



INTERVIEW

Entrevue avec Sylvain Barthélémy et Laurent Bougrain

Sylvain Barthélémy¹, Laurent Bougrain²

By/Par

Les questions de cette interview sont posées à Sylvain Barthélémy et Laurent Bougrain par les organisateurs du colloque international « Les applications des réseaux de neurones en économie, finance, management et environnement », Chtourou Nouri, Feki Rochdi et Bazin Damien. Ce colloque, organisé par l'Ecole Supérieure de Commerce de Sfax, s'est déroulé en juin 2011 aux Iles Kerkennah, en Tunisie.

INTERVIEW

Ethique et économique (E&E.) :- Que sont les réseaux de neurones et à quoi peuvent-ils servir en économie ?

Laurent Bougrain (B.L.) :- Les réseaux de neurones artificiels sont un ensemble d'outils de modélisation comprenant deux grandes familles : les réseaux supervisés comme le perceptron multicouches, utilisés pour la régression et le classement et les réseaux non supervisés comme les cartes auto-organisatrices de Kohonen, utilisés pour la classification. L'idée sous-jacente des réseaux de neurones supervisés est de reprendre la logique des modèles linéaires bien connus des économistes (estimer une relation entre variables descriptives et variables à prédire), mais en introduisant des paramètres et fonctions intermédiaires, permettant de capter les non linéarités. En effet, les modèles de régression et de discrimination basés sur des combinaisons linéaires de fonctions de base fixes ont des propriétés calculatoires et analytiques mais leur utilisation pratique est limitée aux phénomènes linéaires et par la malédiction de la dimensionnalité. Afin d'appliquer de tels modèles à des problèmes de grande taille, il est nécessaire d'adapter les fonctions de base aux données. La solution adoptée par les réseaux de neurones artificiels est de fixer le nombre de fonctions de base (représentées dans les neurones cachés) mais de les choisir paramétrables et d'adapter

¹ Sylvain Barthélémy, économiste et spécialiste des réseaux de neurones, est directeur exécutif de TAC, cabinet français de recherche appliquée en économie et finance (sylvain.barthelemy@tac-financial.com).

² Laurent Bougrain, maître de conférences en informatique à l'Université Nancy 1, est spécialiste de la fouille de données numériques au sein de l'équipe Cortex (laurent.bougrain@loria.fr).

leurs paramètres durant l'apprentissage en fonction des données. Les cartes auto-organisatrices quant à elles, établissent en plus d'un regroupement par similitude des individus, une relation de voisinage entre les groupes. Cette particularité permet de projeter les données dans un espace à une ou plus généralement à deux dimensions (en fonction de la relation de voisinage prédéfinie dans la carte) et donc de visualiser facilement les distances entre les individus. Enfin, pour les deux familles, un mode d'apprentissage itératif permet de suivre l'évolution de la distribution des données dans le temps sans avoir à réapprendre le modèle sur la totalité des données.

Sylvain Barthélémy (B.S.) :- La théorie économique moderne nous enseigne que les équilibres multiples et plus généralement les non linéarités sont très présents en économie. Des études économétriques ont aussi montré que le simple fait d'agréger des indicateurs induit des non linéarités dans les séries de données macroéconomiques, non linéarités qui ont longtemps été négligées et dont les conséquences pratiques peuvent pourtant être majeures. Comme le précise Laurent Bougrain, les réseaux de neurones ont la capacité de bien capter ces non linéarités dans les données, et ceci sans avoir à faire d'hypothèse sur la nature des relations qui peuvent exister entre les variables. Ils peuvent à ce titre se révéler être un outil précieux pour l'économiste qui souhaite mieux comprendre les interactions micro ou macro-économiques. Les cartes auto-organisatrices sont par exemple un outil précieux pour l'économiste qui souhaite établir des typologies sur la base de données quantitatives. Nos équipes ont notamment travaillé dans ce sens dans le cadre d'une étude réalisée en 2001 pour le Ministère de l'Economie et des Finances Français et dans laquelle une première typologie des « très petites entreprises » a été établie sur la base de la théorie des mondes de production de Salais et Storper et de cartes de Kohonen mises en place en combinant des données comptables et des données d'enquête. Mais c'est une application parmi d'autres et des réseaux de neurones peuvent aussi être utilisés pour tenter de prévoir par exemple les crises financières, analyser et prévoir les tensions sur les prix du pétrole, faire de l'économie expérimentale, analyser les similarités dans les caractéristiques de développement des pays, mieux modéliser les fonctions d'utilité et améliorer les modèles de psychologie cognitive, etc...

(E&E) : - Comment considérez-vous l'apport des réseaux de neurones en lien avec les techniques statistiques « conventionnelles » ? Existe-t-il des passerelles entre les deux approches ?

