

Étude de la Performance des Compagnies Pétrolières Nationales via leurs Ratios Financiers : Classification par Réseaux de Neurones Et Classification Automatique (Année 2000)

Mme Moussi Oum El Kheir

Introduction :

Jusqu'en 1960, Les grandes compagnies pétrolières contrôlaient totalement la chaîne d'approvisionnement pétrolier (de la production à la distribution). Le pétrole était produit principalement par les majors dans le cadre des concessions : les compagnies disposaient librement du pétrole en échange du versement d'une redevance et d'impôts à l'état propriétaire du sol. Le pétrole était ensuite transporté et distribué par les autres filiales de la compagnie. A partir des années 70, les pays producteurs prirent progressivement, par le biais, des nationalisations (exemple de l'Algérie en 1971) le contrôle de la production pétrolière. La chaîne pétrolière est fragmentée et de nouvelles habitudes commerciales entrent en jeu. Le système de concessions est remplacé par des relations contractuelles liant les nouvelles compagnies nationales qui ont émergé après les nationalisations (SONATRACH en Algérie) et les anciens concessionnaires.

De nouveaux venus, qui sont donc les compagnies nationales(PEMEX du Mexique, PDVSA du Venezuela, PETROBRAS du Brésil), se sont développées et sont apparues sur la scène pétrolière et traitent avec les grands groupes internationaux qu'on appelle les Majors (Esso (Exxon), Mobil, Chevron, Texaco, Shell)..

Le but de ce travail est d'étudier la performance des compagnies pétrolières nationales et internationales, via leurs ratios financiers, en utilisant deux méthodes différentes :

- La première «classification par les nuées dynamiques», une des les méthodes de la classification automatique. Ce sont des méthodes qui consistent à fournir à partir d'un ensemble d'individus une partition en classes aussi homogènes que possibles.
- La seconde méthode est non paramétrique: « la classification par réseaux de neurones » qui consiste à estimer une fonction de classification pour le partitionnement de l'ensemble des individus en sous ensembles homogènes.
-

La question posée est de savoir si, à l'aide d'un ensemble de ratios financiers observés sur une population d'entreprises, on peut arriver à différencier les compagnies non performantes des compagnies performantes ? Et si une compagnie nationale est nécessairement une compagnie non performante?

Notre échantillon est constitué de 181 compagnies pétrolières opérant dans l'amont ; l'aval ; ou l'électricité ; la majorité étant des compagnies privées. Etant donné que les compagnies étatiques n'ont pas l'obligation de publier leurs bilans annuels contrairement aux compagnies privées,leur nombre est restreint dans cette base, comme l'illustre la figure si dessous:

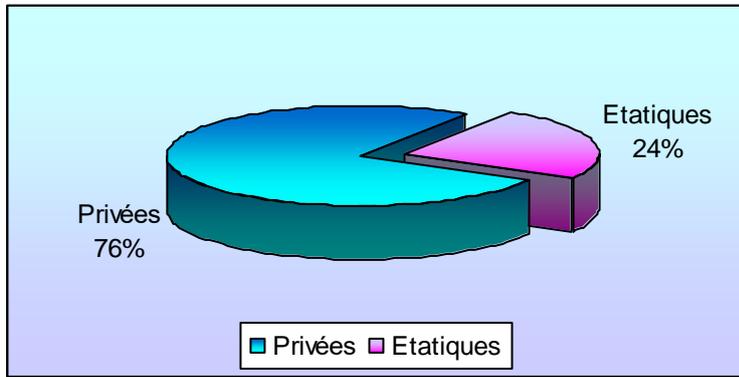


Figure: pourcentage des compagnies privées et étatiques dans la base de données.

Parmi les 181 compagnies pétrolières nationales (étatiques) et internationales (privées); 74 ont été préalablement classées, c'est-à-dire qu'on connaît les classes d'appartenance de ces compagnies. De ce fait 60 compagnies des 74 préalablement classées vont constituer la base d'apprentissage qui sera utilisée pour la construction du modèle; les 14 compagnies restantes seront utilisées pour la validation de ce modèle lequel constituera la base du test. Les 107 compagnies non classées, le seront par le modèle validé.

