

Apprentissage statistique pour le signal

Applications aux Interfaces Cerveau-Machine

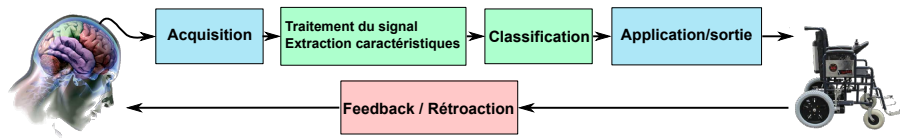
Rémi Flamary

LITIS EA 4108, INSA-Université de Rouen
76800 Saint Etienne du Rouvray, France

6 décembre 2011



Interfaces Cerveau-Machine



Objectifs [Dornhege et al., 2007]

1. Permettre une communication du cerveau vers la machine.
2. Mieux comprendre le fonctionnement du cerveau.

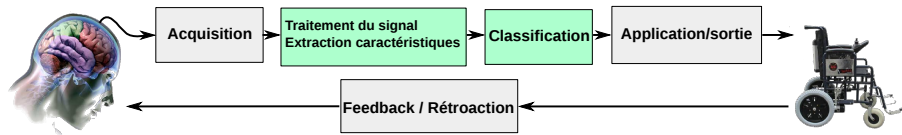
Applications

- ▶ Medium de communication pour des patients à handicap sévère.
- ▶ Contrôle d'un membre robotisé.
- ▶ Jeux vidéo, téléphone portable ...

Fonctionnement

- ▶ Association entre une tâche mentale et une commande.
- ▶ Le sujet peut donc contrôler la machine en effectuant certaines tâches mentales prédéfinies.

Travaux de thèse



Interfaces Cerveau-Machine

- ▶ Traitement du signal, filtrage.
- ▶ Extraction de caractéristiques.
- ▶ Classification.



- ▶ Sélection de paramètres.
- ▶ Sélection de capteurs.
- ▶ Minimisation du temps de calibration.

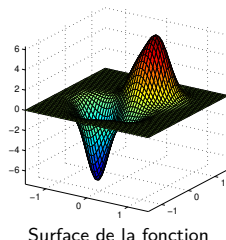
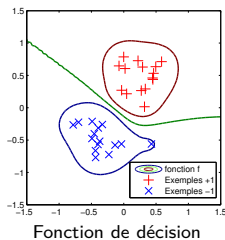
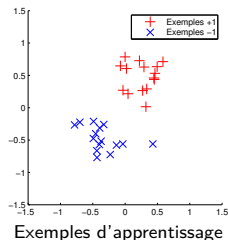
Apprentissage Statistique

- ▶ Séparateurs à vaste marge (SVM).
- ▶ Apprentissage de noyau.
- ▶ Apprentissage multitâche.



- ▶ Apprentissage de paramètres.
- ▶ Régularisation parcimonieuse.
- ▶ Transfert d'information.

Cadre d'apprentissage statistique



Classification supervisée

- ▶ $\{\mathbf{x}_i, y_i\}_{i=1, \dots, n}$ un ensemble d'apprentissage avec $\mathbf{x}_i \in \mathbb{R}^d$ et $y_i \in \{-1, 1\}$.
- ▶ Objectif : obtenir une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ permettant de prédire $y = \text{sgn}(f(\mathbf{x}))$.
- ▶ Minimisation du risque empirique.
- ▶ Régularisation pour éviter le surapprentissage.

Séparateurs à Vaste Marge (SVM)

Problème SVM

$$J_1(K) = \min_{f \in \mathcal{H}} \underbrace{\frac{C}{n} \sum_{i=1}^n H(y_i, f(\mathbf{x}_i))}_{\text{Attache aux données}} + \underbrace{\frac{1}{2} \|f\|_{\mathcal{H}}^2}_{\text{Régularisation}} \quad (1)$$

- ▶ $f(\cdot) \in \mathcal{H}$ et $\mathcal{H}$ un Espace de Hilbert à Noyau Reproductible $K(\cdot, \cdot)$
- ▶ $H(y, f(x)) = \max(0, 1 - y \cdot f(x))^p$ coût Hinge à la puissance $p = \{1, 2\}$.
- ▶ $\|f\|_{\mathcal{H}}$ norme induite par le produit scalaire dans $\mathcal{H}$.
- ▶ C paramètre de régularisation.

