

# Compromis exploration vs. exploitation en apprentissage actif : une approche Bayésienne

Alexis Bondu<sup>1</sup>, Marc Boullé<sup>2</sup>, Vincent Lemaire<sup>2</sup>

<sup>1</sup> EDF R&D ICAME/SOAD, 1 avenue du Général de Gaulle, 92140 Clamart.  
alexis.bondu@edf.fr

<sup>2</sup> OrangeLabs, 2 avenue Pierre Marzin, 22300 Lannion.  
prenom.nom@orange-ftgroup.com

**Résumé** : L'étiquetage des exemples d'apprentissage peut s'avérer être une tâche coûteuse lors d'une classification supervisée. Les stratégies d'apprentissage actif répondent à cette problématique en sélectionnant les exemples non-étiquetés les plus utiles à l'entraînement d'un modèle prédictif. Le choix des exemples à étiqueter peut être vu comme un dilemme entre l'exploration et l'exploitation de l'espace de représentation des données. Dans cet article, une nouvelle stratégie d'apprentissage actif gérant ce compromis est proposée. Cette stratégie est fondée sur un formalisme Bayésien qui minimise les hypothèses faites sur les données. Une validation expérimentale est menée sur des données unidimensionnelles, l'objectif étant d'estimer la position d'un échelon à partir de données bruitées. Notre approche est comparée favorablement à une stratégie ad hoc : la dichotomie probabiliste.

**Mots-clés** : Apprentissage Actif, Bayes, Dichotomie probabiliste

## 1 Introduction

L'apprentissage statistique désigne un vaste ensemble de méthodes et d'algorithmes qui permettent à un modèle d'apprendre un comportement grâce à des exemples. Dans la pratique, la collecte et l'étiquetage des exemples d'apprentissage peut s'avérer être une tâche coûteuse pour de nombreux problèmes de classification supervisée. Ce coût d'étiquetage peut en particulier être dû à l'implication d'un expert humain, d'un matériel de mesure, ou encore d'un temps de calcul important. Les stratégies d'apprentissage actif (Castro & Nowak (2005)) sélectionnent les exemples non-étiquetés jugés les plus informatifs pour construire un ensemble d'apprentissage de manière itérative. Un expert étiquette les exemples sélectionnés. Toutes les stratégies actives ont en commun de chercher à utiliser le moins d'exemples possible et à sélectionner les exemples les plus utiles à l'apprentissage d'un modèle prédictif. Nous nous plaçons dans le cadre de l'échantillonnage sélectif (Roy & McCallum (2001)) où le modèle observe un ensemble fini d'instances : la création de nouvelles instances par la stratégie active n'est pas autorisée.

Dans cet article, une nouvelle stratégie d'apprentissage actif fondée sur une méthode Bayésienne de discrétisation semi-supervisée (BonduICDM2008) est proposée. La Section 2 présente brièvement la méthode de discrétisation semi-supervisée sur laquelle est fondée notre approche. La stratégie active proposée est ensuite formalisée à la Section 3.2. Une validation expérimentale est réalisée. Un cadre applicatif est alors défini : i) les modèles de discrétisation comportent au plus deux intervalles ; ii) le problème traité est une classification binaire ; iii) les données d'apprentissage sont caractérisées par une seule variable explicative. Bien que restrictif, ce cadre applicatif représente une première phase de test et il permet d'optimiser notre critère de sélection d'exemple de manière analytique grâce aux optimisations algorithmiques que nous proposons. Des expériences comparatives ayant pour objectif de détecter la position d'un échelon à partir de données bruitées sont menées dans la Section 4, tout d'abord de manière à illustrer la méthode (Section 4.4) puis de manière à l'évaluer notre approche (Section 4.5). Ces expériences caractérisent l'influence du niveau de bruit présent dans les données sur les performances de notre stratégie active. Dans cette section, notre approche est comparée favorablement à une stratégie ad hoc : la dichotomie probabiliste (Horstein (1963)).

## 2 Discrétisation semi-supervisée Bayésienne

*Notations :* Les données  $D$  sont composées de deux sous-ensembles  $T$  et  $U$  qui correspondent respectivement aux données étiquetées et non-étiquetées, avec  $D = T \cup U$ . L'ensemble  $T$  contient des couples  $(x, y)$ , où  $x \in \mathbb{R}$  et  $y \in \mathbb{Y}$  est une valeur discrète représentant la classe de l'exemple  $x$ . L'ensemble  $U$  contient des réels. Les notations suivantes sont adoptées :  $N$ , le nombre d'exemples observables ( $N = |D|$ ) ;  $N^l$ , le nombre d'exemples étiquetés ( $N^l = |T|$ ) ;  $J$ , le nombre de classes observées dans les données ( $J = |\mathbb{Y}|$ ).

La stratégie d'apprentissage actif présentée dans cet article se base sur une méthode de discrétisation semi-supervisée issue de la famille MODL (Bondu *et al.* (2008)). Dans le cas de la classification semi-supervisée cette méthode discrétise les variables explicatives dans le but d'estimer les distributions conditionnelles aux classes, ces distributions sont supposées constantes par intervalles. L'approche MODL transpose le problème de la discrétisation d'une variable numérique en un problème de sélection de modèles. Une famille de modèles de discrétisation basée sur la statistique d'ordre est définie. Un modèle de discrétisation  $M(I, \{N_i\}, \{N_{ij}\})$  est défini par les paramètres suivants : i)  $I$  est le nombre d'intervalles ; ii)  $\{N_i\}$  est le nombre d'exemples dans chaque intervalle ; iii)  $\{N_{ij}\}$  est le nombre d'exemples de chaque classe dans chaque intervalle. Les paramètres  $\{N_i\}$  spécifient les bornes des intervalles grâce aux rangs des valeurs explicatives et les paramètres  $\{N_{ij}\}$  caractérisent les distributions conditionnelles grâce aux effectifs de chaque valeur à expliquer localement à l'intervalle  $i$ . Une démarche Bayésienne est appliquée pour sélectionner le meilleur modèle de discrétisation, noté  $\mathcal{M}_{map}$  (*Maximum a posteriori*). Le meilleur modèle de discrétisation maximise  $P(M|D)$ , la probabilité du modèle  $M$  connaissant les données  $D$ . En exploitant la formule de Bayes et en considérant que le terme  $P(D)$  est constant quel que soit le modèle, cela revient à maximiser  $P(M)P(D|M)$ . La distribution a priori des modèles  $P(M)$  et la vrai-

