

## **Prédiction du risque de crédit : étude comparative des techniques de Scoring**

### **Prediction of credit risk: comparative study of scoring techniques**

**Siham Lotfi, (Doctorante)**

*Laboratoire Business Intelligence, Gouvernance des Organisations et Finance  
Faculté des Sciences Juridiques, Economiques et Sociales de Casablanca  
Université Hassan II, Casablanca, Maroc*

**Hicham Mesk, (Enseignant-Chercheur, PES)**

*Laboratoire Business Intelligence, Gouvernance des Organisations et Finance  
Faculté des Sciences Juridiques, Economiques et Sociales de Casablanca  
Université Hassan II, Casablanca, Maroc*

<b>Adresse de correspondance :</b>	Faculté des Sciences Juridiques, Economiques et Sociales Km 9, Route El jadida, BP 8110 Oasis, Casablanca Université Hassan II Maroc (Casablanca) <a href="http://www.fdc.ma">www.fdc.ma</a>
<b>Déclaration de divulgation :</b>	Les auteurs n'ont pas connaissance de quelconque financement qui pourrait affecter l'objectivité de cette étude.
<b>Conflit d'intérêts :</b>	Les auteurs ne signalent aucun conflit d'intérêts.
<b>Citer cet article</b>	Lotfi, S., & Mesk, H. (2020). Prédiction du risque de crédit : étude comparative des techniques de Scoring. International Journal of Accounting, Finance, Auditing, Management and Economics, 1(2), 511-527. <a href="https://doi.org/10.5281/10.5281/zenodo.4029645">https://doi.org/10.5281/10.5281/zenodo.4029645</a>

**DOI: 10.5281/zenodo.4029645**  
**Published online: 15 September 2020**

Copyright © 2020 – IJAFAME



## **Prédiction du risque de crédit : étude comparative des techniques de Scoring**

### **Résumé**

Une bonne maîtrise et gestion du risque de crédit est devenue la principale préoccupation des établissements financiers qui ne cessent de développer des modèles d'analyse, d'évaluation et de prédiction de ce risque, notamment avec les normes prudentielles exigées par les banques centrales.

Les méthodes d'évaluation et de prédiction du risque de crédit sont représentées sous la forme de modèles de scoring qui ont pour but la prédiction de la défaillance d'une entreprise grâce à des informations financières et comptables.

L'objectif de notre travail est d'étudier les différentes techniques de crédit scoring, leurs intérêts en tant qu'un outil puissant permettant de prévoir la solvabilité des emprunteurs.

**Mots clés : Prédiction, Risque crédit, Techniques de prédiction, Scoring.**

**Classification JEL : G, G1, G17**

**Type de l'article : Article théorique.**

### **Abstract**

Good control and management of credit risk has become the main concern of financial institutions, which are constantly developing models for analyzing, assessing and predicting this risk, particularly with the prudential standards required by central banks.

Credit risk assessment and prediction methods are represented in the form of scoring models which aim to predict the potential vulnerability of a business using financial information and computable.

The objective of our work is to study the different techniques of credit scoring, their interest as a powerful tool allowing to predict the solvency of the borrowers.

**Keywords: Prediction, Credit risk, Prediction techniques, Scoring.**

**JEL Classification: G, G1, G17**

**Paper type: Theoretical study.**

## 1. Introduction

L'octroi du crédit constitue la principale source de revenus des établissements bancaires, en même temps, il entraîne des risques importants s'il n'est pas bien géré, d'où l'importance de la maîtrise des risques bancaires qui représente le thème central des accords de Bâle.

Le risque de défaillance a des conséquences majeures à la fois sur le pilotage stratégique et financier des banques et sur la perception par le marché de la structure de leur portefeuille d'activités et de la rentabilité.

En effet, la maîtrise de la gestion du risque de crédit ne peut être pratiquée d'une manière efficace sans la connaissance du marché du crédit ainsi que les différents déterminants relatifs au secteur bancaire qui peuvent apprécier ce risque.

Ainsi, la mesure du risque de crédit est primordiale tant il est difficile d'anticiper l'évolution de la situation de l'emprunteur. Cette maîtrise permettra au banquier de prendre une décision d'acceptation ou de refus de contracter le prêt en fonction des données relatives à l'emprunteur.

Généralement, le risque de défaut se calcule selon les trois paramètres de détermination de défaut : l'évaluation de la perte en cas de défaut, le taux de recouvrement en cas de défaut et la probabilité de défaut.