Théorème de représentation

La solution du problème d'optimisation (1) peut être mise sous la forme :

$$f(\cdot) = \sum_i \alpha_i y_i K(\cdot, \mathbf{x}_i) \quad (2)$$

avec $0 \leq \alpha_i$.

Sommaire

Introduction

- Interfaces Cerveau-Machine
- Apprentissage Statistique et SVM

Filtrage vaste marge

- Cadre ICM
- Filtrage Vaste Marge
- Expérimentations numériques
- Conclusions

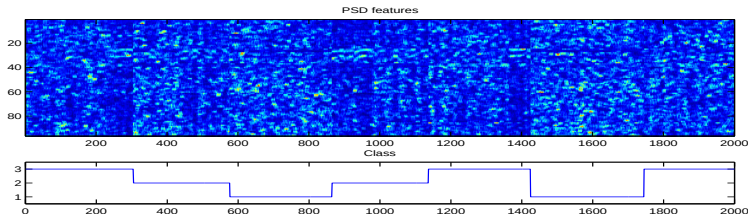
Apprentissage multitâche et applications ICM

- Cadre ICM
- Apprentissage multitâche parcimonieux
- Expérimentations numériques
- Adaptation aux ICM
- Conclusions

Conclusions et perspectives

Bibliographie

Problème ICM



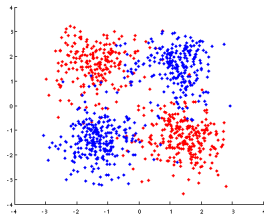
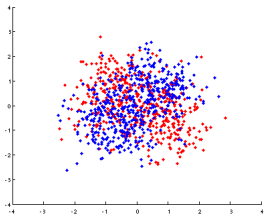
Cadre

- ▶ ICM asynchrones : classification continue d'états mentaux (mouvement main gauche, droite, ...).
- ▶ Signaux ou caractéristiques bruités.
- ▶ Bruit additif + signal convolué (délai, passe bas, ...).

Approches

- ▶ Problème d'étiquetage de séquence (HMM, CRF, ...).
- ▶ Fenêtrage (classification de fenêtres temporelles).
- ▶ Filtrage temporel (atténuation des bruits, filtrage des étiquettes).

Contraintes et motivations



Contraintes

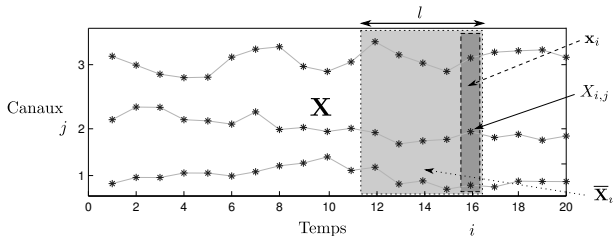
- ▶ Adaptation à une convolution (délai, passe bas).
- ▶ Décision rapide (pas de décodage hors ligne).
- ▶ Interprétabilité.
- ▶ Sélection de canaux ou de capteurs.



Filtrage Vaste Marge

- ▶ Convoluer pour mieux classifier.
- ▶ Classification d'échantillons filtrés.
- ▶ Apprentissage du filtre et du classifieur.
- ▶ Critère vaste marge.
- ▶ Cadre d'apprentissage de noyau.

Définitions



- ▶ $\mathbf{X} \in \mathbb{R}^{n \times d}$ signal de n échantillons $\mathbf{x}_i \in \mathbb{R}^d$.
- ▶ $\mathbf{y} \in \{-1, 1\}^n$ contient les étiquettes.
- ▶ $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_d] \in \mathbb{R}^{l \times d}$ contient d filtres de taille l .

Filtrage

- ▶ $\tilde{\mathbf{X}}$ signal filtré.