semblance des données  $P(D|M)$  sont calculées analytiquement en exploitant le caractère discret de la famille de modèles. La démarche Bayésienne employée adopte des hypothèses faiblement informatives sur les données, cette démarche est dite objective (Robert (2006)). Finalement, le  $\mathcal{M}_{map}$  minimise l'Equation 1. Le critère  $\mathcal{C}_{semi\ super}$  calcule le log négatif d'une probabilité, ce qui correspond à une quantité d'information (Shannon (1948)). Le premier terme correspond au choix du nombre d'intervalles, le deuxième terme au choix des bornes des intervalles. Le troisième terme représente le choix des distributions de la variable à expliquer dans chaque intervalle. L'avant dernier terme représente la probabilité d'observer les valeurs de la variable à expliquer connaissant le modèle de discrétisation. Le dernier terme correspond à une pénalisation due aux exemples non-étiquetés dans chaque intervalle du modèle (notés  $N_i^u$  et  $N_{ij}^u$ ).

$$\mathcal{C}_{semi\ super}(M) = \overbrace{\log(N) + \log C_{N+I-1}^{I-1} \sum_{i=1}^I \log C_{N_i+J-1}^{J-1}}^{-\log P(M)} + \underbrace{\sum_{i=1}^I \log \left( \frac{N_i!}{\sum_{j=1}^J N_{ij}!} \right) - \sum_{i=1}^I \log \left( \frac{N_i^u!}{\sum_{j=1}^J N_{ij}^u!} \right)}_{-\log P(D|M)} \quad (1)$$

### 3 Apprentissage actif

Cette section introduit l'apprentissage actif et souligne l'importance du dilemme entre l'exploitation et l'exploration de l'espace de représentation des données. La stratégie d'apprentissage que nous proposons est ensuite formalisée et présentée en détail.

#### 3.1 Un dilemme entre exploitation et exploration

Lors d'un apprentissage actif, le choix des exemples à étiqueter peut être vu comme le dilemme entre l'exploration et l'exploitation de l'espace de représentation des données (noté  $\mathbb{X}$ ). La sélection d'un exemple non-étiqueté dans une zone non-échantillonnée de  $\mathbb{X}$  contribue à **explorer** les données. Dans ce cas l'espace  $\mathbb{X}$  tend à être échantillonné uniformément, ce qui limite les zones pour lesquelles l'hypothèse apprise par le modèle est potentiellement incorrecte. Plus la dimension de  $\mathbb{X}$  est élevée, plus l'exploration de cet espace requiert un nombre important d'exemples étiquetés. La sélection d'un exemple non-étiqueté dans une zone échantillonnée de  $\mathbb{X}$  contribue à **exploiter** les données. Dans ce cas, la stratégie active se concentre sur une zone déjà peuplée d'exemples étiquetés et affine localement le modèle prédictif. Le dilemme entre l'exploitation et l'exploration des données peut être illustré grâce à deux situations extrêmes. D'une part, une stratégie d'apprentissage actif qui ne fait qu'exploiter les données court le risque d'occulter une grande partie de l'espace  $\mathbb{X}$ . Le modèle prédictif sera spécialisé dans certaines régions de  $\mathbb{X}$  mais sera vraisemblablement de mauvaise qualité en généralisation. D'autre part, une stratégie qui ne fait qu'explorer les données ne se focalise pas sur les régions de  $\mathbb{X}$  où l'étiquetage de nouveaux exemples améliorerait le modèle prédictif.

Dans ces conditions, une stratégie d'apprentissage actif présente peu d'intérêt par rapport à un échantillonnage aléatoire. Ces deux situations extrêmes illustrent la nécessité pour une stratégie d'apprentissage actif de trouver un **compromis** entre l'exploitation et l'exploration des données. Plusieurs approches existent dans la littérature pour aborder ce compromis. Nous citons ci-dessous les 3 approches les plus rencontrées.

**L'exploitation simultanée de plusieurs stratégies :** Plusieurs stratégies d'apprentissage actif peuvent être utilisées conjointement pour répondre au compromis exploitation / exploration. Considérons deux stratégies actives, comme par exemple l'échantillonnage par incertitude (Thrun & Möller (1992)) et l'échantillonnage aléatoire, respectivement dédiées à l'exploitation et à l'exploration des données. A chaque itération, une des deux stratégies est utilisée. Le choix de la stratégie est probabiliste,  $p$  représente la probabilité d'explorer les données et  $1 - p$  la probabilité de les exploiter ; ces probabilités devant être mises à jour au cours de l'apprentissage actif pour ajuster le compromis exploitation / exploration. T. Osugi (Osugi *et al.* (2005)) fait un parallèle avec l'apprentissage par renforcement. Le modèle prédictif est considéré comme étant un "agent" qui peut faire deux actions : explorer ou exploiter les données. A chaque itération, l'agent décide d'exécuter une action et reçoit une récompense [*respectivement* une punition] si cette action est appropriée [*respectivement* inappropriée]. Une métrique est utilisée pour mesurer la variation de l'hypothèse apprise par le modèle entre deux itérations. Une action est d'autant plus récompensée que l'hypothèse apprise par le modèle varie. La probabilité  $p$  est mise à jour après chaque action. Il est également possible d'utiliser des heuristiques issues de l'optimisation combinatoire pour gérer le compromis exploration / exploitation. T. Zoller (Zoller & Buhmann (2000)) utilise un algorithme de recuit simulé pour faire varier  $p$  au cours du temps. Cette heuristique inspirée de la thermodynamique fait décroître la probabilité d'explorer les données selon une fonction prédéfinie, aussi appelée schéma de refroidissement. Cette heuristique concentre l'exploration au début de l'apprentissage actif et exploite par la suite des données représentatives de  $\mathbb{X}$ .

