

Les réseaux de neurones artificiels

Un outil de traitement de données prometteur pour l'anthropologie

Artificial neural network. A relevant data processing tool for anthropology

A. Schmitt, B. Le Blanc, M.-M. Corsini, C. Lafond et J. Bruzek



Édition électronique

URL : <http://journals.openedition.org/bmsap/4463>

ISSN : 1777-5469

Éditeur

Société d'Anthropologie de Paris

Édition imprimée

Date de publication : 1 juin 2001

ISSN : 0037-8984

Référence électronique

A. Schmitt, B. Le Blanc, M.-M. Corsini, C. Lafond et J. Bruzek, « Les réseaux de neurones artificiels », *Bulletins et mémoires de la Société d'Anthropologie de Paris* [En ligne], 13 (1-2) | 2001, mis en ligne le 14 janvier 2010, consulté le 01 mai 2019. URL : <http://journals.openedition.org/bmsap/4463>

LES RÉSEAUX DE NEURONES ARTIFICIELS
UN OUTIL DE TRAITEMENT DE DONNÉES PROMETTEUR POUR L'ANTHROPOLOGIE
ARTIFICIAL NEURAL NETWORKS
A RELEVANT DATA PROCESSING TOOL FOR ANTHROPOLOGY

Aurore SCHMITT ¹, Benoit LE BLANC ², Marc-Michel CORSINI ²
Caroline LAFOND ¹, Jaroslav BRUZEK ¹

RÉSUMÉ

Les réseaux de neurones artificiels sont très peu utilisés en anthropologie biologique, alors que leur application est en plein essor dans de nombreuses disciplines. Ils constituent, entre autre, une alternative intéressante aux statistiques traditionnelles pour le traitement des données. Cette note explique les notions clés du fonctionnement des RNA et plus particulièrement du perceptron multi-couches et décrit quelques applications analogues à certaines problématiques de l'anthropologie biologique.

Mots-clés : traitement des données, réseaux de neurones artificiels, perceptron, prédiction, classification.

SUMMARY

The use of artificial neural networks (ANN) has grown in the last few years in many fields of research, but not in biological anthropology. They are a relevant alternative to current statistics for data processing. This note provides a background for understanding the notion of ANN and describes also few applications close to biological anthropological topics.

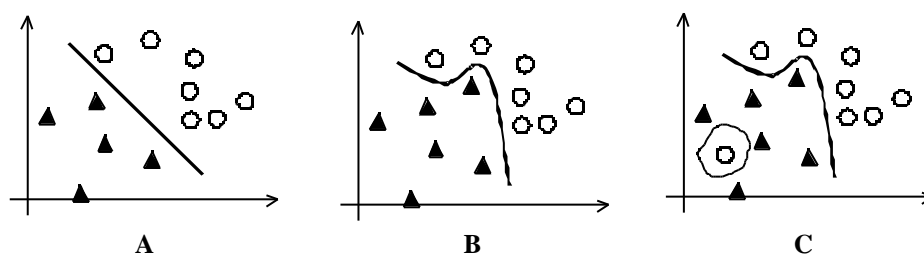
Keywords: data processing, artificial neural networks, perceptron, prediction, classification.

INTRODUCTION

Depuis une dizaine d'années, l'utilisation des réseaux de neurones artificiels (RNA) s'est développée dans de nombreuses disciplines (sciences économiques, écologie et environnement, biologie et médecine...). Ils sont notamment appliqués pour résoudre des problèmes de classification, de prédiction, de catégorisation, d'optimisation, de reconnaissance des formes et de mémoire associative (Drew et Monson, 2000).

-
1. UMR 5894 du CNRS, Laboratoire d'Anthropologie des Populations du Passé, Université Bordeaux 1, Avenue des Facultés, 33405 Talence.
 2. EA 487, Laboratoire de Sciences Cognitives, Université Victor Segalen Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux cedex.