De ce fait, les établissements bancaires sont donc très intéressés à développer des modèles d'analyse, de mesure et d'évaluation du risque de crédit plus précis afin d'apprécier le rendement des prêts accordés.

Parmi ces méthodes de prédiction, on peut citer le crédit Scoring, cette méthode implique l'exploitation des différentes techniques statistiques pour aboutir à une fonction score basée sur les caractéristiques de l'emprunteur et qui aide à la prise de décision dans l'octroi de crédit.

Dans la pratique, on trouve une variété de techniques allant des méthodes conventionnelles telles la méthode de régression linéaire multiple, analyse probit, régression logistique, analyse discriminante ou bien des techniques avancées comme l'algorithmique, le système d'expert (Hand & Henly, 1997).

Dans la littérature financière, d'autres méthodes ont été récemment développées comme les modèles basés sur l'intelligence artificielle dont on peut trouver les réseaux de neurones artificiels, cette technique est en plein essor, utilisée dans de nombreuses disciplines comme (la médecine, le marketing, la finance...), elle constitue, une alternative intéressante aux techniques statistiques traditionnelles pour le traitement des données.

Dans ce présent article, nous allons présenter le concept du crédit Scoring, ensuite nous exposons les différentes méthodes d'évaluation du risque de crédit, ces techniques demeurent nécessaires afin de prévenir et d'anticiper ce type de risque.

La décision d'octroi du prêt constitue une opération risquée lorsque l'emprunteur ne rembourse pas ses dettes à l'échéance fixée, chose qui affecte la situation financière de la banque. Dans ce cadre, la maîtrise du risque de crédit devient aujourd'hui l'une des principales préoccupations des établissements bancaires qui ne cessent de développer les moyens et techniques statistiques de prévision du risque de crédit, car une gestion bien étudiée permet de faciliter et d'améliorer l'efficacité de la prise de décision.

Notre étude consiste à analyser les différentes techniques du scoring qui peuvent être classer en méthodes statistiques basées sur les données comptables comme l'analyse discriminante (Altman, 1968), la régression logistique (Lachenbruch et alii, 1973) et techniques d'intelligence artificielle comme les systèmes d'expert, les réseaux de neurones (Altman et alii, 1994 ; Bardos et Zhu, 1997) ou les algorithmes génétiques (Varetto, 1998)<sup>1</sup>.

À travers cet article, nous avons jugé utile dans un premier lieu de cerner les différents concepts du crédit scoring avec une analyse de son processus de construction, en second lieu,

---

<sup>1</sup> [Catherine Refait-Alexandre](#), 2004 « La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux Dans *Économie & prévision* 2004/1 (n° 162), pages 129 à 147

nous nous intéresserons à la présentation des différentes méthodes de prédiction du risque de crédit. Dans ce cadre, nous nous appliquerons à exposer la théorie d'intermédiation financière et la théorie d'asymétrie d'information essentielles à la compréhension de l'économie bancaire et particulièrement à la relation entre le prêteur (la banque) et l'emprunteur. Enfin, on établira une analyse comparative des importantes techniques du crédit scoring.

## 2. Revue de Littérature

### 2.1. Concepts du crédit Scoring

Le crédit scoring est généralement considéré comme une méthode d'évaluation utilisée par les organismes bancaires pour estimer le risque de défaut et mesurer la solvabilité de chaque entreprise et lui classer soit comme une entreprise saine ou une entreprise défailtante. Il impose d'utiliser les différentes techniques statistiques en vue d'obtenir un modèle de scoring basé sur les caractéristiques de l'emprunteur.

Pour (Flaman, 1997), le crédit scoring est le processus d'assignation d'une note (ou score) à un emprunteur potentiel pour estimer la performance future de son prêt.

(Thomas et al. 2002) stipulent que le crédit Scoring représente un ensemble de modèles de décision et des techniques sous-jacentes qui permettent de décider l'octroi des crédits de consommation.

(Bardos, 2001, 2008a, 2008b) souligne que le scoring est une analyse statistique permettant de prédire la qualité d'un emprunteur. Cette méthode s'inscrit dans le Data Mining, il est défini par R. Anderson (2007) comme l'ensemble des modèles statistiques capables de transformer des informations (qualitatives, quantitatives) en indicateurs numériques mesurables afin de prendre la décision d'octroi ou de rejet du prêt.