Les variables utilisées sont au nombre de 11, ce sont les ratios financiers de Rentabilité; d'Analyse de cash flow; d'Efficacité et de Rendement; et de Structure du capital pour la mesure de la performance, choisies d'après " The Energy Intelligence Group Top 100: Ranking the world's oil companies", de l'année 2000 :

- **Return On Average Capital Employed (ROACE):** Earnings Before Interest and Tax as a percentage of average capital employed.
- **Return On Average Total Assets (ROATA):** Earnings before interest and tax and before exceptional items over average total assets.
- **Return On Average Equity (ROAE):** Income available to common and preference stockholders over average shareholders' equity.
- **Net Return On Sales (NROS):** Net income divided by Net sales revenue. This provides a measure of the company's efficiency at generating profits from each unit of sales.
- **Capex as % of Op. Cash Flow (%) (Cap/CF):** Capex (including expenditures on intangible fixed assets) as % of Net Cash Provided by Operating Activities.
- **Long Term Debts To Total Debt (LTD/TTD):** This ratio is based on long-term debt as reported by the companies, which may vary significantly in terms of the period defined as long-term, divided by short-term notes and loans plus long-term debt.
- **Shareholders' Equity as % of Total Assets (SE/TA):** Shareholders' Equity as percentage of Total Assets.
- **Dividend To Cash Flow (%) (DT/CF):** Dividends payable to ordinary shareholders divided by Net cash provided by operating activities.
- **Net Income Per Employee (NIPE):** Net income divided by average number of employees. Period-end number of employees is used if average is not available.
- **Cash Flow To Revenue (CF/TR):** The net cash flow from operating activities divided by total sales and operating revenue.
- **Cash Flow Per Employee (CF/PE):** Net cash provided by operating activities divided by average number of employees. Period-end number of employees is used if average is not available.

A partir de ces variables qui sont des ratios de performances, caractérisant chaque compagnie, la population va être scindée en deux classes distinctes, une contiendra les compagnies

performantes et l'autre les compagnies non performantes (la non performance d'une compagnie est définie par le rang qu'elle occupe dans le classement mondial des compagnies)

Par hypothèse, les majores (**BP Amoco, Chevron Texaco, Conoco, Exxon Mobil, Shell Group, Total Fina Elf**) sont considérées comme des compagnies performantes. Aussi nous considérons que la classe qui les contiendra, sera la classe des compagnies performantes. On utilisera cette hypothèse dans l'application de la méthode des nuées dynamiques.

La classification par la méthode des nuées dynamiques (k_means)

La classification par nuées dynamiques, consiste à choisir N objets au hasard, qui constituent la première nuée. On forme n classes initiales en regroupant chaque objet avec l'élément de la nuée le plus proche. On forme ensuite une nouvelle nuée en prenant les objets, les plus proches des centres de gravité des classes. On itère le processus jusqu'à ce que les classes obtenues soient stables et les plus homogènes possibles tout en minimisant l'inertie intra-classes et on maximisant l'inertie inter-classes. Ainsi, l'application de cette méthode nous a permis d'avoir les résultats suivants :

- **Décomposition de l'inertie pour la meilleure partition obtenue**

Inertie	
Intra-classes	35384897,672
Inter-classes	21482428,766
Totale	56867326,438

- **Les Individus centraux pour chaque classe**

Classe	ROACE	ROATA	ROAE	NROS	Cap/CF	LTDTTD	SE/TA	DTCF	NICF	CFTR	CFP
1 (Gas Natural)	13,522	10,383	15,91	10,17	127,117	82,650	33,726	10,357	69,71	27,90	189,
2 (Devon EnergyCorp.)	25,275	21,581	25,18	26,49	79,061	100,000	47,770	2,000	411,80	59,60	925,

D'après le tableau ci-dessus, la première classe est constituée par les individus proches de l'individu Gas Natural, par contre la deuxième classe est constituée par les individus proches de l'individu Devon EnergyCorp.

Un classifieur est l'algorithme qui réalise automatiquement une classification. Il est conçu pour attribuer un objet représenté par un vecteur d'entrée à une classe d'appartenance.

- Le classifieur par nuées dynamiques, a pu partitionner l'ensemble des compagnies pétrolières en deux classes, la première contient 141 compagnies (privées et étatiques), la deuxième ne contient que 40.