(B.S.) : - L'apport des réseaux de neurones en statistique est majeur, mais pour l'instant assez limité en économie. Plusieurs raisons à cela. La première est qu'ils ne sont généralement pas enseignés aux économistes et qu'ils ont encore l'image d'outils « complexes » destinés aux spécialistes de l'intelligence artificielle. La seconde est que le traitement statistique des « ruptures » n'est pas absent de l'économétrie moderne, qui dispose déjà de bon nombre de techniques permettant d'aborder les non linéarités, comme par exemple au travers des modèles à changement de régime markoviens ou des modèles STAR. Mais les économistes travaillent dans un environnement complexe, avec généralement peu de données historiques, et bien

souvent des imprécisions dans les mesures et des données manquantes. Tous ces éléments créent un climat d'incertitude statistique où règne avant tout le doute sur la validité des relations établies. Même de simples relations comme la courbe de Phillips ou la théorie de la parité des pouvoirs d'achat font encore couler beaucoup d'encre aujourd'hui. Pourquoi ? En grande partie parce qu'elles semblent se vérifier dans certains cas seulement, et pas dans d'autres. Bien sûr, les réseaux de neurones n'apporteront pas de solution statistique miracle et ne sont pas le Graal de l'économètre. Mais ils pourront capter ces phénomènes non linéaires de façon relativement simple, et illustrer les interactions économiques complexes plus clairement que l'économétrie traditionnelle.

(B.L.) : - Jusqu'à il n'y a pas si longtemps, les deux communautés avaient le sentiment d'être éloignées pour des raisons d'une part de différences dans la terminologie utilisée (les statisticiens raisonnent plus en termes de variables descriptives et de modèles alors que les connexionnistes parlaient d'entrées/sorties et de systèmes dynamiques) et d'autre part d'applications visées (analyse socio-économique avec des modèles explicites versus reconnaissance de forme performante). Des travaux ont montré que la régression multiple s'apparentait au perceptron simple et que l'analyse discriminante était un cas particulier du perceptron multicouches. De plus, les perceptrons multicouches sont souvent utilisés pour estimer les probabilités a posteriori sans avoir à estimer les densités. D'autre part, les centres des classes des cartes auto-organisatrices de Kohonen convergent vers le barycentre des classes si la fonction de voisinage est omise ou très restreinte.

(E&E) : - Parmi la multitude de modèles fondés sur les approches de l'intelligence artificielle tels que les réseaux de neurones, quelles sont les mieux adaptés aux problématiques posées par les économistes qui cherchent à expliquer et prévoir un certain nombre de phénomènes sur lesquels le nombre d'observations est généralement limité ?

(B.S.) : - Il y a effectivement une grande variété de modèles disponibles et non pas comme on le pense souvent, un seul type de réseaux de neurones. Ces dernières années nous avons avec nos équipes beaucoup travaillé sur les modèles quantitatifs de signalement avancé des crises économiques et financières dans les pays en voie de développement, notamment sur la base des travaux initiés par Kaminsky, Lizondo et Reinhart. Ces auteurs ont identifié au milieu des années 90 un certain nombre d'indicateurs macroéconomiques clés permettant de prévoir ces crises économiques et financières dans les pays en voie de développement. Cependant, de simples combinaisons linéaires de ces variables ne suffisent pas à prévoir les crises suffisamment à l'avance, et le modélisateur doit avoir recours à des outils plus performants, permettant de capter des signaux de rupture, de dépassement de seuil et de combinaisons d'indicateurs particulièrement vulnérables. C'est la direction que nous avons pris, et les réseaux de neurones (en particulier les cartes auto-organisatrices), les arbres de décision, « random forest » et SVM ont été des atouts majeurs pour identifier des « patterns » de développement plus risqués ou plus vulnérables que d'autres (patterns

que nous n'avions pas pu identifier avec des méthodes plus traditionnelles, comme par exemple des analyses linéaires discriminantes). Il faut ajouter que la capacité de la plupart des réseaux de neurones de travailler avec des données manquantes et/ou des ruptures dans les historiques est aussi un avantage majeur qui facilite grandement la tâche de l'économiste du développement, qui doit gérer en permanence des bases de données incomplètes.

(B.L.) : - Dernièrement les machines à vecteurs supports, aussi appelés séparateurs à vastes marges (SVM), ont montré des propriétés intéressantes car ils permettent une modélisation non linéaire et la fonction d'erreur est convexe. De plus, leur complexité dépend non du nombre de variables mais du nombre d'individus. Comparativement aux perceptrons multicouches, les fonctions de bases, c'est-à-dire de vecteurs supports, sont non paramétriques mais leur nombre varie en fonction de la complexité de la tâche. Ils sont en particulier adaptés à la discrimination à deux classes. Les arbres de décision présentent eux aussi un intérêt car ils sont capables généralement de traiter des variables quantitatives et qualitatives. De plus, les prédictions sont explicites puisque le modèle se présente sous la forme d'un arbre où chaque branche correspond à une règle du type $SI X1 = \dots$ et $SI X2 > \dots$ et $SI \dots$. Alors X appartient à la classe B.