**Le pré-clustering :** H. Nguyen améliore la stratégie d'échantillonnage par incertitude (Thrun & Möller (1992)) en appliquant un pré-clustering aux données d'apprentissage, ce qui permet de diversifier les exemples sélectionnés. Seuls les centroïdes des clusters sont candidats à l'étiquetage. Les exemples d'un même cluster sont supposés appartenir à la même classe et les centroïdes sont considérés comme étant représentatifs des clusters. H. Nguyen propose un critère qui sélectionne le centroïde contribuant le plus à l'erreur courante. L'utilisateur fait varier la taille des clusters au cours des itérations successives de l'apprentissage actif. La diminution de la taille des clusters au cours des itérations répond au compromis exploration / exploitation. Lorsque les clusters contiennent beaucoup d'exemples, les centroïdes sont éloignés les uns des autres. Dans ce cas, la sélection des exemples à étiqueter favorisera l'exploration des données. Au contraire, lorsque les clusters contiennent peu d'exemples, les centroïdes sont potentiellement proches les uns des autres. Dans ce cas, la stratégie basée sur le pré-clustering exploite les données. A l'instar du *recuit simulé* (Zoller & Buhmann (2000)), cette stratégie explore davantage les données au début de l'échantillonnage sélectif qu'à la fin.

**Utilisation d'une mesure de dissimilarité :** Certaines stratégies d'apprentissage actif abordent le compromis exploration / exploitation en mesurant la dissimilarité entre les exemples sélectionnés. Xu (Xu *et al.* (2007)) propose une stratégie active multi-critères appliquée à la recherche documentaire. Cette approche maximise la distance entre le nouvel exemple et l'exemple étiqueté le plus proche et étiquette des exemples les plus éloignés les uns des autres, au sens de la métrique utilisée. Brinker (Brinker (2003)) présente une stratégie basée sur une mesure de dissimilarité qui étiquette un ensemble d'exemples à chaque itération, de manière à réduire le plus possible l'espace des versions du modèle prédictif. Chaque exemple candidat à l'étiquetage correspond à un hyperplan dans l'espace préhilbertien induit par le noyau. Cette stratégie sélectionne des exemples non-étiquetés dont les hyperplans correspondants sont orientés selon des directions les plus éloignées les unes des autres. Une mesure d'angle est définie en exploitant l'astuce du noyau.

Finalement, les 3 méthodes présentées ci-dessus ainsi que l'état de l'art (Anonyme. (2008), Settles (2009)) montrent que le compromis entre l'exploitation et l'exploration est une question centrale à l'apprentissage actif. La stratégie active présentée dans la suite de cet article propose de gérer ce compromis de manière originale en se fondant sur un formalisme Bayésien qui minimise les hypothèses faites sur les données, et ce sans impliquer de paramètre utilisateur.

### 3.2 Une nouvelle stratégie

Cette section présente une stratégie originale d'apprentissage actif fondée sur la méthode de discrétisation semi-supervisée décrite à la Section 2. La qualité d'un modèle de discrétisation est donnée par la probabilité du modèle connaissant les données. Le critère  $C_{semi\ super}$  est une expression analytique de  $P(M|D)$ , au sens des hypothèses de modélisation de l'approche MODL (Boullé (2006)). Notre stratégie d'apprentissage actif cherche à étiqueter l'exemple qui maximisera la qualité du futur modèle, sans connaître la classe du nouvel exemple et sans connaître le meilleur modèle à l'itération suivante. Notre démarche prend en compte ces incertitudes en menant un calcul d'espérance sur tous les cas possibles. Un critère dont l'optimisation désigne l'exemple  $x_{t+1} \in U$  qui maximise l'espérance de  $P(M|D, x_{t+1})$  est établi.

Soit  $P_{(\cdot|D)}(M) = P(M|D)$  la distribution a posteriori des modèles de discrétisation, connaissant les données. Notre stratégie active sélectionne l'exemple  $x_{t+1} \in U$  qui maximise l'espérance de  $P(M|D, x_{t+1})$  sur l'ensemble des modèles  $M \in \mathbb{M}$  de la manière suivante :

$$ArgMax_{x_{t+1} \in U} E_{\substack{P_{(\cdot|D)} \\ M \in \mathbb{M}}} [P(M|D, x_{t+1})] = ArgMax_{x_{t+1} \in U} \sum_{M \in \mathbb{M}} P(M|D) \times P(M|D, x_{t+1})$$

L'étiquette  $y_{t+1}$  est inconnue, mais la probabilité  $P(y|M, D, x_{t+1})$  de la classe  $y \in \mathbb{Y}$  connaissant le modèle et les données, peut être évaluée. Grâce à la formule de la probabilité totale, nous écrivons :

1.  $\sum_{y \in \mathbb{Y}} P(y|M, D, x_{t+1}) = 1$
2.  $P(M|D, x_{t+1}) = \sum_{y \in \mathbb{Y}} P(y|M, D, x_{t+1})P(M|D, x_{t+1}, y)$

Finalement,

$$ArgMax_{x_{t+1} \in U} \sum_{M \in \mathbb{M}} P(M|D) \times \left[ \sum_{y \in \mathbb{Y}} P(y|M, D, x_{t+1}) \times P(M|D, x_{t+1}, y) \right]$$

Cette expression est ensuite développée en exploitant la formule de Bayes :

$$ArgMax_{x_{t+1} \in U} \sum_{M \in \mathbb{M}} \left[ \frac{P(M) \times P(D|M)}{P(D)} \times \sum_{y \in \mathbb{Y}} \left[ P(y|M, D, x_{t+1}) \times \frac{P(M) \times P(D, x_{t+1}, y|M)}{P(D, x_{t+1}, y)} \right] \right] \quad (2)$$

La probabilité jointe  $P(D, x_{t+1}, y)$  est décomposée comme suit :

$$P(D, x_{t+1}, y) = P(D) \times P(x_{t+1}|D) \times P(y|D, x_{t+1}) \quad (3)$$

L'exemple  $x_{t+1}$  est considéré comme étant tiré uniformément dans l'ensemble  $U$ , puisqu'à l'itération  $t$ , tous les exemples non-étiquetés ont la même probabilité d'être sélectionnés. Les termes  $P(D)$  et  $P(x_{t+1}|D)$  de l'équation 2 restent donc constants lorsque que le modèle  $M$  varie. Cette équation se réécrit comme suit :

$$ArgMax_{x_{t+1} \in U} \sum_{M \in \mathbb{M}} \left[ \overbrace{P(M) \times P(D|M)}^A \times \sum_{y \in \mathbb{Y}} \left[ \underbrace{P(y|M, D, x_{t+1})}_C \times \underbrace{\frac{P(M) \times P(D, x_{t+1}, y|M)}{P(y|D, x_{t+1})}}_D \right] \right] \times Cste \quad (4)$$