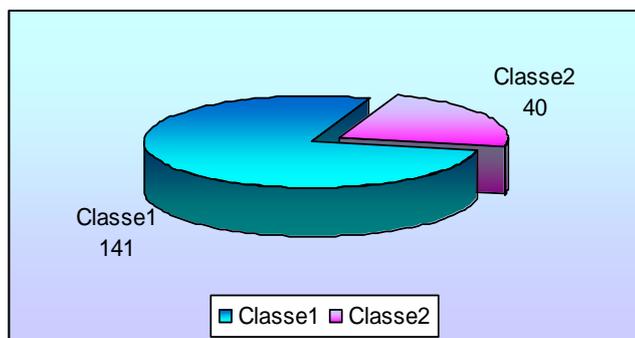


Figure: L'effectif des compagnies dans chaque classe.

- Le classifieur par nuées dynamiques a affecté 62% des compagnies privées à la première classe, parmi elles se trouvent les majeures. Et 38% de ces compagnies privées sont dans la deuxième classe. Ce qui nous permet de conclure que la première classe est celle des compagnies performantes.

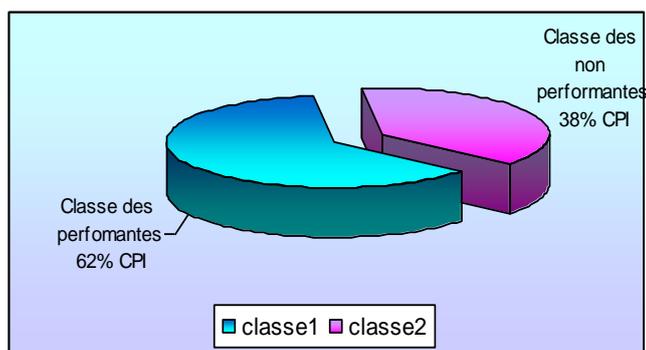


Figure: La répartition des compagnies privées entre les deux classes.

- Le classifieur par nuées dynamiques a attribué 92% des compagnies pétrolières étatiques à la première classe celle des compagnies performantes, comme PDV, ENI, PEMEX. Contre 8% seulement dans la deuxième classe celle des non performantes. La SONATRACH a été mal classée, elle a été affectée à la deuxième classe, celle des non performantes.

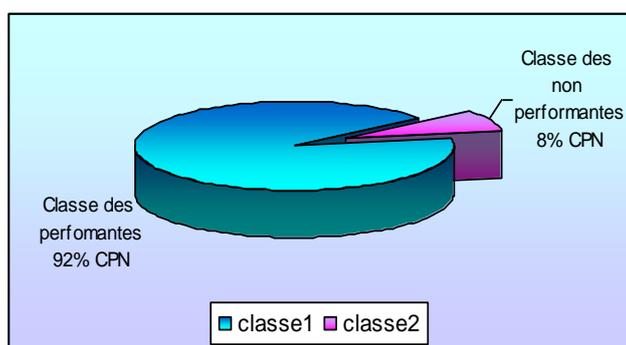


Figure: La répartition des compagnies étatiques entre les deux classes.

L la comparaison des sorties du classifieur des nuées dynamiques avec les sorties désirées, nous a permis d'avoir les résultats suivants :

Résultats de la classification	Des nuées dynamiques			
	Les individus Bien classés		Les individus Mal classés	
	effectif	%	effectif	%
Base d'apprentissage	42	70%	18	30%
Base du test	6	43%	8	57%

Pour la base d'apprentissage, le classifieur a pu classer correctement 42 compagnies de l'ensemble utilisé en apprentissage, et donc a fourni le pourcentage de 70% de bonnes réponses. Et dans la base test, il n'a pu classer correctement que 6 compagnies, et donc il a fourni 43% de bonnes réponses.

La classification par la méthode des réseaux de neurones

La classification par la méthode des réseaux de neurones, est une méthode de classification non paramétrique qui consiste à estimer une fonction de classification, à partir de l'assimilation d'un certain nombre d'exemples, afin de tirer une généralisation pour le partitionnement de l'ensemble des individus en n classes. On a choisi d'appliquer un classifieur supervisé. Etant donné que les valeurs de nos inputs sont des valeurs continues, et l'on désire que les résultats du classifieur soient obtenus en tenant compte des sorties désiré (c'est à dire de la classification préalable de certains individus).

L'élaboration d'un classifieur supervisé par réseau de neurones, nécessite le suivi de ces étapes :

- **La décomposition de la base de données en trois:**
 - Base d'apprentissage qui servira à l'estimation des paramètres du classifieur.
 - Base de test qui permettra d'évaluer la performance du classifieur.
 - Le reste de la base de données, dont la classification n'est pas connue, l'utilisation du classifieur validé permettra de classer les individus de cette base.
- **L'entrée des trois bases séparément:**
 - La base d'apprentissage.
 - La base du test.
 - La base du reste des individus.