- ▶ Convolution :

$$\tilde{X}_{i,j} = \sum_{m=1}^l F_{m,j} X_{i+1-m,j}$$

Noyau gaussien sur échantillons filtrés

$$\begin{aligned} \tilde{K}_{i,j} &= K(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j) = \exp\left(-\frac{\|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|^2}{2\sigma^2}\right) \\ &= \exp\left(-\frac{1}{2\sigma^2} \sum_v^d ((X_{i,v} - X_{j,v}) * \mathbf{f}_v)^2\right) \end{aligned}$$

Filtrage vaste marge

Problème d'apprentissage

$$\min_{f, \mathbf{F}} \underbrace{\frac{1}{2} \|f\|^2 + C \sum_{i=1}^n H(y_i, f(\tilde{\mathbf{x}}_i))}_{\text{Problème SVM}} + \lambda \Omega(\mathbf{F}) \quad (3)$$

- ▶ λ un paramètre de régularisation et $\Omega(\cdot)$ un terme de régularisation.
- ▶ Problème d'apprentissage SVM avec $\mathbf{F}$ en paramètre.
- ▶ Le problème est convexe lorsque le filtrage $\mathbf{F}$ est fixé.

Formulation apprentissage de noyau

$$\min_{\mathbf{F}} \{J(\mathbf{F}) = J_1(\mathbf{F}) + \lambda \Omega(\mathbf{F})\} \quad (4)$$

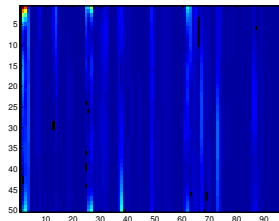
- ▶ $J_1(\mathbf{F})$ fonction de coût SVM, continue et différentiable.
- ▶ Quel type de régularisation pour $\mathbf{F}$?

Régularisation

Frobenius au carré

- ▶ Différentiable.
- ▶ Lissage du noyau gaussien.
- ▶ $F_{i,j}$ régularisés indépendamment.

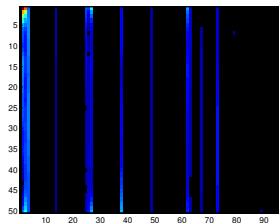
$$\Omega_F(\mathbf{F}) = \sum_{i,j} F_{i,j}^2$$



Norme mixte $\ell_1 - \ell_2$

- ▶ Non différentiable.
- ▶ $F_{i,j}$ régularisés par canaux.
- ▶ Sélection automatique de canaux.

$$\Omega_{1,2}(\mathbf{F}) = \sum_j \left(\sum_i F_{i,j}^2 \right)^{\frac{1}{2}}$$



Algorithmes d'optimisation

$$\min_{f, \mathbf{F}} \quad \frac{1}{2} \|f\|^2 + C \sum_{i=1}^n H(y_i, f(\tilde{\mathbf{x}}_i)) + \lambda \Omega(\mathbf{F})$$

Régularisation Frobenius

- ▶ Problème différentiable.
- ▶ Méthode de descente de Gradient Conjugué (KF-SVM).
- ▶ Résolution d'un problème SVM à chaque calcul de coût.
- ▶ Convergence vers un point stationnaire.

Régularisation $\ell_1 - \ell_2$

- ▶ Problème non différentiable.
- ▶ Méthode de Majoration-Minimisation ou IRLS (SKF-SVM).
- ▶ Résolution itérative d'un problème avec une Frobenius pondérée.
- ▶ Convergence vers un point stationnaire.

Données ICM

Méthode	Modèles linéaires							
	taille des filtres $l = 10$				taille des filtres $l = 50$			
	S1	S2	S3	Moy.	S1	S2	S3	Moy.
SVM	0.254	0.377	0.553	0.395	0.254	0.377	0.553	0.395
Avg-SVM	0.228	0.342	0.534	0.368	0.133	0.236	0.475	0.282
WinSVM	0.214	0.316	0.540	0.357	0.146	0.223	0.482	0.284
SWinSVM	0.215	0.314	0.470	0.333	0.146	0.218	0.460	0.274
KF-SVM	0.205	0.304	0.512	0.340	0.126	0.231	0.423	0.260
SKF-SVM	0.205	0.294	0.473	0.324	0.128	0.222	0.438	0.263

TABLEAU: Erreurs de test pour les données ICM (BCI Compet. III).