Le terme "A" de l'équation 4 est déduit du critère  $\mathcal{C}_{semi\ super}$  :

$$P(M)P(D|M) = \frac{1}{N} \times \frac{1}{C_{N+J-1}^{I-1}} \times \prod_{i=1}^I \frac{1}{C_{N_i+J-1}^{J-1}} \times \prod_{i=1}^I \left[ \frac{\prod_{j=1}^J N_{ij}!}{N_i!} \times \frac{N_i^{u_i}!}{\prod_{j=1}^J N_{ij}^{u_j}!} \right]$$

Le terme "B" est calculé de la même façon, en ajoutant le couple  $(x_{t+1}, y)$  à l'ensemble d'apprentissage  $T$ . Le terme "C" est évalué par la prédiction du modèle  $M$ . Le modèle courant estime la probabilité d'observer la classe  $y$  connaissant l'exemple  $x_{t+1}$ . Cette prédiction se base sur la proportion des exemples étiquetés par la classe  $y$  dans l'intervalle contenant l'exemple  $x_{t+1}$ . Le terme "D" représente la probabilité d'observer la classe  $y$ , connaissant l'exemple  $x_{t+1}$  et les données. Ce terme est difficile à évaluer car il n'implique aucun modèle particulier. Pour évaluer ce terme, nous choisissons d'intégrer sur l'ensemble des modèles  $M \in \mathbb{M}$ . En exploitant la formule de la probabilité totale [ $\sum_{M' \in \mathbb{M}} P(M'|D) = 1$ ], nous pouvons écrire :

$$P(y|D, x_{t+1}) = \sum_{M' \in \mathbb{M}} P(M'|D) \times P(y|D, M', x_{t+1}) = \sum_{M' \in \mathbb{M}} \frac{P(D|M')P(M')}{P(D)} \times P(y|D, M', x_{t+1})$$

La probabilité des données  $P(D)$  reste constante lorsque le modèle  $M$  varie. Finale-ment, l'espérance de  $P(M|D, x_{t+1})$  est évaluée par le critère  $\mathcal{C}_{actif}(x_{t+1})$  :

$$\mathcal{C}_{actif}(x_{t+1}) = \sum_{M \in \mathbb{M}} \left[ P(M)P(D|M) \times \sum_{y \in \mathbb{Y}} \left[ \frac{P(y|M, D, x_{t+1}) \times P(M) \times P(D, x_{t+1}, y|M)}{\sum_{M' \in \mathbb{M}} P(M') \times P(D|M') \times P(y|D, M', x_{t+1})} \right] \right] \quad (5)$$

Notre stratégie sélectionne l'exemple  $x_{t+1} \in U$  maximisant  $\mathcal{C}_{actif}(x_{t+1})$ .

## 4 Validation expérimentale

Notre stratégie active est évaluée dans un cas très simple : l'estimation d'une fonction échelon à partir de données bruitées (Castro & Nowak (2008)). Ce problème d'apprentissage constitue une première application de notre stratégie de sélection d'exemples. Dans ce cas particulier, où les données sont unidimensionnelles et où le modèle de discrétisation inclue un ou deux intervalles, le critère  $\mathcal{C}_{actif}$  peut être optimisé analytiquement en un temps raisonnable. Dans cette section, nous définissons le problème d'apprentissage considéré lors de nos expériences. Une stratégie concurrentes ad hoc est ensuite présentée. Nous proposons des optimisations algorithmiques pour notre approche. Enfin deux types de résultats sont discutés : i) les résultats illustratifs montrent le comportement de notre stratégie active ; ii) les résultats comparatifs évaluent la performance de chaque stratégie en fonction du nombre d'exemples étiquetés.

### 4.1 Jeu de données

Le jeu de données utilisé lors de nos expériences comporte 100 exemples uniformément distribués dans l'intervalle  $[0, 1]$ . Le domaine de définition de la variable explicative  $x$  est divisé en deux intervalles  $[0, \theta[$  et  $[\theta, 1]$ , avec  $\theta$  la position de l'échelon. L'objectif de ce problème d'apprentissage est d'évaluer  $\theta$  à partir de données bruitées. Les exemples d'apprentissage de l'intervalle  $[0, \theta]$  [respectivement  $[\theta, 1]$ ] sont majoritairement étiquetés par la classe "1" [respectivement par la classe "2"]. La probabilité que l'étiquette d'un exemple soit erronée est constante sur l'intervalle  $[0, 1]$  et est notée  $p \in [0, 0.5]$ . La Figure 1 trace la probabilité d'observer la classe "1" connaissant la valeur de la variable explicative  $x$ .

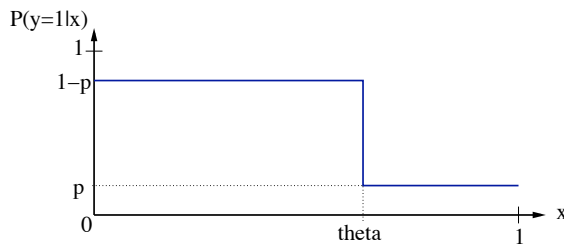


FIGURE 1 – Echelon bruité : probabilité d'observer la classe "1" connaissant  $x$ .

### 4.2 Stratégies concurrentes

Deux stratégies concurrentes sont évaluées : l'aléatoire et la dichotomie probabiliste.

Echantillonnage aléatoire : L'échantillonnage aléatoire sélectionne les exemples uniformément. Cette stratégie de référence permet de mesurer l'apport des autres approches.