- **La normalisation des trois bases**

La normalisation des trois bases à cause de la distinction des unités des variables, selon la formule suivante:

$$N(E_k) = \frac{E_k - \min_k}{\max_k - \min_k}$$

Où E_k : la valeur brute de la $k^{\text{ième}}$ entrée.

$N(E_k)$: la valeur normalisée de la $k^{\text{ième}}$ entrée.

\min_k : la valeur minimale enregistrée pour la $k^{\text{ième}}$ variable sur toute la base.

\max_k : la valeur maximale enregistrée pour la $k^{\text{ième}}$ variable sur toute la base.

- **L'architecture du réseau de neurones**

Notre réseau de neurones, est un réseau multicouche, il se constitue de quatre couches :

Onze inputs dans la couche d'entrée.

➤ Six neurones dans la première couche cachée.

➤ Trois neurones dans la deuxième couche cachée.

➤ un neurone dans la couche de sortie.

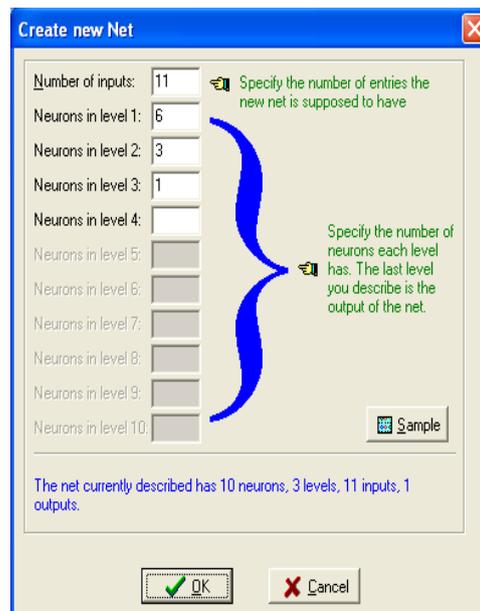


Figure: le choix d'un réseau de neurones.

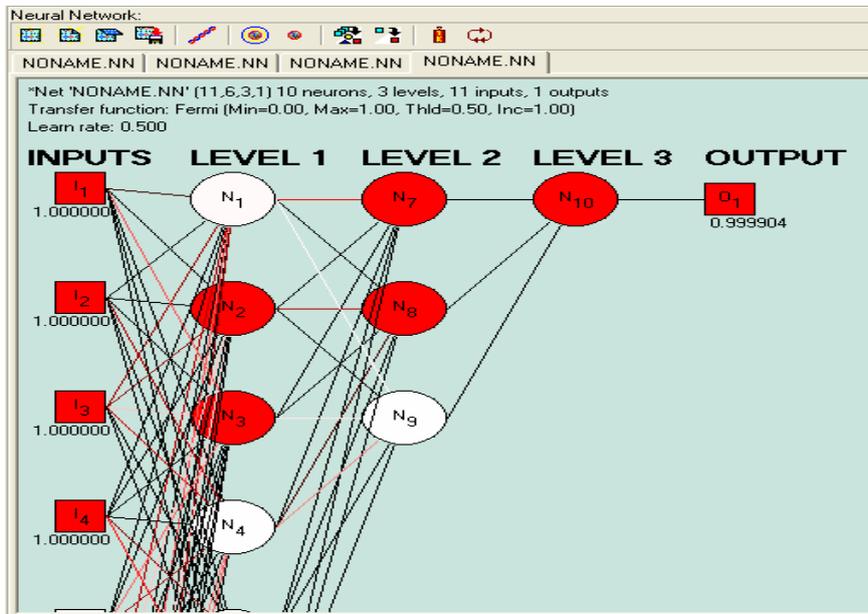


Figure: l'architecture du réseau de neurones multicouches

- **La phase d'apprentissage:**

Cette phase consiste en utilisant la base d'apprentissage, à estimer un classifieur (modèle) qui minimisera la moyenne des carrés des erreurs (average deviation).

Après l'apprentissage des données, la moyenne des carrés des erreurs est 0.030270.