La technique de scoring permet d'utiliser des modèles statistiques afin de transformer des données (qualitatives, quantitatives) en indicateurs numériques mesurables à des fins d'aide à la décision d'octroi ou de rejet de crédit.

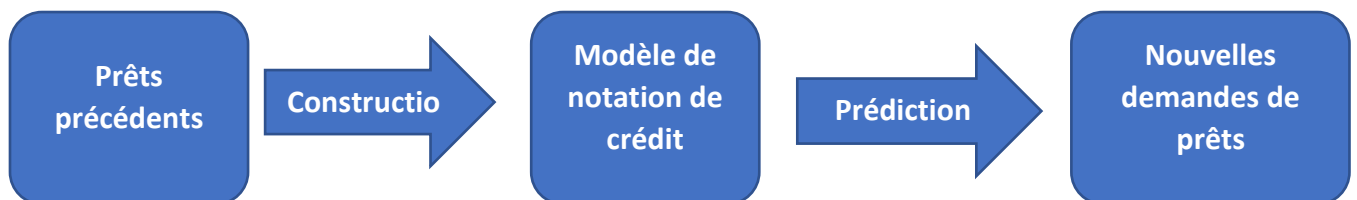
Dans le même d'ordre d'idée, Gilbert Saporta définit le Crédit Scoring comme un ensemble d'outils d'aide à la décision permettant aux organismes financiers de mesurer le risque de crédits octroyés.

Les méthodes de scoring ont été utilisées dans plusieurs domaines autres que le marché du crédit, en assurance automobile, en prospection publicitaire, en épidémiologie, pour distinguer entre les malades et les sains, etc.... Mais généralement, cette méthode est utilisée par les banques afin de prévoir le risque de défaut d'un emprunteur (un particulier ou une firme).

À travers l'historique des données des prêts précédents, le scoring permet d'attribuer une note appelée « score », qui permet de prédire la probabilité de défaut pour les nouveaux crédits.

Le processus du crédit scoring peut être résumé comme suit :

**Figure 1 : Processus du Crédit Scoring**



Source: Liu<sup>2</sup> (2001)

<sup>2</sup> Liu, Y. (2001). New issues in credit scoring application. Work report No, 16.

### 3. Processus de construction d'un modèle de Scoring

La construction d'une fonction de score repose sur la disposition de deux groupes, un groupe constitué des emprunteurs défaillants et un autre groupe composé des emprunteurs sains. L'élaboration de ce modèle nécessite aussi un certain nombre de critères :

- La détermination du critère de défaillance.
- Le choix des variables explicatives de ce défaut.
- Le choix de la technique qui permet d'évaluer de ce défaut.

#### 3.1. La détermination du critère de défaillance

Il est nécessaire de disposer des données historiques sur les défauts qui doivent couvrir un cycle économique complet de même, d'élaborer un échantillon représentatif d'emprunteurs en situation de défaut. Ces emprunteurs doivent faire partie à des populations homogènes afin de chercher et détecter les caractéristiques économiques financières les plus prédictifs le plus les emprunteurs sains de ceux en défaut.

L'élaboration du modèle de scoring nécessite aussi la détermination d'un horizon de prédiction. Prenant l'exemple des données de l'année N-3 pour prévoir la défaillance de l'année en cours, dans ce cas, l'horizon de prédiction sera de trois ans.

#### 3.2. Le choix des variables explicatives

La construction d'une base de données est fondée sur la collecte des données utiles et fiables qui affectent la capacité du remboursement de l'emprunteur. Ces données représentent des variables collectées de l'historique et qui doivent présenter des dimensions différentes du risque de défaut et informer sur les caractéristiques de l'emprunteur, du crédit et du prêteur (la banque).

Une variété de variables peut s'introduire dans le modèle, on trouve :

- **Les données qualitatives de l'emprunteur** : elles concernant les caractéristiques démographiques, l'activité professionnelle, l'ancienneté dans l'activité, l'expérience, sa localisation géographique, le ménage auquel il appartient et les avoirs de la famille, existence d'incident dans le passé. Ces données concernent les scores des clients particuliers.
- **Les données comptables et financières** : elles concernent généralement les ratios financiers (solvabilité, activité, taille, endettement, liquidité,) pour les entreprises.
- **Les caractéristiques du crédit** : elles concernent le montant du crédit demandé, le taux d'intérêt, le délai de remboursement, montant des remboursements, du différé et, le cas échéant, des types de garanties ....
- **Les caractéristiques du prêteur** : la banque doit disposer des informations concernant l'agent chargé de crédit (son niveau d'étude, ses caractéristiques socio démographiques, sa situation familiale, son niveau d'étude).