*Dichotomie probabiliste* : M. Horstein généralise la stratégie dichotomique au cas de données bruitées (Horstein (1963)). Nous appelons cette stratégie “la dichotomie probabiliste”. Cette approche ad hoc à l’estimation de la position d’un échelon suppose que le niveau de bruit  $p$  est connu. Comme le montre l’étape (I) de l’Algorithme 1, la distribution de probabilité de la position de l’échelon est notée  $P_\theta(x)$ . Initialement aucune étiquette n’est disponible. Une première étape initialise la distribution  $P_\theta(x)$  selon un a priori uniforme. A chaque itération, l’exemple non-étiqueté  $x^* \in U$  qui “coupe” la distribution  $P_\theta(x)$  en deux parties égales est sélectionné (étape A).  $x^*$  est l’exemple le plus proche de la médiane de  $P_\theta(x)$ . L’étiquette  $f(x^*)$  est ensuite demandée à un expert (étape B). Le couple  $(x^*, f(x^*))$  est ajouté à l’ensemble d’apprentissage  $T$ . La distribution  $P_\theta(x)$  est mise à jour à chaque itération, selon la valeur de la nouvelle étiquette  $f(x^*)$  (étape C). Cette étape prend également en considération le niveau de bruit  $p$ .

Notations :

- $P_\theta(x)$  la distribution de probabilité a priori de  $\theta$ .
- $p$  la probabilité qu’une étiquette soit erronée.
- Un échelon bruité, tel que l’intervalle  $[0, \theta]$  [respectivement  $[\theta, 1]$ ] soit majoritairement constitué d’exemples de la classe “1” [respectivement de la classe “2”].
- $n$  le nombre d’exemples d’apprentissage souhaité.
- Les ensembles  $U$  et  $L$  d’exemples non étiquetés et étiquetés, avec  $L \cup U = \Phi$
- L’ensemble d’apprentissage  $T$  avec  $|T| < n$

*/\*initialisation des ensembles L et U\*/*

$L = \emptyset$  et  $U = \Phi$

*/\*initialisation de  $P_\theta(x)$ \*/*

(I)  $P_\theta(x) \leftarrow \frac{1}{N} \quad \forall x \in \Phi$

**Répéter**

(A) Trouver l’exemple non-étiqueté  $x^* \in U$  tel que :  $\sum_{x \in [0, x^*]} P_\theta(x) = \frac{1}{2}$

(B) Demander l’étiquette  $f(x^*)$ , ajouter  $(x^*, f(x^*))$  à  $T$  et  $x^*$  à  $L$ , retirer  $x^*$  de  $U$ .

(C) Mettre à jour  $P_\theta(x)$  tel que :

**Si**  $f(x^*) = 1$  **Alors**

- (i)  $P_\theta(x) \leftarrow 2.p.P_\theta(x) \quad \forall x \in [0, x^*]$
- (ii)  $P_\theta(x) \leftarrow 2.(1-p).P_\theta(x) \quad \forall x \in ]x^*, 1]$

**Fin Si**

**Si**  $f(x^*) = 2$  **Alors**

- (iii)  $P_\theta(x) \leftarrow 2.(1-p).P_\theta(x) \quad \forall x \in [0, x^*]$
- (iv)  $P_\theta(x) \leftarrow 2.p.P_\theta(x) \quad \forall x \in ]x^*, 1]$

**Fin Si**

**Tant** que  $|T| < n$

Algorithme 1: Dichotomie probabiliste



### 4.3 Optimisations algorithmiques de notre approche

Cette section présente une partie des optimisations algorithmiques mises en œuvre pour réduire la complexité de notre approche, dans le cas de données unidimensionnelles et d'un modèle incluant un ou deux intervalles. Pour calculer l'espérance de  $P(M|D, x_{t+1})$ , la famille de modèles est parcourue deux fois de manière imbriquée (l'équation comporte deux sommes sur  $\mathbb{M}$ ). La complexité du parcours de la famille de modèles est de l'ordre de  $\mathcal{O}(N^{(I-1+J)})$ , dans notre cadre applicatif  $I = 2$  et  $J = 2$ . Les deux parcours de  $\mathbb{M}$  nécessaires au calcul de  $P(M|D, x_{t+1})$  sont imbriqués (voir Equation 5). Par ailleurs, le parcours des exemples non-étiquetés a une complexité de l'ordre de  $\mathcal{O}(N)$ . Finalement, la sélection de l'exemple  $x_{t+1} \in U$  maximisant  $\mathcal{C}_{actif}(x_{t+1})$  a une complexité de l'ordre de  $\mathcal{O}(N^7)$ . Malgré le cadre applicatif très simple considéré ici, la complexité de notre stratégie active est très élevée lorsque celle-ci est implémentée naïvement. Les optimisations algorithmiques proposées par la suite ne dégradent pas le calcul de  $\mathcal{C}_{actif}(x_{t+1})$ , le parcours de  $\mathbb{M}$  étant toujours exhaustif. Le parcours exhaustif de la famille de modèles  $\mathbb{M}$  peut être effectué une seule passe.

$$\mathcal{C}_{actif}(x_{t+1}) = \sum_{M \in \mathbb{M}} \left[ P(M)P(D|M) \times \sum_{y \in \mathbb{Y}} \left[ \underbrace{\frac{P(y|M, D, x_{t+1}) \times P(M) \times P(D, x_{t+1}, y|M)}{\sum_{M' \in \mathbb{M}} P(M') \times P(D|M') \times P(y|D, M', x_{t+1})}}_A \right] \right] \quad (6)$$

La partie "A" de l'équation 6 dépend uniquement de la classe  $y \in \mathbb{Y}$  et de l'instance  $x_{t+1} \in U$ . Ce terme est indépendant du modèle  $M$  impliqué dans la première somme sur  $\mathbb{M}$ . L'équation 6 peut être réorganisée de la manière suivante :

$$\begin{aligned} \mathcal{C}_{actif}(x_{t+1}) &= \frac{\overbrace{\sum_{M \in \mathbb{M}} P(M)P(D|M) \times P(y_1|M, D, x_{t+1}) \times P(M)P(D, x_{t+1}, y_1|M)}^A}{\underbrace{\sum_{M' \in \mathbb{M}} P(M')P(D|M') \times P(y_1|D, M', x_{t+1})}_B} \\ &+ \frac{\overbrace{\sum_{M \in \mathbb{M}} P(M)P(D|M) \times P(y_2|M, D, x_{t+1}) \times P(M)P(D, x_{t+1}, y_2|M)}^C}{\underbrace{\sum_{M' \in \mathbb{M}} P(M')P(D|M') \times P(y_2|D, M', x_{t+1})}_D} \quad (7) \end{aligned}$$