Données

- ▶ 3 sujets, 96 canaux, densités spectrales de puissance (PSD).
- ▶ 3 sessions d'apprentissage, 1 session de test.
- ▶ Sélection des paramètres par validation.

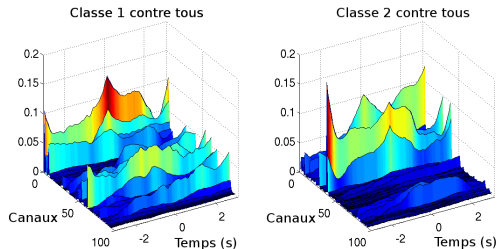
Résultats

- ▶ Apport de filtres longs..
- ▶ 3 meilleures : KF-SVM, SKF-SVM et SWinSVM.
- ▶ Nb canaux sélectionnés sur 96 : S1 : 13, S2 : 5, S3 : 1 .
- ▶ Performances similaires avec un noyau gaussien.

Interprétation des filtres

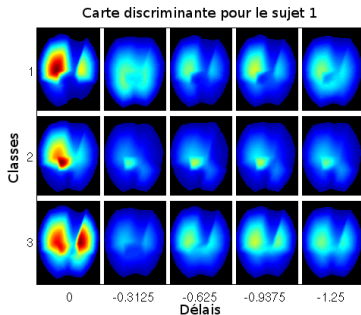
Visualisation des filtres

- Un canal : bande passante sur un capteur.
- Visualisation des instants discriminants.

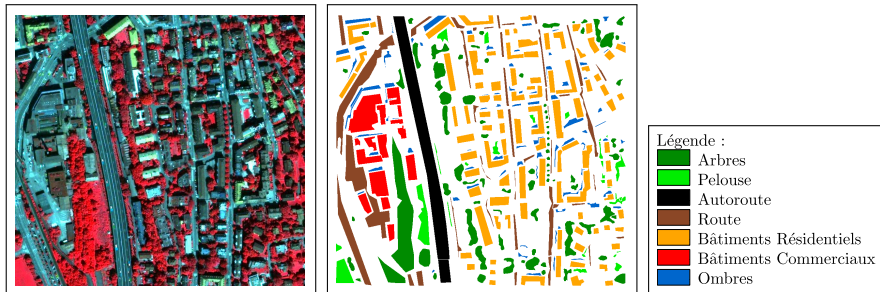


Carte discriminante

- Projection des filtres sur le scalp.
- Moyenne sur les bandes passantes.
- Visualisation similaire aux CSSSP [Dornhege et al., 2006].



Segmentation d'image



Classification de pixels

- ▶ Possibilité d'étendre le filtrage vaste marge à la 2D.
- ▶ Apprentissage d'un filtre de convolution 2D par canal (couleur).
- ▶ Image très haute résolution de Zurich.

Résultats en segmentation d'image

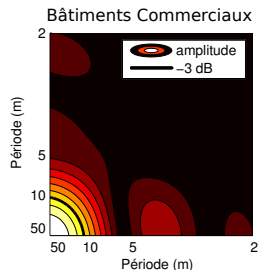
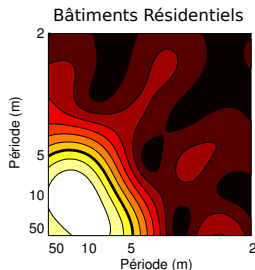
Reconnaissance

- Problème 7 classes, 5000 exemples en apprentissage.
- Meilleure discrimination des classes similaires (bâtiments commerciaux et résidentiels).

Méthode	Taux erreur	Kappa
SVM	0.249	0.685
AvgSVM	0.163	0.796
WinSVM	0.170	0.785
KF-SVM	0.147	0.816

Interprétation

- Amplitude de la TF des filtres pour la composante rouge.
- Filtres passe-bas.
- Bande passante différente selon de type de bâtiments.