Dans le cadre du traitement des données, les RNA constituent une méthode d'approximation de systèmes complexes, particulièrement utile lorsque ces systèmes sont difficiles à modéliser à l'aide des méthodes statistiques classiques. Les RNA sont également applicables dans toutes les situations où il existe une relation non linéaire entre une variable prédictive et une variable prédite. Par leur nature et leur fonctionnement, les RNA peuvent détecter les interactions multiples non linéaires parmi une série de variables d'entrée, ils peuvent donc gérer des relations complexes entre les variables indépendantes et les variables dépendantes (*fig. 1*).



*Fig. 1 - Discrimination entre les paramètres « triangle » et « rond ».
Discriminations linéaire (A), non linéaire (B) et non continue (C).*

Avec les méthodes statistiques traditionnelles, il faut d'abord penser à un modèle, le tester, penser à un autre modèle jusqu'à ce qu'on obtienne un modèle suffisamment précis. Un réseau est entraîné sur des données grâce à un mécanisme d'apprentissage qui agit sur les constituants du réseau pour réaliser au mieux la tâche souhaitée. La famille de RNA la plus utilisée ces dernières années comme outil d'aide à la décision est le perceptron multi-couches (PMC).

Même si on peut fournir des exemples d'application des RNA dans des disciplines aussi nombreuses que variées, l'utilisation la plus proche des recherches anthropologiques est l'amélioration des diagnostics en sciences médicales. En effet, les RNA sont utilisés, par exemple, pour la prévision des infarctus (Baxt, 1990), pour le diagnostic des pathologies pulmonaires (Patil *et al.*, 1993), du diabète (Armoni, 1998), des cancers (Han *et al.*, 2001), de la maladie d'Alzheimer (Hamilton *et al.*, 1997), des lésions crâniennes post-traumatiques (Arana *et al.*, 1998), etc. Le diagnostic clinique et la prévision des maladies se font à partir de l'analyse de symptômes et de facteurs de risque. Or, l'interaction entre ces éléments est souvent complexe. Un mauvais diagnostic peut conduire au décès de l'individu. Il est donc nécessaire de pouvoir évaluer le risque avec le plus de précision possible. Les RNA sont de plus en plus utilisés dans cette discipline comme outil d'aide à la décision (Baxt, 1995).

Dans le domaine de l'anthropologie, les problématiques de prédiction sont courantes. Or, le rapport entre les variables est rarement linéaire. L'interaction entre les paramètres, même si elle est soupçonnée, est difficile à modéliser. L'utilisation des RNA et particulièrement des perceptrons multi-couches présente une alternative intéressante aux statistiques traditionnelles. L'objectif de cette note n'est pas de décrire le fonctionnement détaillé des réseaux de neurones, les articles et ouvrages sur ce sujet étant nombreux (Davallo et Naïm, 1990 ; Bourret *et al.*, 1991 ; Touzet, 1992 ; Abdi, 1994 ; Jodouin, 1994 ; Cross *et al.*, 1995 ; Penny et Frost, 1996). Nous souhaitons montrer l'intérêt de cet outil d'aide à la décision pour l'anthropologie biologique, domaine dans lequel il est, à ce jour, très peu exploité.

QU'EST-CE QU'UN RÉSEAU DE NEURONE ARTIFICIEL ?

Les réseaux de neurones fonctionnent en répartissant les valeurs des variables dans des automates (les *neurones*). Ces unités sont chargées de combiner entre elles leurs informations pour déterminer la valeur du paramètre de discrimination. C'est de la connexion de ces unités entre elles qu'émerge la capacité de discrimination du RNA.

Chaque neurone reçoit des informations numériques en provenance de neurones voisins ; à chacune de ces valeurs est associée un *poids* représentatif de la force de la connexion. Chaque neurone effectue localement un calcul dont le résultat est transmis ensuite aux neurones aval.