Les termes sont calculés indépendamment les uns des autres. Etant donné leur caractère additif, les termes "A", "B", "C" et "D" de l'équation 7 sont calculés en parallèle. La complexité temporelle de notre stratégie active se réduit alors à  $\mathcal{O}(N^4)$ . D'autres optimisations présentées dans (Anonyme. (2008)) permettent de gagner un facteur constant qui a été mesuré en pratique comme étant supérieur à 5000. Ces optimisations exploitent l'additivité du critère et l'ordre de parcours des modèles pour éviter de ré-estimer tous les termes de  $\mathcal{C}_{actif}$  pour chaque modèle de  $\mathbb{M}$ .

## 4.4 Résultats illustratifs

Les expériences présentées ici ont été réalisées sur le jeu de données décrit à la Section 4.1. Ces expériences illustrent la méthode proposée et montrent l'évolution de  $\mathcal{C}_{actif}(x_{t+1})$  (Equation 5) au cours des itérations d'un apprentissage actif. Les exemples sélectionnés, ainsi que leurs étiquettes, sont représentés. La position de l'échelon est fixé à  $\theta = 0.5$ .

*échelon pur :*

La Figure 2 détaille la sélection d'exemples, lors des premières itérations d'un apprentissage actif. Les courbes représentent<sup>1</sup>  $\mathcal{C}_{actif}(x_{t+1})$  (axe vertical), en fonction de la position de l'exemple candidat à l'étiquetage (axe horizontal). La valeur maximum de chaque courbe est symbolisée par "▼" et correspond à la position de l'exemple sélectionné à chaque itération. Les exemples de la classe "1" [respectivement "2"] sont symbolisés par "●" [respectivement "●"]. Initialement aucun exemple n'est étiqueté. Lors de la première itération (graphique "A", Figure 2), le critère  $\mathcal{C}_{actif}(x_{t+1})$  est maximal pour  $x_{t+1} = 0$  et  $x_{t+1} = 1$ . Ce graphique illustre l'intérêt d'utiliser la version semi-supervisée du critère MODL (voir Section 2) : la pénalisation des exemples non-étiquetés engendre la sélection d'un des exemples situés aux extrémités de l'intervalle  $[0, 1]$ . En cas d'égalité, un des exemples est choisi aléatoirement. Ici, l'exemple  $x_{t+1} = 1$  est étiqueté par la classe "2". A la deuxième itération (graphique "B", Figure 2), l'exemple  $x_{t+1} = 0$  est étiqueté par la classe "1". A la troisième itération la courbe semble très plate (graphique "C", Figure 2). Le critère  $\mathcal{C}_{actif}(x_{t+1})$  est pourtant maximal en deux points. Finalement, l'exemple  $x_{t+1} = 0.28$  est sélectionné et étiqueté par la classe "1". A partir de 2 étiquettes par classe, notre stratégie adopte un comportement qui semble proche de la dichotomie. Les graphiques "E, F" de la Figure 2 montrent que les exemples sélectionnés se situent quasiment à mi-chemin des étiquettes "1" et "2" les plus proches. Notre stratégie converge à 0.5, après avoir étiqueté 9 exemples.

*échelon bruité :*

La même expérience a été réalisée en impliquant un exemple bruité, dont l'étiquette est erronée. Dans ce cas, notre stratégie se comporte de deux manières différentes :

1. Dans un premier temps, notre stratégie cherche à préciser la position de l'échelon autour de l'exemple bruité. Ce comportement est cohérent car rien ne laisse présager qu'une étiquette est erronée.
2. Notre stratégie "détecte" l'exemple bruité en l'encadrant par deux exemples de classe opposée. Notre stratégie poursuit sa recherche en adoptant le même comportement que précédemment (Figure 2).

Finalement, notre stratégie se comporte correctement face à des données bruitées. La position exacte du front montant de l'échelon est déterminée en étiquetant 12 exemples.

---

1. L'espérance est normalisée de telle sorte que son maximum soit égal à 1.

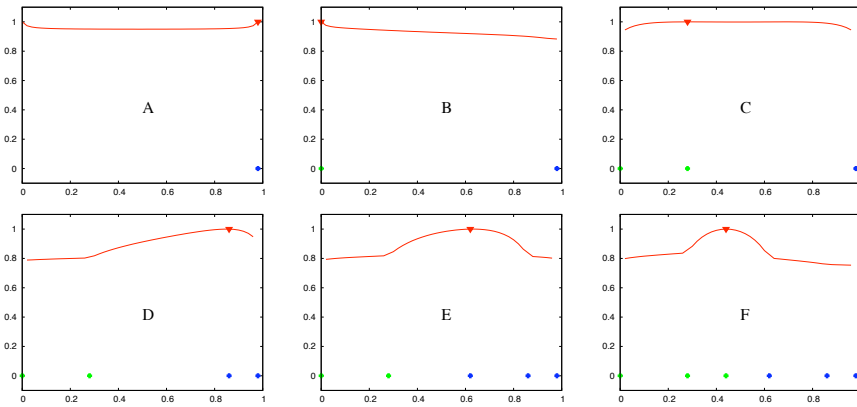


FIGURE 2 – Visualisation des exemples étiquetés pour l'échelon pur. L'axe vertical représente  $\mathcal{C}_{actif}(x_{t+1})$  et l'axe horizontal correspond à position de  $x_{t+1}$ .