(E&E) : - Quelles sont les principales difficultés techniques posées par l'utilisation des réseaux de neurones ?

(B.L.) : - Les perceptrons multicouches sont souvent qualifiés d'approximateurs universels de fonctions. Cependant, il est parfois délicat de les paramétrer puisque la fonction d'erreur est non convexe. Il est aussi nécessaire d'empêcher le sur-apprentissage c'est-à-dire que le modèle ne restitue très bien que les données apprises. Ce phénomène apparaît lorsque le modèle est surdimensionné par rapport à la complexité du problème. Plusieurs solutions existent. On peut vérifier les performances du modèle sur un ensemble de nouvelles données et ce pour un nombre variable de neurones cachés. Il est également possible d'utiliser un modèle surdimensionné, capable donc de modéliser une grande famille de fonctions, mais de pénaliser l'apparition de poids forts car ils génèrent une plus grande variabilité du modèle. On ajoute un terme de pénalisation et on parle de régularisation. Une autre solution est le « early stopping », qui stoppe l'algorithme avant d'atteindre les meilleures performances possibles. L'apprentissage est arrêté quand les performances diminuent sur un ensemble de nouvelles valeurs, signe que le modèle se spécialise non plus sur le problème mais sur les données d'apprentissage. Il perd ses capacités de généralisation qui nous intéressent. Rappelons que le problème du sur-apprentissage n'est pas spécifique aux réseaux de neurones mais il prend une dimension importante car la variabilité du réseau est plus grande que pour un modèle linéaire.

(B.S.) : - Il est vrai que les réseaux de neurones posent un certain nombre de difficultés techniques, dont certaines peuvent être plus particulièrement gênantes pour les économistes. Le sur-apprentissage, dont vient de parler Laurent Bougrain, fait

effectivement partie de ces difficultés. Outre les solutions présentées plus haut, c'est une question particulièrement difficile à résoudre en économie, car découper un échantillon de données historiques en sous-ensembles afin d'éviter un sur-apprentissage (et éviter ainsi que le réseau de neurone ne soit pas généralisable) n'est pas toujours chose facile. L'économie mondiale est en effet animée par des fluctuations ou cycles économiques de fréquences plus ou moins régulières, cycles qu'il faudra veiller à bien répartir entre les échantillons (en particulier « les pics et les creux ») afin d'éviter l'estimation d'un modèle de prévision de l'inflation en récession pour réaliser des prévisions en phase d'expansion (questions que l'on retrouve en économétrie avec les difficultés que posent les chocs asymétriques). Il ne faut pas oublier non plus que l'économétrie traditionnelle pose elle aussi un certain nombre de difficultés au statisticien et à l'économiste, comme par exemple les hypothèses sur les lois de probabilité sous-jacentes des indicateurs utilisés dans les estimations économétriques, les difficultés posées par les trends stochastiques et l'éventuelle stationnarisation des séries temporelles, la nécessité d'avoir des données sur des historiques très longs pour obtenir des résultats fiables, ce qui est rarement possible en macroéconomie (en dehors de quelques pays OCDE).

(E&E) : - On remarque parfois certaines réticences concernant l'utilisation des réseaux de neurones, essentiellement sur l'aspect « boîte noire » de ces techniques. Habités aux techniques statistiques classiques, le chercheur en sciences sociale se trouve confronté à certains problèmes (interpréter les paramètres, effectuer des tests de robustesse, ...). Pensez vous, qu'avec le développement actuel des techniques d'apprentissage et de la puissance des calculs, l'argument « boîte noire » peut être dépassé ?

(B.S.) : - Effectivement, le fait que la plupart des réseaux de neurones soient vus comme des « boîtes noires » rebute bon nombre d'économistes. Mais même s'il est vrai qu'il reste encore aujourd'hui toujours difficile de bien comprendre ce qui se passe lors du fonctionnement d'un réseau de neurone, un certain nombre de méthodes permettent néanmoins de réduire cet effet « boîte noire ». Parmi ces techniques, on peut citer la logique floue ou les « ensembles » de réseaux de neurones, qui permettent par exemple de mieux tester et anticiper la sensibilité des réseaux à certains chocs sur les variables d'entrée. Enfin, il convient de préciser que tous les réseaux de neurones ne présentent pas ces problèmes d'opacité du modèle, et que les réseaux de Kohonen permettent au contraire de mieux comprendre les structures de données et les éventuels groupes de données qui peuvent « se cacher » à l'intérieur d'une base de données.