La famille de réseau majoritairement employé est le *perceptron multi-couches* (PMC). À lui seul ce type de réseau recouvre plus de 95 % des applications scientifiques et industrielles. Il comporte quelques dizaines à quelques centaines de neurones dans les cas usuels, voir plusieurs milliers pour les applications graphiques.

Le PMC est un modèle de *réseau à propagation par couche* (fig. 2). Les neurones y sont organisés en couches successives : une couche d'entrée, une couche de sortie et entre les deux une ou plusieurs couches intermédiaires, appelées aussi *couches cachées*. Il n'existe pas de connexion entre les neurones d'une même couche, en revanche tout neurone d'une couche est connecté à tous les neurones de la couche suivante. La « couche » d'entrée n'est pas une réelle couche de neurones car elle se contente de coder les variables d'observation. La couche de sortie code la variable de discrimination. Les valeurs d'activité des neurones sont propagées dans le réseau, de l'entrée vers la sortie, sans retour arrière. La présence d'une couche cachée permet de modéliser des relations non linéaires entre les entrées et la sortie. En théorie une seule couche cachée suffit, mais le fait de disposer d'une seconde couche cachée permet de modéliser plus facilement une fonction de discrimination non continue. En pratique, la plupart des problèmes sont résolus avec un ou deux niveaux, trois au maximum.

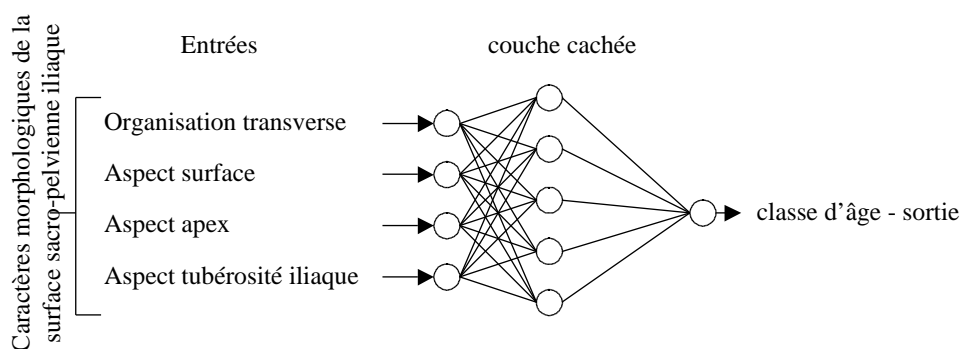


Fig. 2 - Schéma d'un perceptron multi-couche illustrant l'estimation de l'âge au décès à partir de l'observation de critères osseux de la surface sacro-pelvienne iliaque. Les entrées correspondent à l'observations de critères morphologiques sur la surface sacro-pelvienne iliaque, la sortie est la prédiction de l'âge au décès.

L'objectif général d'un RNA est de trouver la configuration des poids de connexion entre neurones pour qu'il associe à chaque configuration d'entrée, une réponse adéquate. L'utilisation d'un RNA se fait en deux temps. Tout d'abord une phase d'*apprentissage* qui est chargée d'établir des valeurs pour chacune des connexions du réseau, puis une phase d'*utilisation* proprement dite, où l'on présente au réseau une entrée et où il nous indique en retour « sa » sortie calculée.

Dans le cas du PMC, on utilise un *apprentissage supervisé*. Les valeurs des poids de connexion sont créées tout d'abord au hasard et le système cherche par itérations successives à obtenir une modélisation des données. A chaque étape, une entrée est présentée au réseau, il propage ces valeurs vers les neurones de sortie. Cette sortie calculée est comparée avec la réponse attendue et le système modifie les poids en conséquence. Cette altération des connexions est obtenue par l'*algorithme de rétropropagation du gradient d'erreur*. Ce calcul est chargé de rétropropager dans le réseau les erreurs constatées sur les sorties. En théorie, on ne peut jamais être sûr que cet algorithme finisse par déterminer un ensemble de poids convenable pour tous les couples d'entrées-sorties. En pratique, on ne construit pas un seul RNA, mais plusieurs modèles en jouant sur les paramètres de cet algorithme, et en cherchant à obtenir un modèle qui colle au mieux aux données.