Conclusions sur le filtrage vaste marge



Conclusions

- ▶ Discrimination d'échantillons.
- ▶ Apprentissage supervisé de filtres.
- ▶ Apprentissage de métrique de noyau.
- ▶ Sélection/pondération automatique des canaux.
- ▶ Interprétabilité (temporelle, spatiale, fréquentielle).
- ▶ Pré-traitement optimal utilisable avec d'autres méthodes (GMM, CRF).

Section

Introduction

Interfaces Cerveau-Machine

Apprentissage Statistique et SVM

Filtrage vaste marge

Cadre ICM

Filtrage Vaste Marge

Expérimentations numériques

Conclusions

Apprentissage multitâche et applications ICM

Cadre ICM

Apprentissage multitâche parcimonieux

Expérimentations numériques

Adaptation aux ICM

Conclusions

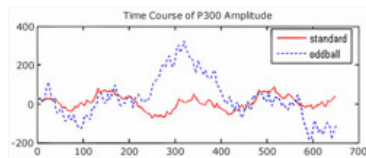
Conclusions et perspectives

Bibliographie

Problème ICM

Classification de signaux

- ▶ ICM synchrones.
- ▶ Reconnaissance de potentiels évoqués.
- ▶ Signaux apparaissant après un stimulus.



Clavier Virtuel P300

- ▶ Clignotement des lignes et colonnes du clavier.
- ▶ Apparition du Potentiel évoqué P300 lorsque la bonne lettre est illuminée.
- ▶ Reconnaissance du P300 permet au sujet d'épeler.

SIX(S)					
A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	1	2	3	4
5	6	7	8	9	-

Objectifs ICM

Minimisation du temps de calibration

- ▶ Apprendre classifieur unique pour tous les sujets [Fazli et al., 2009].
- ▶ Transfert d'information [Alamgir et al., 2010].

Sélection de capteurs

- ▶ Méthode Forward Backward [Rakotomamonjy and Guigue, 2008].
- ▶ Méthode de filtrage [Cecotti et al., 2011].
- ▶ Régularisation parcimonieuse [Tomioka and Müller, 2010].

Solution

- ▶ Apprentissage multitâche parcimonieux



Apprentissage multitâche

Problème d'apprentissage

On cherche T fonctions de décision $f_t(\cdot)$ en résolvant le problème suivant

$$\min_{f_1, \dots, f_T} C \underbrace{\sum_{t=1}^T \sum_{i=1}^{n_t} L(f_t(\mathbf{x}_{i,t}), y_{i,t})}_{\text{Attache aux données}} + \underbrace{\lambda \Omega(f_1, \dots, f_T)}_{\text{Régularisation}} \quad (5)$$

Avec

- ▶ $L(\cdot, \cdot)$ une fonction de perte.
- ▶ $\Omega(\dots)$ un terme de régularisation promouvant la similarité entre les tâches.

Termes de régularisation

- ▶ Variance des fonctions [Evgeniou and Pontil, 2004].
- ▶ Clustering des fonctions [Jacob et al., 2008].
- ▶ Parcimonie jointe [Obozinski et al., 2009].
- ▶ **Parcimonie jointe + variance [Rakotomamonjy et al., 2011]**

Apprentissage multitâche parcimonieux

Apprentissage de noyaux multiples

Fonctions de décision $f_t(\cdot)$ de la forme

$$f_t(\cdot) = \sum_{m=1}^M f_{t,m}(\cdot) + b_t \quad \forall t \in \{1, \dots, T\} \quad (6)$$

où $f_{t,m}(\cdot)$ sont des fonctions appartenant au EHNR $\mathcal{H}_m$ induit par le noyau $K_m(\cdot, \cdot)$.

Régularisation jointe

Nous utilisons dans nos travaux la régularisation suivante :

$$\Omega_{p,q}(f_1, \dots, f_T) = \sum_{m=1}^M \left(\sum_{t=1}^T \|f_{t,m}\|_{\mathcal{H}_m}^q \right)^{p/q} \quad (7)$$

qui est une norme mixte $\ell_p - \ell_q$ à la puissance p , avec $0 < p \leq 1$ et $1 \leq q \leq 2$.