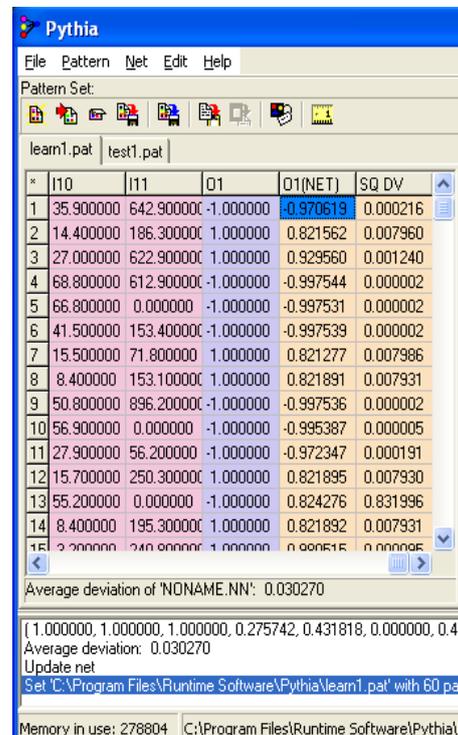


Figure: Apprentissage des données.

- **La phase de test**

Dans cette phase, on va tester la performance du classifieur élaboré. La performance du classifieur est mesurée par le nombre d'individus mal classés dans la base du test.

On remarque d'après la figure, que l'application du classifieur sur la base du test donne une moyenne quadratique des erreurs très petite, égale à 0.114931.

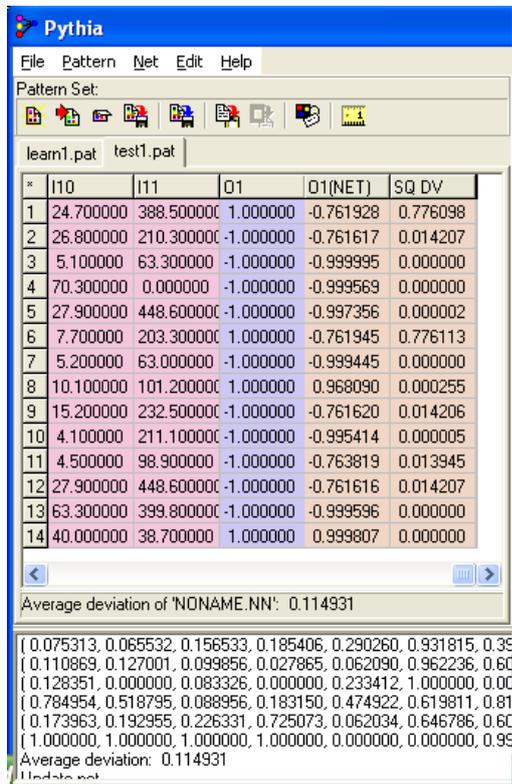


Figure: L'application du classifieur dans la base du test.

La comparaison des sorties du classifieur du réseau de neurones avec les sorties désirés, nous a permis d'avoir les résultats suivants:

Résultats de la classification	Des réseaux de neurones			
	Les individus Bien classées		Les individus Mal classées	
	effectif	%	effectif	%
Base d'apprentissage	57	95%	3	5%
Base du test	12	85%	2	15%

Ainsi, Le classifieur dans la base d'apprentissage a bien classé 57 individus et a donc fourni 95% de bonnes réponses. IL a fourni 85% de bonnes réponses dans la base de test.

Après validation du classifieur, on fait rentrer le reste des individus non classés. On obtient les résultats suivants :

➤ Le classifieur a attribué 48 compagnies pétrolières à la première classe celle des performantes, et 133 compagnies à la deuxième classe celle des non performantes.

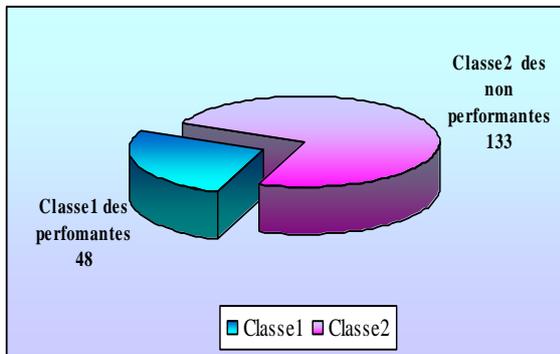


Figure: l'effectif des compagnies dans chaque classe.

➤ d'après la figure, 73 % des compagnies privées se trouvent dans la deuxième classe celle des non performantes, contre 27 % dans la première celle des performantes.