Dans le cas où son apprentissage est *non supervisé*, le réseau doit trouver par lui-même les régularités sous-jacentes des données.

QUELQUES APPLICATIONS

À l'heure actuelle, peu de travaux appliquent les PMC dans le domaine de l'Anthropologie. Cependant, il existe quelques études à l'interface de cette discipline dans lesquelles on rencontre des problématiques analogues.

Le premier groupe d'exemples concerne l'utilisation des PMC en taxinomie. Culverhouse et ses collaborateurs (1994) ont tenté de distinguer plusieurs espèces de protozoaires en élaborant un PMC. Le réseau réalisé a donné de meilleurs résultats que les techniques courantes de catégorisation des organismes marins. Marcondes et Borges (2000) ont élaboré un PMC pour différencier les mâles de plusieurs espèces de diptères. Ils ont effectué un test sur un échantillon indépendant de 58 spécimens. Seulement 2 spécimens ont été mal discriminés. Ces exemples de classification s'apparentent dans leur problématique aux travaux effectués en paléoanthropologie pour différencier des groupes ou taxons fossiles, ainsi que des populations archéologiques ou actuelles, par des particularités de forme et de format.

Le deuxième groupe d'exemple est l'estimation de l'âge au décès des adultes à partir du squelette. Cette estimation est un problème épineux tant en anthropologie qu'en médecine légale. En effet, les indicateurs osseux ou dentaires évoluent avec l'âge selon une multitude de facteurs qui interagissent de façon complexe. La relation entre l'âge chronologique et les indicateurs n'est donc pas linéaire. C'est la raison pour laquelle l'estimation de l'âge au décès est actuellement un véritable échec. De plus, la plupart des techniques de prédiction utilisées ne tiennent pas compte de cette variabilité et sont donc des sources d'erreurs supplémentaires. Amariti et ses collègues (2000) ont tenté d'améliorer l'estimation de l'âge au décès à partir de la formation de la dentine secondaire en élaborant un PMC. L'erreur moyenne est de l'ordre de 8 ans avec le PMC, alors qu'elle est de 11 ans avec une régression linéaire.

L'estimation de l'âge au décès des adultes avec le PMC a également été faite à partir d'indicateurs osseux (Schmitt, 2001) en tenant compte de la variabilité de la sénescence et en évitant les écueils méthodologiques systématiques. Ce travail a privilégié la fiabilité de l'estimation au détriment de la précision par l'attribution des spécimens à une classe d'âge. L'étude prospective sur le potentiel du PMC, a donné des résultats intéressants. Il permet de distinguer les individus appartenant aux classes d'âge extrêmes, soit la classe 20-29 et les plus de 60 ans. L'identification de ces classes d'âge est primordiale dans les études paléobiologiques. En effet, les méthodes classiques d'estimation de l'âge sous-estiment systématiquement l'effectif des individus de plus de 60 ans et conduisent à des interprétations erronées sur la mortalité des populations du passé. De nombreuses publications affirment que peu ou aucun individu n'atteignait l'âge de 60 ans dans ces populations. La reconnaissance des individus âgés est donc un

apport majeur aux études paléobiologiques, qu'il s'agisse de l'analyse du recrutement funéraire ou de l'épidémiologie.

Un domaine de recherche très courant en anthropologie du vivant est la détermination de l'âge biologique des sujets au cours de la croissance à partir de radiographies du poignet. La méthode la plus commune consiste à comparer, sur radiographie, le stade de maturation de la totalité des os du poignet de l'individu étudié avec des radiographies de référence qui englobent la totalité des phases de maturation. Le PMC a été utilisé comme alternative à cette méthode (Gross *et al.*, 1995). Les résultats sont aussi performants que ceux obtenus par la méthode classique, mais ne sont pas meilleurs. Cependant, les variables d'entrée du PMC sont des mesures alors que la méthode classique se base sur la comparaison de radiographies.