## 4.5 Résultats comparatifs

Cette section présente des expériences d'évaluation, réalisées sur un échelon dont la position est fixée à  $\theta = 0.675$ . L'objectif de ces expériences est de caractériser l'influence du bruit d'étiquetage  $p$  sur les performances des stratégies actives,  $p$  variant entre 0 et 0.20. Les stratégies sont toutes évaluées en utilisant le même modèle prédictif, seule la sélection des exemples étiquetés influe sur la performance du modèle prédictif. Le modèle utilisé est le maximum a posteriori ( $\mathcal{M}_{map}$ ) défini à la Section 2 qui inclut un ou deux intervalles. Le modèle a le choix de couper ou non le domaine de définition de la variable explicative  $x$  et évalue la position de la coupure éventuelle. La qualité des stratégies actives est évaluée grâce à l'AUC en fonction du nombre d'exemples étiquetés.

Le protocole expérimental est le suivant : i) les expériences sont répétées 100 fois de manière à obtenir une AUC moyenne et une mesure de variance pour chaque valeur de  $|T|$ ; ii) au début de chaque expérience aucun exemple n'est étiqueté,  $L = \emptyset$ ; iii) à chaque itération de l'apprentissage actif un exemple est sélectionné puis étiqueté; iv) les expériences s'arrêtent lorsque  $|T| = 20$ .

A la différence des autres stratégies évaluées, la dichotomie probabiliste est renseignée du niveau de bruit  $p$ . Lors de nos expériences, trois cas sont envisagés :

Cas où la dichotomie probabiliste est correctement renseignée du niveau bruit : La Figure 3 trace l'AUC moyenne observée sur 100 expériences (axe vertical) en fonction du nombre d'exemples étiquetés (axe horizontal). Les moustaches sur les courbes représentent la variance des résultats ( $\pm 2\sigma$ ). Les graphiques de la Figure 3 correspondent à quatre niveaux de bruit différents.

Considérons dans un premier temps le cas où le bruit d'étiquetage est nul ( $p = 0$ ). La stratégie aléatoire présente une AUC constante, égale à 0.5, lorsque le nombre

d'exemples étiquetés est inférieur à 12 (courbe rouge). Dans ce cas, le maximum a posteriori comporte un seul intervalle et estime les distributions conditionnelles aux classes de manière uniforme. Lorsque suffisamment d'exemples sont étiquetés pour produire un  $\mathcal{M}_{map}$  à deux intervalles ( $|T| > 12$ ) la stratégie aléatoire progresse et atteint une AUC de 0.95 à la fin de l'expérience.

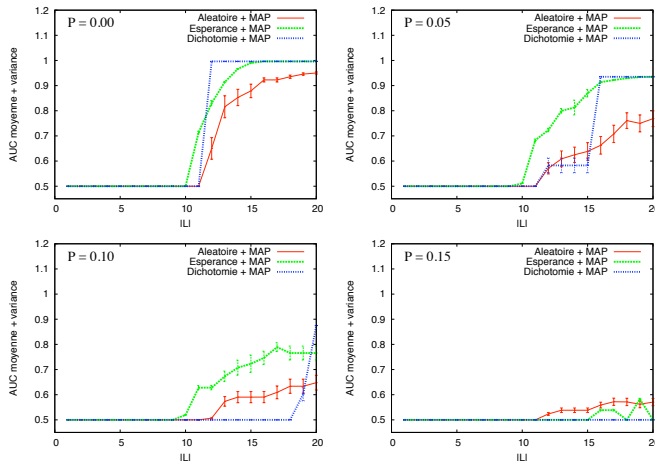


FIGURE 3 – Performance moyenne des stratégies actives vs. le nombre d'exemples étiquetés. Cas où la dichotomie probabiliste est correctement renseignée du niveau de bruit

Notre stratégie basée sur la maximisation de l'espérance de  $P(M|D, x_{t+1})$  présente deux caractéristiques intéressantes (courbe verte) : i) cette stratégie est plus performante que l'aléatoire quel que soit le nombre d'exemples étiquetés ; ii) cette stratégie requiert moins d'exemples étiquetés que l'aléatoire pour produire un modèle optimal à deux intervalles (l'AUC est supérieure à 0.5 pour  $|T| > 10$ ). Finalement, notre stratégie se comporte honorablement et atteint la performance optimale (AUC=1) après avoir étiqueté 15 exemples.

Comme la stratégie aléatoire, la dichotomie probabiliste présente une AUC égale à 0.5 lorsque le nombre d'exemples étiquetés est inférieur à 12 (courbe bleue). Cette stratégie progresse ensuite très rapidement et atteint la performance optimale dès que le  $\mathcal{M}_{map}$  inclut deux intervalles ( $|T| = 13$ ). Dans le cas d'un bruit d'étiquetage nul la dichotomie probabiliste est la stratégie la plus performante.

Lorsque le niveau de bruit augmente ( $p = 0.05, p = 0.10, p = 0.15$ ) les performances des trois stratégies diminuent. Notre stratégie reste plus performante que la stratégie aléatoire, pour  $p = 0.05$  et  $p = 0.10$ . La dichotomie probabiliste se dégrade très nettement lorsque  $p$  croît. Plus le bruit d'étiquetage augmente, plus le nombre d'exemples étiquetés nécessaires à l'obtention d'un  $\mathcal{M}_{map}$  à deux intervalles est important.

Les expériences réalisées dans ce paragraphe montrent que notre stratégie active est plus performante que la dichotomie probabiliste lorsque le niveau de bruit est important. Les expériences présentées ici sont pourtant favorables à la dichotomie probabiliste, car

le niveau de bruit  $p$  est connu.

*Cas où la dichotomie probabiliste est renseignée par un niveau de bruit nul :* D'autres expériences ont été réalisées en renseignant la dichotomie probabiliste d'un niveau de bruit nul, alors que  $p$  varie. Dans ce cas, notre stratégie offre de meilleures performances que la dichotomie probabiliste et cette tendance s'accroît lorsque le niveau de bruit présent dans les données augmente.

*Cas où la dichotomie est renseignée par  $p_{false} \in [0, 0.20]$ , alors que  $p = 0$  :* Les dernières expériences réalisées renseignent la dichotomie probabiliste par un niveau de bruit erroné, noté  $p_{false}$ , le véritable niveau de bruit  $p$  étant nul. Dans ce cas, les performances de la dichotomie probabiliste sont très mauvaises dès  $p_{false}$  atteint 0.1 : le  $\mathcal{M}_{map}$  comporte un seul intervalle quel que soit le nombre d'exemples étiquetés.