Ces caractéristiques doivent figurer dans la base de données de la banque afin que la méthode du scoring puisse être pratiquée pour la construction de la fiche d'évaluation.

#### 3.3. Le choix de la technique à utiliser

Avec le développement des différents besoins des systèmes de crédit scoring, on assiste à une variété des méthodes d'évaluation du risque de crédit avec un objectif identique qui est l'augmentation de l'efficacité des prises de décision.

Le principe de ces méthodes est d'identifier les variables qui déterminent la probabilité de défaut afin de pondérer leurs poids dans un score quantitatif. Ces systèmes de crédit scoring sont mis en place à partir de quatre principales formes de modélisation multivariée :

- Le scoring par le modèle linéaire.
- Le scoring par le modèle d'analyse discriminante.
- Le scoring par le modèle logit.

En plus de ces techniques, on trouve d'autres méthodes d'intelligence artificielle citées par S. Tuffery (2007)<sup>3</sup> :

- Les réseaux de neurones.
- Les arbres de décision
- Les systèmes experts.
- Algorithmes Génétiques

Dans ce cadre, on peut poser les questions suivantes :

1. *Quelles sont ces techniques de prédiction et d'évaluation du risque de crédit ?*
2. *Qu'est ce qui caractérise les techniques intelligentes par rapport aux autres méthodes traditionnelles de crédit scoring sachant que les deux catégories disposent d'un objectif commun qui est le classement des entreprises en saines et défaillantes ?*

#### **4. Les méthodes de prédiction du risque de crédit**

La revue de littérature nous offre des diverses méthodes de prédiction du risque de crédit classées sous forme de méthodes basées sur le jugement humain, les méthodes statistiques et les méthodes basées sur l'intelligence artificielle.

##### **4.1. L'Analyse financière**

L'analyse Financière est une technique de l'approche traditionnelle. Il s'agit probablement de la méthode à la fois la plus ancienne et la plus utilisée en analyse du risque qui a pour objectif d'étudier le passé afin de prévoir le présent et l'avenir, (Vernimmen, 1998).

L'établissement de crédit va établir différents ratios (de structure, d'endettement, de rentabilité et de liquidité) et calculs pour vérifier la performance de l'entreprise à travers son compte de résultat et son bilan. Selon (Ndaynou, 2001). Cette analyse se concentre sur deux éléments :

- **Le flux de liquidité futur** : est obtenu par le calcul de la différence entre les entrées et les sorties de flux réalisés par l'entreprise. Ce flux de liquidité permet d'avoir une visibilité sur la capacité du débiteur à rembourser ses dettes sans mettre en péril son activité durant un emprunt. La banque peut continuer à consulter l'évolution des bénéficiaires et voir si elles sont suffisantes par rapport au besoin en fonds de roulement.
- **Le fonds de roulement** permet d'apprécier l'équilibre financier de l'organisation. Il indique si l'entreprise est pérenne et pourra assurer ses engagements. Pour le calculer il existe deux méthodes : Soit par le haut du bilan avec la différence entre les ressources stables (capitaux propres et dettes à long terme) moins les emplois stables (actif immobilisé net). Soit par le bas du bilan avec la différence entre l'actif circulant d'exploitation et les dettes à court terme. »

##### **4.2. L'analyse discriminante multivariée**

L'analyse discriminante est un outil statistique qui permet de distinguer à partir d'un ensemble d'entreprises séparées en 2 groupes les entreprises saines et les entreprises défaillantes en tenant compte leurs caractéristiques spécifiques.

L'analyse discriminante permet de décrire, d'expliquer et de prédire l'appartenance d'un ensemble d'observations sur la base d'une série de variables prédictives à des groupes prédéfinis.

Du point de vue pratique, l'analyse discriminante permet une combinaison linéaire optimale entre des variables prédictives pour différencier les entreprises saines des entreprises en détresse. L'analyse discriminante exige que la variable à expliquer soit de nature

---

<sup>3</sup> Stéphane Tufféry, « Data Mining et Statistique décisionnelle, l'intelligence des données », Editions Technip (2e Ed), 2007, p.25.



qualitative et les variables soient indépendantes et distribuées selon une loi normale. Sa formule se présente comme suit :

$$Z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Avec (Z) représente le score de l'entreprise.