Formulation variationnelle du problème $\ell_1 - \ell_q$

► $s > 0$ et $\|f_{t,k}\| = \|f_{t,k}\|_{\mathcal{H}_k}$ tel que au moins un $\|f_{t,k}\| > 0$.

► Alors

$$\min_{\{d_{t,m}\}} \left\{ \sum_{k,m} \frac{\|f_{t,m}\|^2}{d_{t,m}} : d_{t,m} \geq 0, \sum_k \left(\sum_t d_{t,m}^{1/s} \right)^s \leq 1 \right\} = \left(\sum_m \left(\sum_t \|f_{t,m}\|^q \right)^{1/q} \right)^2$$

où $q = \frac{2}{s+1}$.

► À l'optimum nous avons :

$$d_{t,m}^* = \frac{\|f_{t,m}\|^{\frac{2s}{s+1}} \left(\sum_u \|f_{u,m}\|^{\frac{2}{s+1}} \right)^{\frac{1-s}{2}}}{\sum_v \left(\sum_u \|f_{u,v}\|^{\frac{2}{s+1}} \right)^{\frac{s+1}{2}}}. \quad (8)$$

Nous savons donc résoudre le problème pour des fonctions $f_{t,m}$ fixées.

Résolution du problème $\ell_1 - \ell_q$

Algorithme alterné

- ▶ Si $\mathbf{d}$ est fixé, alors problème SVM avec le noyau :

$$K(\mathbf{x}, \mathbf{x}') = \sum_k d_{t,m} K_m(\mathbf{x}, \mathbf{x}') \quad (9)$$

- ▶ Minimisation à $\mathbf{d}$ fixé pour obtenir les $f_t(\cdot)$ puis minimisation à $f_t(\cdot)$ fixés pour obtenir $\mathbf{d}$.
- ▶ Convergence vers un minimum global.

Algorithme Explicite Implicite Multipas [Beck and Teboulle, 2009]

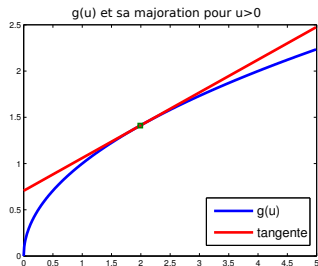
- ▶ Si les fonctions de décision sont linéaires.
- ▶ Opérateur proximal pour le terme de régularisation non différentiable.
- ▶ Algorithme utilisable directement sur le problème d'apprentissage multitâche parcimonieux.
- ▶ Convergence rapide en valeur objectif sous certaines conditions.

Résolution du problème $\ell_p - \ell_q$

Terme de régularisation

$$\Omega_{p,q}(\cdot) = \sum_{m=1}^M g(\|f_{\cdot,m}\|_q)$$

avec $\|f_{\cdot,m}\|_q = (\sum_t \|f_{t,m}\|_{\mathcal{H}_m}^q)^{1/q}$ et
 $g(u) = u^p, u > 0$ avec $p < 1$.



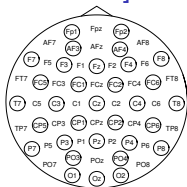
Algorithme Majoration-Minimisation

- ▶ Avec $p < 1$, solution plus parcimonieuse mais problème non convexe.
- ▶ Le terme $\Omega_{p,q}(\cdot)$ avec $p < 1$ est concave.
- ▶ Majoré dans son quadrant positif par sa tangente.
- ▶ Résolution itérative d'un problème $\ell_1 - \ell_q$ pondéré.

Expérimentations numériques

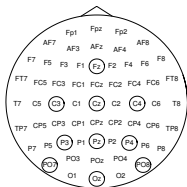
Données P300 EPFL [Hoffmann et al., 2008]

- ▶ P300 pour sélectionner parmi 3×2 images.
- ▶ 8 sujets (une tâche par sujet).
- ▶ 32 électrodes.



Données P300 UAM [Ledesma-Ramirez et al., 2010]

- ▶ Clavier virtuel classique 6×6 .
- ▶ 30 sujets (une tâche par sujet).
- ▶ 10 électrodes.