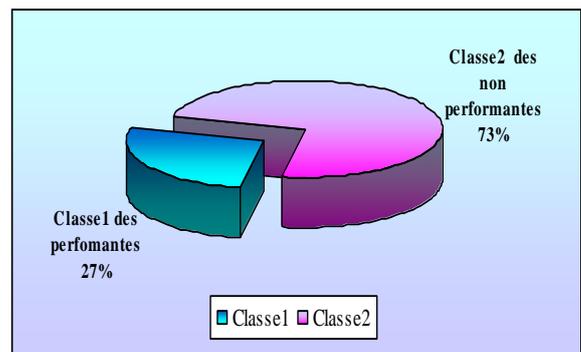


Figure: la répartition des compagnies privées entre les deux classes.

➤ le classifieur a affecté 52% des compagnies étatiques à la classe des performantes, et 48% dans la classe des non performantes.

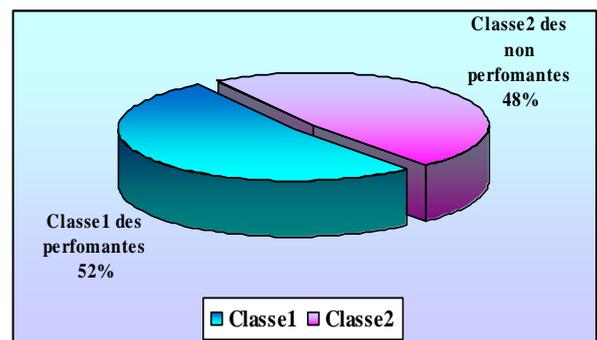


Figure: la répartition des compagnies étatiques entre les deux classes.

Les compagnies étatiques comme **PdV, Pemex, Petrochina, Petrobras, SONATRACH, ENI, ONGC, Sinopec, Gazprom, Slovnaf, Norsk Hydro, Petro-Canada, cpc** ont été affectées à la première classe celle des compagnies performantes avec les majores et d'autres compagnies privées.

Conclusion :

On peut conclure d'après les résultats obtenus par l'application des deux méthodes, que la performance n'est pas spécifique aux compagnies pétrolières internationales puisqu'il existe des compagnies nationales performantes comme PDV, Sonatrach, Pemex, et d'autres ; leur point commun étant l'objectif de réaliser des profits et d'assurer une extension pour la pérennité de l'entreprise. Toutefois, elles diffèrent du fait que, les compagnies pétrolières internationales ciblent une rentabilité de court terme, alors que les compagnies pétrolières nationales sont soumises aux contraintes de la politique énergétique nationale, ciblant une rentabilité de long terme.

Par ailleurs, l'application du classifieur du réseau de neurone a permis d'avoir des résultats meilleurs que ceux du classifieur des nuées dynamiques (95% de bonnes réponses dans la base d'apprentissage et 85% dans la base de test).

Toutefois, la préparation des données pour une représentation complète des classes n'est pas évidente, car si une région est non exploitée dans l'apprentissage le réseau ne pourra pas effectuer une généralisation sur les données appartenants à cette région.

- ◆ La définition de l'architecture optimale du réseau se fait d'une manière empirique, aucune théorie n'a pu trancher sur cette question.
- ◆ La durée de la phase d'apprentissage et parfois très longue, particulièrement si on a une grande base de données.
- ◆ Les paramètres d'apprentissage doivent être choisis soigneusement pour ne pas avoir un phénomène d'oscillation, ou une convergence longue.
- ◆ La classification par réseau de neurones appartient au système boîte noire. elle ne livre aucune information sur les règles de décision, ni sur les caractéristiques de la classe. Si les résultats ne sont pas corrects, rien n'oriente l'analyste pour mieux comprendre la nature des phénomènes qu'il traite, et de lui permettre de trouver des solutions.
- ◆ L'aptitude de généralisation est fortement liée au choix de la base de données et de l'architecture, un mauvais choix peut conduire à des résultats inacceptables en généralisation.

- L'application de la méthode des réseaux de neurones nécessite une large base d'apprentissage, qui dépend du nombre fourni des sorties désirées. Dans cette étude, le nombre de sorties désiré est très restreint.

Références :

Thèses et mémoires

Bordji Naziha, Bouchlough Fouzia « Etude de la performances des compagnies pétrolières ». Mémoire d'ingénieur d'état en statistique (INPS), encadré par Mme Moussi.

