

Conclusion

Un des meilleurs livres en théorie de l'apprentissage est à mon sens le bouquin "A theory of learning and generalization", de M. Vidyasagar. Outre l'état de l'art très complet que l'auteur y présente, on y trouve des directions de recherche à mon avis intéressantes: notamment, le contrôle (voir "Is it possible to introduce explicitly a notion of 'time' into the learning problem formulation?" (Open Problem 1 dans le livre cité ci-dessus), et son chapitre sur le contrôle) et le processus empirique (il écrit "...to make a connection between PAC learning theory and the theory of empirical processes. While there have been a few other papers that followed up this connection, my opinion is that by and large this connection remains unexplored, or perhaps merely unexplained to a wide audience..." dans l'introduction, une réponse partielle se trouve dans mes chapitres sur l'apprentissage non-identiquement distribués (les modèles déformables et le recalage d'images), l'apprentissage Bayésien et l'apprentissage non-indépendant (le contrôle)). Ses problèmes ouverts 1 et 5 sont étudiés dans ma thèse (l'apprentissage non-indépendant ayant conduit à définir une deuxième notion combinatoire, facteur de la VC-dimension dans les complexités d'échantillon, dépendant de la "vitesse" de l'ergodicité uniforme, et la loi limite des déviations étant évaluée par les résultats sur les classes de Donsker - résultat classique en processus empirique, malheureusement sous une forme peu maniable).

En conclusion de cette thèse, quelques commentaires sur les domaines que j'aurais aimé étudier pour être plus complet. Tout d'abord, des manques dans les sujets traités:

- La partie statistique de l'apprentissage a vu ses grandes questions à peu près résolues, quoique manquant sans doute d'un formalisme unifié pour résumer les diverses branches, rarement confrontées, qui traitent de ces problèmes. Ma partie sur l'apprentissage à VC-dimension infinie cherche à entamer une telle unification.
- Le parallélisme, sur lequel j'ai travaillé sans prendre le temps de me soucier de la valorisation du travail réalisé.
- Le contrôle. Il y a des milliards de trucs à faire, et ça me tente bien. Je me suis limité à la théorie par manque de temps, mais je vais m'y consacrer bien à fond dans les années à venir si j'en ai l'occasion.
- La complexité de l'apprentissage. Beaucoup de problèmes intéressants en apprentissage sont NP-complets. Voir dans quelle mesure on peut approximer rapidement des solutions est capital. Et malheureusement passé sous silence au profit des bornes statistiques qui attirent plus souvent les théoriciens.
- L'implémentation c'est agréable, utile, et j'en ferais bien. En fait j'ai implémenté un tas de trucs divers, mais je n'ai pas encore réalisé une synthèse de tout ça en un cadre facile à manier. Une bibliothèque commode à bricoler m'apparaît un point utile à avoir sous la main.

et ensuite juste quelques remarques sur les manques en général dans la littérature, dans des domaines que je n'ai pas traités du tout:

1. Nous manquons sans doute de résultats d'approximation permettant de justifier des résultats empiriques, comme le fait que les réseaux de neurones sigmoïdaux permettent dans un grand nombre de cas de résoudre des problèmes avec un plus petit nombre de paramètres que d'autres méthodes (notamment par rapport aux espaces de fonctions dont les applications sont linéaires en les paramètres - RBF, SVM). Je n'ai malheureusement pas la moindre idée d'hypothèses générales permettant de justifier de tels faits empiriques énoncés par des praticiens.
2. L'introduction d'a priori me semble avoir beaucoup d'avenir, même si cette introduction a souvent répugné les adeptes de formalisme. Contrairement à Cherkassky qui a expliqué à Icann 2001 en lecture invitée qu'il n'y avait pas de salut en apprentissage en dehors de la minimisation du risque structurelle, je crois que les théorèmes notamment négatifs montrent qu'au pire cas on ne peut espérer mieux que la consistance universelle, et que l'introduction d'a priori me semble une excellente façon de justifier des méthodes efficaces en pratique.
3. Les réseaux de neurones (notamment récurrents dans le cas du contrôle) semblent un fort bel outil, notamment lorsque les nombres d'exemples sont élevés. La rétropropagation m'apparaît encore, malgré les années et les paradigmes "révolutionnaires" qui s'accumulent, un très bon moyen algorithmique de fouiller de vastes espaces de fonctions. Tenter des applications un peu "folles" m'apparaît intéressant, en dépit du peu d'intérêt théorique de telles choses. Les problèmes de contrôle semblent un bon moyen

d'utiliser les réseaux de neurones pour un grand nombre de tâches originales et utiles (notamment, peut-être, dans le cas d'un signal cible dynamiquement fourni par l'utilisateur - ce qu'on appelle la neuro-interface). En plus faut pas boudier son plaisir, c'est bien la théorie, mais c'est vraiment sympa de faire des applis rigolotes bien visuelles bien concrètes bien fun, et puis c'est moins crispant que les théorèmes sur les classes de Donsker (que j'aime bien aussi, hein).

4. A mon sens les articles actuels qui traitent de nouvelles méthodes pour encoder des données qui ne rentrent pas dans les cadres classiques discrets/continus/vectoriels/..., par exemple arbres, fichiers graphiques, langages, sont très importants.
5. Peut-être des versions hardware tirant parti de connexions 3D (pour le moment on simule toujours, à bas-niveau, sur des circuits en gros 2D), si des physiciens ou informaticiens trouvent des solutions pour faire ça.
6. Peut-être aussi que la modularité permettrait un essor des applications...