Données P300 EPFL

Méthode	AUC	# Var	p-val
SepSVM	75.6 ± 0.8	256	-
Sep ℓ_1 SVM	73.4 ± 1.3	118 $\pm$ 30	0.00006
Full SVM	67.1 ± 0.6	256	0.00006
Full ℓ_1 SVM	67.0 ± 0.6	238 $\pm$ 7	0.00006
MTL _{1,2}	76.5 $\pm$ 0.6	191 $\pm$ 26	0.00019
MTL _{p,2}	76.1 $\pm$ 0.5	134 $\pm$ 33	0.00897
MTL _{1,q}	76.5 $\pm$ 0.6	201 $\pm$ 23	0.00014

Méthodes

SepSVM SVM indépendants par tâche.

Full SVM SVM unique pour toutes les tâches.

MTL Apprentissage multitâche parcimonieux.

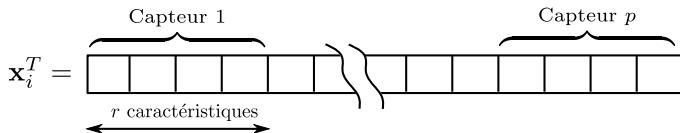
Protocole

- Sélection des paramètres par validation.
- Performances en AUC et nombre de variables sélectionnées.
- Test de Wilcoxon pour comparer avec SVM.

Résultats

- MTL statistiquement meilleur.
- Parcimonie plus agressive avec $p < 1$.
- Pas de sélection de canaux.

Applications aux ICM



Jusqu'à maintenant...

- ▶ Sélection de **variables**.
- ▶ Similarité entre tâches sous la forme de **parcimonie jointe**.

Adaptations

- ▶ Exemples structurés (groupes de variables).
- ▶ Sélection de capteurs.
- ▶ Similarité entre tâches [Alamgir et al., 2010].

Prédiction linéaire

- ▶ Fonction pour la tâche t :

$$f_t(\mathbf{x}) = \mathbf{w}_t^T \mathbf{x} + b_t$$

avec $\mathbf{w}_t \in \mathbb{R}^d$ et $b_t \in \mathbb{R}$.

- ▶ Terme de régularisation MTL s'applique sur

$$\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_T] \in \mathbb{R}^{d \times T}$$

Régularisation multitâche pour ICM

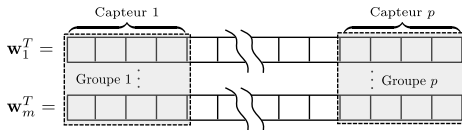
Nous avons proposé le terme suivant :

$$\Omega_{\text{mtl}}(\mathbf{W}) = \underbrace{\lambda_r \sum_{g \in \mathcal{G}'} \|\mathbf{w}_g\|_2}_{\text{Norme mixte}} + \underbrace{\lambda_s \sum_{t=1}^T \|\mathbf{w}_t - \hat{\mathbf{w}}\|_2^2}_{\text{Similarité}} \quad (10)$$

où λ_r et λ_s pondèrent les termes de régularisation.

Norme mixte

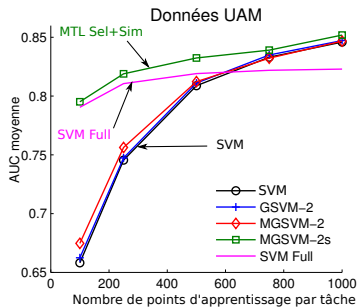
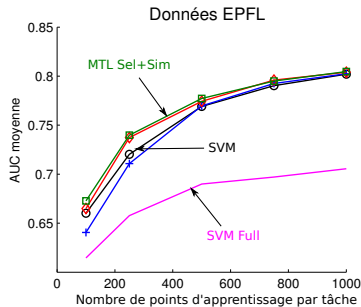
$\mathcal{G}'$ contient des groupes de variables de $\mathbf{W}$:



Similarité

- ▶ $\hat{\mathbf{w}} = \frac{1}{T} \sum_t \mathbf{w}_t$ est la normale à l'hyperplan du classifieur moyen.
- ▶ Minimisation de la variance des classifieurs [Evgeniou and Pontil, 2004].