(B.L.) : - Les connaissances dans ce domaine sont effectivement insuffisantes. Il est possible d'obtenir de bons modèles mais il est très difficile de comprendre la prédiction du modèle. Il est cependant possible d'appliquer des méthodes de sélection de variables soit en-dehors du modèle (approche enveloppante) soit au sein du modèle (approche embarquée). Par exemple, d'autres algorithmes que l'algorithme de descente du gradient de l'erreur permettent d'ajuster les paramètres (ou poids synaptiques) et de déterminer lesquels augmentent le moins l'erreur en cas d'absence. Par extension, les

variables descriptives pondérées par ces paramètres sont estimées moins utiles et supprimées.

(E&E) : - Comment peut-on envisager vaincre les difficultés qui bloquent la diffusion de l'utilisation des réseaux de neurones en économie et en sciences sociales en général ?

B.L. : Certains choix ne sont pas intuitifs pour le néophyte : combien de couches cachées utiliser et combien de neurones attribuer à chaque couche cachée, quelle valeur prendre pour le coefficient d'apprentissage... Un peu de pratique et plusieurs essais avec différentes valeurs permettent d'appivoiser les réseaux de neurones artificiels.

(B.S.) : - Aux difficultés techniques précisées par Laurent Bougrain, je pense que deux autres difficultés pratiques ralentissent l'adoption des réseaux de neurones par les économistes. La première est la très mauvaise image des réseaux de neurones auprès des économistes et plus particulièrement des macroéconomistes, qui y voient une sorte de « boule de cristal » dont on ne comprendrait pas le fonctionnement. Cette image ne correspond pas à la réalité de ces outils (ou de leur dérivés) et un très grand nombre de nouvelles méthodes de « data mining » plus ou moins intensives en calcul sont très prometteuses et permettent d'aller bien plus loin que l'économétrie traditionnelle dans l'exploration des données. Je pense que des domaines de la macroéconomie tels que l'économie du développement sont tout à fait adaptés à ce type de méthodes. La seconde difficulté généralement évoquée concerne plus spécifiquement les économètres. Ils ont en effet l'habitude d'utiliser des logiciels dans lesquelles les modules permettant d'estimer des réseaux de neurones ne sont malheureusement pas présents (EViews, TSP, Rats, Fair, Ox, ...). Mais la popularité croissante d'outils ouverts tels que R, très populaire auprès de la nouvelle génération d'économistes, devrait je pense changer rapidement les choses.

(E&E) : - Comment voyez-vous l'avenir de l'utilisation des réseaux de neurones en général et en sciences sociales en particulier ?

(B.L.) : - Les chercheurs s'intéressent désormais à de nouvelles familles de réseaux comme le réservoir computing (réseaux récurrents à l'architecture fixe mais aléatoire), le deep neural networks (réseaux avec de nombreuses couches), etc. Les réseaux de neurones classiques de part leur performance méritent qu'on leur prête attention. Toutefois je pense que trop peu d'ouvrages sur leur utilisation pratique existent. Malgré cela, le bouche à oreille devrait fonctionner même s'il est un peu plus lent semble-t-il dans les sciences sociales pour les raisons évoquées précédemment. Les réseaux de neurones devraient faire partie des outils de modélisation de tout analyste en connaissant leurs qualités et leurs défauts.

(B.S.) : - L'avenir des réseaux de neurones et l'avenir des réseaux de neurones en sciences sociales sont deux choses très différentes. Pour ce qui est de l'avenir des réseaux de neurones et algorithmes assimilés, ils ont à mon avis de beaux jours devant eux et dans bien des domaines scientifiques comme la santé, bien sûr, mais aussi le marketing, la défense, les télécommunications, les transports, la finance, la robotique, etc... En

revanche, leur adoption par les chercheurs en sciences sociales mettra je pense du temps, car leur objectif est avant tout de « mieux comprendre » les interactions entre les agents et non d'en reproduire quantitativement le fonctionnement à l'aide de réseaux de neurones. C'est pour cette raison que les économistes évoquent si souvent la difficulté posée par l'aspect « boîte noire » des réseaux de neurones et préfèrent bien souvent estimer une simple équation avec les moindres carrés ordinaires plutôt que d'utiliser un réseau de neurones un peu complexe... même si les propriétés statistiques des variables sont rarement vérifiées (leur normalité par exemple). La publication de résultats de recherches obtenus via des cartes de Kohonen, des SVM, des arbres de classification d'autres d'algorithmes similaires sera déterminante pour l'adoption de ces méthodes par la communauté dans le futur. Ils donneront je l'espère non seulement des idées d'application de ces outils en économie, mais ouvriront de nouveaux champs d'exploration pratiques qui étaient jusque-là impensables sans les puissances de calculs actuelles.