Pour conclure, les expériences qui ont été réalisées montrent que notre stratégie basée sur la maximisation de l'espérance de  $P(M|D, x_{t+1})$  domine la dichotomie probabiliste, particulièrement lorsque celle-ci est mal renseignée du niveau de bruit présent dans les données.

## 5 Conclusion et perspectives

Cet article présente une nouvelle stratégie d'apprentissage actif basée sur une méthode de discrétisation semi-supervisée issue de la famille MODL (Boullé (2006)). Cette stratégie sélectionne l'exemple non-étiqueté qui maximise l'espérance de la probabilité des modèles de discrétisation, connaissant les données et un exemple supplémentaire noté  $x_{t+1}$ . Notre démarche aboutit à un critère d'optimisation, noté  $\mathcal{C}_{actif}(x_{t+1})$ . Notre approche est validée dans le cadre de données unidimensionnelles, en considérant des modèles de discrétisation comportant un ou deux intervalles. Dans le cas de l'estimation de la position d'un échelon à partir de données bruitées, notre approche est évaluée et comparée favorablement à une stratégie ad hoc : la dichotomie probabiliste. L'étude comparative réalisée à la Section 4.5 montre que notre stratégie active est plus performante que la dichotomie probabiliste, particulièrement lorsque le niveau de bruit présent dans les données n'est pas connu. Ce résultat est très encourageant et permet d'envisager de futurs travaux.

Le critère d'optimisation  $\mathcal{C}_{actif}(x_{t+1})$  ne présente aucune limite théorique pour le traitement de données multi-classes caractérisées par plusieurs variables. En pratique, notre stratégie active est limitée par sa complexité algorithmique. Dans le cadre de nos expériences, plusieurs optimisations algorithmiques ont été mises en œuvre (Anonyme. (2008)) permettant de gagner trois ordres de grandeurs et un facteur constant d'environ 5000. Ces optimisations ne permettent pas en pratique de résoudre des problèmes d'apprentissage plus complexes car elles sont spécifiques au cas de données unidimensionnelles et aux modèles comportant au plus deux intervalles. Le parcours de l'ensemble des modèles de discrétisation lors de l'évaluation de l'espérance de  $P(M|D, x_{t+1})$  est une étape coûteuse en temps de calcul. Cette difficulté pourrait être contournée par l'échantillonnage de la famille de modèles  $\mathbb{M}$ . Des travaux futurs pourraient étendre notre stratégie active à des problèmes d'apprentissage plus complexes, par exemple en

exploitant une méthode du type “*Monte Carlo Markov Chain*” (Robert (2006)) pour échantillonner la famille de modèles  $\mathcal{M}$ .

Notre stratégie d'apprentissage actif pourrait être exploitée par d'autres méthodes d'apprentissage nécessitant elles aussi une dichotomie sur des données bruitées. Des techniques incrémentales permettent la construction d'arbres de décision en prenant en compte des données d'apprentissage qui s'enrichissent itérativement (Utgoff (1989)). Ces méthodes ne permettent pas la sélection des nouveaux exemples à étiqueter. Un arbre binaire actif pourrait être défini grâce au critère  $\mathcal{C}_{actif}$  présenté dans cet article. Le critère  $\mathcal{C}_{actif}$  désignerait, localement à chacune des feuilles, l'exemple qui maximise l'espérance de  $P(M|D, x_{t+1})$ . Des travaux futurs sur un arbre actif, exploitant notre stratégie d'apprentissage actif pourraient être réalisés.

## Références

- ANONYME. (2008). *Manuscrit de these du premier auteur*. Phd thesis, XXX.
- BONDU A., BOULLÉ M. & LEMAIRE V. (2008). A Non-parametric Semi-supervised Discretization Method. In *ICDM (International Conference on Data Mining)*, Pise.
- BOULLÉ M. (2006). MODL : A bayes optimal discretization method for continuous attributes. *Machine Learning*, **65**(1), 131–165.
- BRINKER K. (2003). Incorporating Diversity in Active Learning with Support Vector Machines. In *ICML : International Conference on Machine Learning*, p. 59–66.
- CASTRO R. & NOWAK R. (2008). *Foundations and Application of Sensor Management*, chapter Active Learning and Sampling. Springer-Verlag.
- CASTRO, R. AND WILLETT R. & NOWAK R. (2005). Faster rate in regression via active learning. In *NIPS (Neural Information Processing Systems)*, Vancouver.
- HORSTEIN M. (1963). Sequential decoding using noiseless feedback. In *IEEE Transactions Information Theory*, volume 9, p. 136–143.
- OSUGI T., KUN D. & SCOTT S. (2005). Balancing exploration and exploitation : A new algorithm for active machine learning. In *ICDM '05 : Proceedings of the Fifth IEEE International Conference on Data Mining*, p. 330–337, Washington, DC, USA : IEEE Computer Society.
- ROBERT C. (2006). *Le choix bayésien Principes et pratique*. Springer.
- ROY N. & MCCALLUM A. (2001). Toward optimal active learning through sampling estimation of error reduction. In *Proc. 18th International Conf. on Machine Learning*, p. 441–448 : Morgan Kaufmann, San Francisco, CA.
- SETTLES B. (2009). *Active Learning Literature Survey*. Computer Sciences Technical Report 1648, University of Wisconsin–Madison.
- SHANNON C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, **27**(3), 379–423.
- THRUN S. B. & MÖLLER K. (1992). Active exploration in dynamic environments. In J. E. MOODY, S. J. HANSON & R. P. LIPPMANN, Eds., *Advances in Neural Information Processing Systems*, volume 4, p. 531–538 : Morgan Kaufmann Publishers, Inc.
- UTGOFF P. (1989). Incremental Induction of Decision Trees. *Machine Learning*, **4**, 161–186.

- XU Z., AKELLA R. & ZHANG Y. (2007). Incorporating diversity and density in active learning for relevance feedback. In *ECIR (European Conference on Information Retrieval)*, volume 4425, p. 246–257.
- ZOLLER T. & BUHMANN M. (2000). Active learning for hierarchical pairwise data clustering. In *ICPR '00 : Proceedings of the 15th International Conference on Pattern Recognition*, p. 186–189.