
The Forgetron: A Kernel-Based Perceptron on a Fixed Budget

I) Problème général

Cet article présente et analyse un algorithme d'apprentissage en ligne appelé FORGETRON.

Le problème rencontré lors de l'implémentation de ce type d'algorithme est la taille mémoire exigée pour la sauvegarde des hypothèses de l'apprentissage qui peuvent grandir d'une façon rapide, le problème devient éminent lorsqu'il s'agit des appareils autonomes de petite taille tel que les téléphones portable.

II) Méthode présentée

FORGETRON est composé de deux mots Perceptron et Forget le perceptron est utilisé pour les mises à jour et les plus anciens exemples sont oubliés (supprimés)

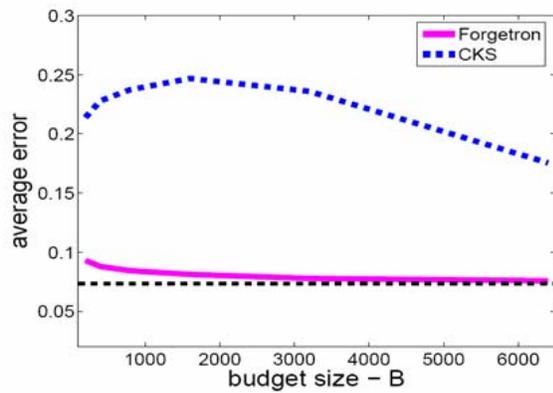
Le FORGETRON est appliqué en deux étapes :

- la première consiste à utiliser l'algorithme de Perceptron pour construire les mises à jours.
- Dans le cas où $|\mathbf{It}| \leq \mathbf{B}$ (*tg B entier positif, permet de paramétrer la taille mémoire*) on ne fait rien et on saute la seconde étape dans le cas contraire (c-a-d $|\mathbf{It}| > \mathbf{B}$) on applique la seconde étape qui se fait à son tour en deux pas
 - o *Shrinking* consiste à trouver et maximiser ce qu'on appelle le coefficient de shrinking $\phi_t \in (0, 1)$ pour tenter de minimiser voire éliminer le dommage causé par l'étape de suppression
 - o Suppression, supprime l'exemple le plus ancien.

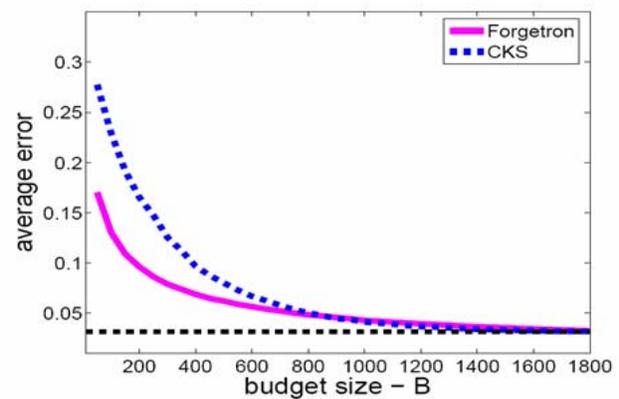
III) Résultats

Les deux expérimentations citées dans l'article montrent que le Forgetron se comporte mieux que l'algorithme CKS (Suppression des exemples redondants) surtout pour une valeur de \mathbf{B} particulièrement petite

Les deux expérimentations sont réalisées sur deux ensembles de données en l'occurrence MNIST 60 000 exemples et census-income 200 000 exemples.



Census-income Dataset



MNIST Dataset

IV) Etat de l'art

L'utilisation du perceptron en conjonction avec Kernel pour régler le problème de classification en ligne n'est pas une idée nouvelle, plusieurs travaux dans le domaine ont été déjà réalisés la majorité de ces travaux utilise le même principe en général, un ensemble d'exemples est observé à chaque étape, la fonction de classification est construite incrémentalement en sauvegardant un sous ensemble des exemples observés dans leur mémoires internes, parce que l'ensemble des exemples peut croître d'une façon rapide on risque d'avoir besoin de plus de mémoire que celle physiquement disponible.

Crammer, canola et singer [Online classification budget, 2003] par exemple ont tenté de régler ce problème par la suppression des exemples redondants.

V) Perspectives

Les perspectives évoquées par les auteurs correspondent surtout à employer l'analyse présentée dans cet article pour dériver une famille des algorithmes en ligne dont le Forgetron est seulement un cas spécial