Une méthode souvent utilisée pour la classification en anthropologie est l'analyse discriminante. Il est évident que dans les applications où cette technique apporte des résultats satisfaisants (détermination du sexe à partir du bassin), il n'est pas utile de recourir au PMC. Néanmoins, il existe des situations où les taux de classement ne sont pas très élevés, notamment en ce qui concerne la discrimination entre sexe à partir de dimensions prises sur les os longs. Pour illustrer l'utilisation du PMC comme alternative aux fonctions discriminantes, nous pouvons citer l'étude sur la classification de l'origine géographique de certaines variétés de pommes de terre (Anderson *et al.*, 1999). La comparaison entre les résultats obtenus avec une analyse discriminante et le PMC montre que ce dernier est plus performant. Cet exemple est loin des problématiques anthropologiques, cependant il est intéressant puisqu'il s'agit d'un problème de classification classique.

LES RÉSEAUX DE NEURONES ARTIFICIELS EN PRATIQUE

Tout comme pour les traitements statistiques, les données doivent avoir été recueillies sur des cas représentatifs de la population étudiée et la problématique doit être posée en termes clairs.

En ce qui concerne la classification ou la prédiction de valeurs, l'apprentissage des PMC se fait de façon supervisée, la variable de classement ou la variable à prédire doivent être connus. Dans le cas de l'estimation de l'âge au décès à partir du squelette, les collections à observer sont celles pour lesquelles on possède cette information. Il existe cependant d'autres problématiques, telles que la catégorisation à laquelle correspond d'autres types de RNA.

En effet, l'alternative à l'architecture en couches du PMC est un réseau où les flux d'information ne suivent pas une direction privilégiée. Ce type de réseau est appelé

réseau bouclé, dont un cas particulier est celui des *réseaux totalement connectés*. Dans ce cas, on ne cherche plus à connaître des valeurs de « sortie » du réseau, mais à stabiliser globalement les activités individuelles de chacun des neurones. Les réseaux de Hopfield ou encore les cartes auto-organisatrices de Kohonen (Kohonen, 1990) en sont des illustrations. Les réseaux de Kohonen, ont permis, par exemple, la reconnaissance des variables significatives dans les séquences génomiques des grandes banques de données internationales.

L'élaboration d'un RNA, quelle que soit son architecture, requiert une certaine expérience. C'est la raison pour laquelle la collaboration avec des spécialistes est la démarche la plus souhaitable, que ce soit pour superviser l'utilisation de logiciels simples, tels que SPSS®, MATLAB®, STATISTICA® qui proposent des modules de RNA, ou pour l'élaboration de RNA à partir de logiciels plus sophistiqués. L'application des réseaux de neurones artificiels étant en plein essor dans tous les domaines de la recherche, cette interdisciplinarité est facilement envisageable.

Les réseaux de neurones artificiels constituent donc un outil de traitement des données qui, au même titre que les statistiques traditionnelles, peut s'appliquer dans le domaine de l'anthropologie biologique, les relations complexes entre variables ne manquant pas dans ce domaine. Ils ne constituent pas le remède miracle, mais ils sont une alternative possible aux approches statistiques courantes qui ne sont pas toujours adaptées à nos problématiques.