- MAAMRA. S. "Les réseaux de neurones théorie et application dans le secteur des hydrocarbures". Mémoire d'ingénieur d'Etat en recherche opérationnelle, U.S.T.H.B. 2005.
- ABALOU. M. "Stratégies des compagnies nationales en matière de gestion de porte feuilles : entre la rentabilité et la politique d'Etat», rapport de stage.
- KOUICI. S. "Réseaux de neurones artificiels et ACP pour l'amélioration de la recherche de l'information". Thèse de Magistère en recherche opérationnelle U.S.T.H.B. 2000.
- GHERBI. S. "La fiscalité des hydrocarbures vers une meilleure adaptation aux enjeux de la mondialisation". Mémoire de troisième cycle professionnel en finances publiques I.E.D.F. 2002.
- DEBAR. H. "Application des réseaux de neurones à la détection d'intrusion sur les système informatiques".Thèse de Doctorat en informatique, université PARIS 6, 1993
- HOANG. E. H. "Evaluation stratégique d'entreprise par méthodes neuronales ". Thèse de Doctorat en sciences de gestion, université PANTHEON-ASSA PARIS II ,2000.
- ELIE. F. "Conception de réalisation d'un système utilisant des réseaux de neurones pour l'identification et la caractérisation, à bord de satellites, de signaux transitoires de type sifflement ".Thèse de Doctorat en électronique, université d'ORLÉANS, 1997.
- LUCIDARME. P. "Apprentissage et adaptation pour des ensembles de robots réactifs coopérants". Thèse de Doctorat en génie informatique, automatique et traitement de signal, université MONTPELLIER II, 2003.
- TEYTAUD. O. "Apprentissage, réseaux de neurones et application ". Thèse de Doctorat en informatique, université de LYON II.
- ZAAF. I, SMAILI. C. " Système de reconnaissance de mots isolés basés sur deux approches : Réseaux de Neurones (ANN) et Programme Dynamiques (DTM) ". Projet de fin d'études d'ingénieur en informatique, USTHB, 2000.
- BOUKIATALA et ISSANI " Classes de risques en assurance Responsabilité Civil et Elaboration d'un système de Tarification par Réseaux de Neurones ". Projet de fin d'études d'ingénieur en informatique, USTHB, 2000.
- FALLAH. K, GUIATNI. M. "Application des réseaux de neurones dans le contrôle d'accès par identification de visage". Projet de fin d'études d'ingénieur d'Etat en Génie Electrique, EMP, 2004.
- GALINARI. P. " Apprentissage Numérique et Réseaux de Neurones". DEA, IARFA, université, paris 6, 2001.

Ouvrages

- GREMILLET. A. "Les ratios financiers et le nouveau plan comptable". Les éditions d'organisation, paris 1984.
- DREYFUS. G. " Réseaux de neurones méthodologie et application". 2^{ème} édition EYROLLES, Paris.
- PERSONNAZ. L, RIVALS. I. " Réseaux de neurones formels pour la modélisation la commande et la classification", CNR édition .Paris 2003.
- TRIVIA. S, GASCUEL. O, LECHEVALLIER. Y. "Statistique et méthode neuronales". Dunod Paris, 1997.
- SAPORTA.G. "Probabilités, analyses des données, et statistiques". Edition TECHNIP.Paris, 1990.
- VOLLE, M., " Analyse des données." Economica.2^e édition, Paris, 1981.
- John S. Denker. " Les réseaux de neurones". Inter édition 1990.

- CAUSER. J.P, HARVATOPOULOT. Y, SARNIN, P. "Guide pratique d'analyse des données", Les éditions d'organisation, 1994.
- HRONEC .S.M. " Vital Sign : des indicateurs, coûts, qualité, délai pour optimiser la performance de l'entreprise", les éditions d'organisation 1995.
- Davalo Eric, Naïm Patrick. " Des réseaux de neurones". Edition EYROLLES. 1989.

Reuves

- L'avenir des sociétés nationales des pays exportateurs d'hydrocarbures N° 9/1994.
- la revue de SONATRACH N° 48-Mars 2006.
- Rapport annuel de la SONATRACH 2001.
- THE Energy Intelligence Group Top 100: Ranking the world's oil companies. Edition 2002.

Cédés et logiciels

- Collection Microsoft Encarta 2005.
- Collection Microsoft Encarta 2006.
- Le monde diplomatique.
- Logiciel pythia.
- Logiciel XLstat.