

Une étude des algorithmes de construction d'architecture des réseaux de neurones multicouches.

Norbert Tsopzé^{1**} Engelbert Mephu Nguifo*
Gilbert Tindo**

*CRIL-CNRS, IUT de Lens, SP 16 Rue de l'Université 62307 Lens Cedex
{tsopze,mephu}@cril.univ-artois.fr

**Département d'Informatique - Université de Yaoundé I BP 812 Yaoundé
tsopze.norbert@gmail.com, gtindo@uycdc.uninet.cm

Résumé. Le problème de choix d'architecture d'un réseau de neurones multicouches reste toujours très difficile à résoudre dans un processus de fouille de données. Ce papier recense quelques algorithmes de recherche d'architectures d'un réseau de neurones pour les tâches de classification. Il présente également une analyse théorique et expérimentale de ces algorithmes. Ce travail confirme les difficultés de choix des paramètres d'apprentissage (modèle, nombre de couches, nombre de neurones par couches, taux d'apprentissage, algorithme d'apprentissage,...) communs à tout processus de construction de réseaux de neurones et les difficultés de choix de paramètres propres à certains algorithmes.

1 Introduction

Un réseau de neurones est un ensemble de neurones interconnectés qui communiquent entre eux et avec l'extérieur. Un réseau de neurones se présente comme un graphe où les noeuds sont les différentes unités de réseau et les arcs représentent les connexions entre ces unités. Le nombre de couches, le nombre de neurones par couche et les interconnexions entre les différentes unités du réseau définissent l'architecture (encore appelée topologie) de celui-ci. Un neurone peut être appelé unité ou cellule. Comme tout système d'apprentissage supervisé, les systèmes d'apprentissage supervisé à base des réseaux de neurones fonctionnent en deux phases : la phase d'apprentissage qui consiste à construire à partir des observations (exemples présentés sous forme (x, y) où y représente l'observation de la fonction f en x) un système capable d'approximer la fonction f dont l'expression analytique n'est pas facile à trouver ; la phase de classement qui utilise le modèle construit en phase d'apprentissage pour produire des décisions (prédire un nouvel exemple qui ne faisait pas partie des observations de la base d'apprentissage). Définir la structure du réseau pour de tel système n'est pas une tâche évidente (J.Han et Hamber, 2001; A.Cornuéjols et Miclet, 2002). En effet, il n'existe aucune méthode permettant de définir et de justifier la structure d'un réseau de neurones (J.Han et Hamber, 2001).

¹Le Service de Coopération et d'Action Culturelle (SCAC) de l'ambassade de France à Yaoundé (Cameroun) a financé le séjour du premier auteur au CRIL pendant la réalisation de ce travail. Ce travail est partiellement financé par le ministère français des affaires étrangères.