L'analyse discriminante de Fisher est une technique descriptive permettant une représentation graphique de combinaisons de variables initiales. Elle permet de produire une suite de variables discriminantes, les entreprises des mêmes groupes soit les plus proches soit celles d'autres groupes différents. La dispersion à l'intérieur du groupe est présentée par la matrice de variance covariance  $W_k$  :

$$W = \frac{1}{n} \times \sum_k n_k \times W_k$$

#### 4.3. La régression logistique

Desjardins (2005) définit la régression logistique comme une technique permettant d'ajuster une surface de régression à des données lorsque la variable dépendante est dichotomique. Cette technique est utilisée pour des études ayant pour but de vérifier si des variables indépendantes peuvent prédire une variable dépendante dichotomique.<sup>4</sup>

Cette méthode est très utilisée en médecine (caractérisation des sujets malades par rapport aux sujets sains par exemple), en marketing et commence à prendre de l'ampleur dans les sciences économiques et de gestion vu son vif succès.

La régression logistique est une technique statistique qui s'utilise lorsque la variable dépendante est qualitative, le plus souvent binaire ou dichotomique. Quant aux variables explicatives, ils peuvent être soit qualitatives soit quantitatives.

L'intérêt majeur de cette technique est de mesurer la relation de l'association entre chaque variable indépendante et la variable dépendante, en tenant compte de l'effet des autres variables intégrées dans le modèle.

Le modèle fournit la probabilité qu'un événement se produise ou non (dans notre cas défaut ou non défaut) et les variables indépendantes X sont celles susceptibles d'influencer la survenue ou non de l'événement. La fonction logistique s'écrit de la manière suivante :

$$Y = P(Y = 1) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \dots + \beta_k X_k$$

Avec :

$$Y = \text{Log} \left[ \frac{p}{1-p} \right]$$

$$p = \frac{e^{\beta_0 + \sum_{i=1}^k \beta_i x_i}}{1 + e^{\beta_0 + \sum_{i=1}^k \beta_i x_i}}$$

Le fondateur de la régression logistique est Ohlson (1980), il est le premier à publier une étude pour prédire la défaillance en utilisant un modèle logit. Ohlson a utilisé neuf variables comptables afin de prédire la probabilité de défaut de l'emprunteur en supposant que cette probabilité est distribuée logistiquement.

Ainsi, la fréquence cumulative des défauts prend la forme fonctionnelle logistique et peut ainsi, prendre des valeurs contenues entre zéro et un.

Les ratios utilisés par Ohlson sont comme suit :

---

<sup>4</sup> Julie DESJARDINS, « L'analyse de la régression logistique », Université de Montréal, Tutorial in Quantitative Methods for Psychology, 2005, Vol. 1(1), p. 35-41.

- **Log (total des actifs / niveau d'indice du prix)**
- **Total des dettes / total des actifs**
- **Fonds de roulement / total des actifs**
- **Dettes exigibles / actif circulant**
- **Résultat net / total des actifs**
- **Fonds générés par l'exploitation / total des dettes**
- **Le résultat net**
- **Une variable binaire qui prend 1, si total des dettes est supérieur au total des actifs, sinon elle est égale à 0**
- **Une variable qui prend '1' si le résultat net des deux dernières années est négatif, sinon elle est égale à '0'.**

Face aux contraintes liées à l'application des méthodes paramétriques, d'autres techniques empruntées à l'intelligence artificielle sont apparues. Selon certaines études comparatives, plus précisément celles qui appliquent les approches d'intelligence artificielle à la prévision de la défaillance de l'entreprise, confirment leur supériorité par rapport à l'analyse discriminante en matière de la qualité de prévision. A côté des outils paramétriques, l'utilisation des réseaux neurones par Altman et al (1994), Bardos et Zhu (1997), les systèmes d'experts par les établissements bancaires et les grandes entreprises s'est intensifiée.

## **5. Les techniques intelligentes**

Face à la montée en puissance des progrès de l'information (2000), les établissements bancaires et les grandes entreprises ont fait appel à d'autres modèles telles les méthodes de l'intelligence artificielle, celle des réseaux neurone qui sont utilisées pour la prédiction du phénomène de défaillance. Ces modèles relèvent du domaine de l'intelligence artificielle, plus particulièrement la branche liée à l'apprentissage automatique.