Expérimentations numériques



Données

- ▶ P300 EPFL : choix d'image, 8 sujets.
- ▶ P300 UAM : clavier virtuel, 30 sujets.
- ▶ Variation du nombre de points d'apprentissage par tâche.

Résultats

- ▶ AUC Moyenne pour les tâches.
- ▶ Apport de la similarité entre tâches (peu d'exemples).
- ▶ Apport de l'adaptation au sujet (>800 exemples).

Conclusions sur l'apprentissage multitâche et applications ICM



Conclusions

- ▶ Apprentissage multitâche dans un cadre d'apprentissage de noyaux.
- ▶ Algorithmes efficaces de résolution du problème d'apprentissage.
- ▶ Norme mixte $\ell_p - \ell_q$ permet plus de souplesse en terme de parcimonie.
- ▶ Interprétabilité des noyaux/variables sélectionnés.
- ▶ Adaptation à la classification de potentiels évoqués.
- ▶ Gain important de performances lorsque peu d'exemples d'apprentissage sont disponibles.

Section

Introduction

Interfaces Cerveau-Machine

Apprentissage Statistique et SVM

Filtrage vaste marge

Cadre ICM

Filtrage Vaste Marge

Expérimentations numériques

Conclusions

Apprentissage multitâche et applications ICM

Cadre ICM

Apprentissage multitâche parcimonieux

Expérimentations numériques

Adaptation aux ICM

Conclusions

Conclusions et perspectives

Bibliographie

Conclusions et perspectives

Approche méthodologique

- ▶ Motivation pratique (ICM).
- ▶ Théorie (régularisation $\ell_p - \ell_q$).
- ▶ Algorithmes (repondération).
- ▶ Expérimentations numériques.

Bilan

- ▶ Publications en revues (IEEE TNN, IEEE TSP).
- ▶ Compétitions internationales (BCI Competition IV, MLSP 2011).
- ▶ Recherche reproductible (code source).

Perspectives

- ▶ Représentation discriminante des données.
- ▶ Apprentissage de métrique (exemples, tâches).
- ▶ Problèmes de non stationnarité, bruit de classification.



Section

Introduction

Interfaces Cerveau-Machine

Apprentissage Statistique et SVM

Filtrage vaste marge

Cadre ICM

Filtrage Vaste Marge

Expérimentations numériques

Conclusions

Apprentissage multitâche et applications ICM

Cadre ICM

Apprentissage multitâche parcimonieux

Expérimentations numériques

Adaptation aux ICM

Conclusions

Conclusions et perspectives

Bibliographie

References

- [Alamgir et al., 2010] Alamgir, M., Grosse-Wentrup, M., and Altun, Y. (2010).
Multi-task Learning for Brain-Computer Interfaces.
In AI & Statistics.
- [Beck and Teboulle, 2009] Beck, A. and Teboulle, M. (2009).
A fast iterative shrinkage-thresholding algorithm for linear inverse problems.
SIAM Journal on Imaging Sciences, 2 :183–202.
- [Cecotti et al., 2011] Cecotti, H., Rivet, B., Congedo, M., Jutten, C., Bertrand, O., Maby, E., and Mattout, J. (2011).
A robust sensor-selection method for P300 brain–computer interfaces.
Journal of Neural Engineering.
- [Dornhege et al., 2006] Dornhege, G., Blankertz, B., Krauledat, M., Losch, F., Curio, G., and Muller, K. (2006).
Optimizing spatio-temporal filters for improving brain-computer interfacing.
Advances in Neural Information Processing Systems, 18 :315.
- [Dornhege et al., 2007] Dornhege, G., Millán, J., Hinterberger, T., McFarland, D., and Müller, K. (2007).
Toward brain-computer interfacing, volume 74.
Mit Press Cambridge, MA.
- [Evgeniou and Pontil, 2004] Evgeniou, T. and Pontil, M. (2004).
Regularized multi-task learning