BIBLIOGRAPHIE

- ABDI (H.) 1994, *Les réseaux de neurones*, Presses Universitaires de Grenoble, Grenoble.
- AMARITI (M.-L.), RESTORI (M.), DE FERRARI (F.), PAGANELLI (C.), FAGLIA (R.), LEGNANI (G.) 2000, A histological procedure to determine dental age, *Journal of Forensic Odontostomatology* 18: 1-5.
- ANDERSON (K.A.), MANGNUSON (B.A.), TSCHIRGI (M.L.), SMITH (B.) 1999, Determining the geographic origin of potatoes with trace metal analysing using statistical and neural network classifiers, *Journal of Agricultural and Food Chemistry* 47: 1568-1575.
- ARMONI (A.) 1998, Use of neural networks in medical diagnosis, *Medical Diagnosis Computing* 15 : 100-104.
- ARANA (E.), MARTI-BONMATI (L.), PAREDES (R.), BAUTISTA (D.) 1998, Comparison of logistic regression and neural network models, *Investigative Radiology* 33 : 738-745.
- BAXT (W.G.) 1990, Use of an artificial neural network for data analysis in clinical decision making: the diagnosis of acute coronary occlusion, *Neural computation* 2: 480-489.

- BAXT (W.G.) 1995, Application of artificial neural networks to clinical medicine, *The Lancet* 346: 1135-1138.
- BOURRET (P.), REGGIA (J.), SAMUELIDES (M.) 1991, *Réseaux de neurones artificiels : une approche connexionniste de l'Intelligence Artificielle*, Teknea.
- CROSS (S.S.), HARRISON (R.F.), KENNEDY (R.L.) 1995, Introduction to neural networks, *The Lancet* 346 : 1075-1079.
- CULVERHOUSE (P.F.), ELLIS (R.), SIMPSON (R.G.), WILLIAMS (R.), PIERCE (R.W.), TURNER (J.T.) 1994, Automatic categorisation of five species of *Cymatocylys* (Protozoa, Tinitinnida) by artificial neural network, *Marine Ecology progress series* 107: 273-280.
- DAVALO (E.), NAÏM (P.) 1990, *Des réseaux de neurones*, Eyrolles, Paris.
- DREW (P.J.), MONSON (J.R.T.) 2000, Artificial neural networks, *Surgery* 127: 3-11.
- GROSS (G.W.), BOONE (J.M.), BISHOP (D.M.) 1995, Pediatric skeletal age: determination with neural networks, *Radiology* 195: 689-695.
- HAMILTON (D.), O'MAHONY (D.), COFFEY (J.), MURPHY (J.), O'HARE (N.), FREYNE (P.), WALSH (B.), COAKLEY (D.) 1997, Classification of mild Alzheimer's disease by artificial neural network analysis of SPET data, *Nuclear Medicine Communications* 18: 805-810.
- HAN (M.), SNOW (P.B.), BRANDT (J.M.), PARTIN (A.W.) 2001, Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma, *Cancer* 91(S8): 1661-1666.
- JODOUIN (J.F.) 1994, *Les réseaux de neurones, principes et définitions*, Hermès, Paris.
- KOHONEN (T.) 1990, The self-organizing map. *Proceedings of the IEEE* 9: 1464-1479.
- MARCONDES (C.B.), BORGES (P.) 2000, Distinction of males of the *Lutzomyia intermedia* (Lutz & Neiva, 1912) species complex by ratios between dimensions and by an artificial neural network (Diptera: Psychodidae, Phlebotominae), *Memorias do Instituto Oswaldo Cruz* 95: 685-688.
- PATIL (S.), HENRY (J.W.), RUBENFVIE (M.), STEIN (P.D.) 1993, Neural Network in the clinical diagnosis of acute pulmonary embolism, *Chest* 104: 1685-1689.
- PENNY (W.), FROST (D.) 1996, Neural networks in clinical medicine, *Medical Decision Making* 16: 386-398.
- SCHMITT (A.) 2001, *Variabilité de la sénescence du squelette humain. Réflexions sur les indicateurs de l'âge au décès : à la recherche d'un outil performant*, Thèse de doctorat, Université Bordeaux 1, inédit.
- TOUZET (C.) 1992, *Les réseaux de neurones artificiels. Introduction au connexionnisme*, EC2 Ed., Nanterre.