, pp. 59–70, Dec. 2014.
- [6] U. Dackermann, B. Skinner, and J. Li, “Guided-Wave-Based Condition Assessment of In-Situ Timber Utility Poles using Machine Learning Algorithms,” 2007.
- [7] X. Li, “Structural Damage Classification using Support Vector Machine,” *thèse de doctorat*, 2012.
- [8] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.
- [9] J. C. Platt, “Sequential Minimal Optimization : A Fast Algorithm for Training Support Vector Machines,” pp. 1–21, 1998.
- [10] C. Hsu, C. Chang, and C. Lin, “A Practical Guide to Support Vector Classification,” no. 1, pp. 1–16, 2010