### **5.1. Les Systèmes experts**

« Un système-expert est un outil informatique d'intelligence artificielle, conçu pour simuler le savoir-faire d'un spécialiste, dans un domaine précis et bien délimité, grâce à l'exploitation d'un certain nombre de connaissances fournies explicitement par des experts du domaine. » J.C. Pomerol.

Davis (1986), précise que les systèmes experts sont utiles pour résoudre des problèmes complexes comme la stratégie d'entreprise, l'évaluation, la prise de décision et la résolution de problèmes.

Cette méthode d'intelligence artificielle est considérée comme :

- Un système de décision : dans ce cas le système propose des choix à suivre
- Un système d'aide à la décision : dans la mesure où le système est capable de proposer des choix tout en gardant une marge du pouvoir décisionnel d'interprétation.
- Un système d'aide à l'apprentissage : le système joue le rôle d'un formateur afin de développer les connaissances.

Dans le cadre de décision d'octroi du crédit, les systèmes experts sont basés sur les règles relatives aux caractéristiques du client identifiées par les responsables du crédit. Ces règles permettent d'évaluer le risque de défaut de chaque emprunteur et d'attribuer un score selon leurs caractéristiques.

### **5.2. Les réseaux de neurones artificiels**

Parmi les méthodes non paramétriques, on trouve les réseaux de neurones qui occupent une place prépondérante dans le domaine de classification et de prédiction. Cette approche neuronale a récemment attiré l'attention des chercheurs. Elle s'est développée dès les années 80, pour faire face aux limites de l'analyse discriminante dans la prédiction du risque de crédit



des entreprises (Anandarajan, Lee et Anandarajan, 2001). Ils ont été utilisés pour mesurer le risque de défaillance par les ratios financiers, (Leshno et Spector, 1996 ; Bardos et Zhu, 1997).

Ces modèles neuronaux sont utilisés dans différents domaines comme le marketing, le data mining, la médecine..., (Liu et Yu, 2005), mais ils sont moins exploités dans le champ de la prédiction du risque de défaillance.

Le fonctionnement d'un réseau de neurones passe par trois phases :

- **La phase d'input**
- **La phase de la fonction de transfert**
- **La phase d'apprentissage.**

### 5.2.1. La phase d'Input

Les réseaux reçoivent les informations sur une couche réceptrice contenant des neurones. Ces derniers traitent ces informations avec ou sans l'aide d'une ou plusieurs couches cachées contenant une ou plusieurs neurones produisant un signal ou plusieurs de sorties.

Chaque neurone appartenant soit à la première couche (réceptrice), soit aux couches cachées ou à la couche de sortie est liée aux autres neurones par des connexions (similaires aux synapses du cerveau) auxquelles sont affectés des poids (eux-mêmes assimilables aux potentiels synaptiques)<sup>5</sup>.

Les réseaux de neurones, dans le cadre du credit scoring, permettent de mettre en relation les inputs (la base de données qui est composée des dossiers de crédits) et les outputs (le résultat du crédit : bons payeurs ou mauvais payeurs) sans supposer que cette relation est linéaire.

Le processus de traitement se présente comme suit :

- **Une couche d'input :**

La couche d'input est composée des neurones qui reçoivent des signaux de l'environnement. Cette couche sert à distribuer les valeurs d'entrées aux neurones des couches supérieures, éventuellement multipliées ou modifiées d'une façon ou d'une autre. Le nombre de neurones d'entrée dépend du nombre de variables dans le vecteur d'input.

- **Des couches cachées :**

Les neurones dans ces couches n'ont aucune interaction directe avec l'environnement. Ils se situent entre la couche d'entrée et la couche de sortie. La détermination du nombre des neurones dans les couches cachées est la tâche la plus difficile dans les réseaux de neurones en général. Il est préférable d'utiliser un nombre réduit de neurones cachés, mais il est insuffisant aussi d'utiliser que deux neurones pour modéliser un ensemble de données<sup>6</sup>.

- **Une couche d'output :**

Les neurones de cette couche émettent des signaux à l'environnement. Elle calcule une somme pondérée de toutes ses entrées. Le nombre des neurones de sortie est rattaché au problème étudié.

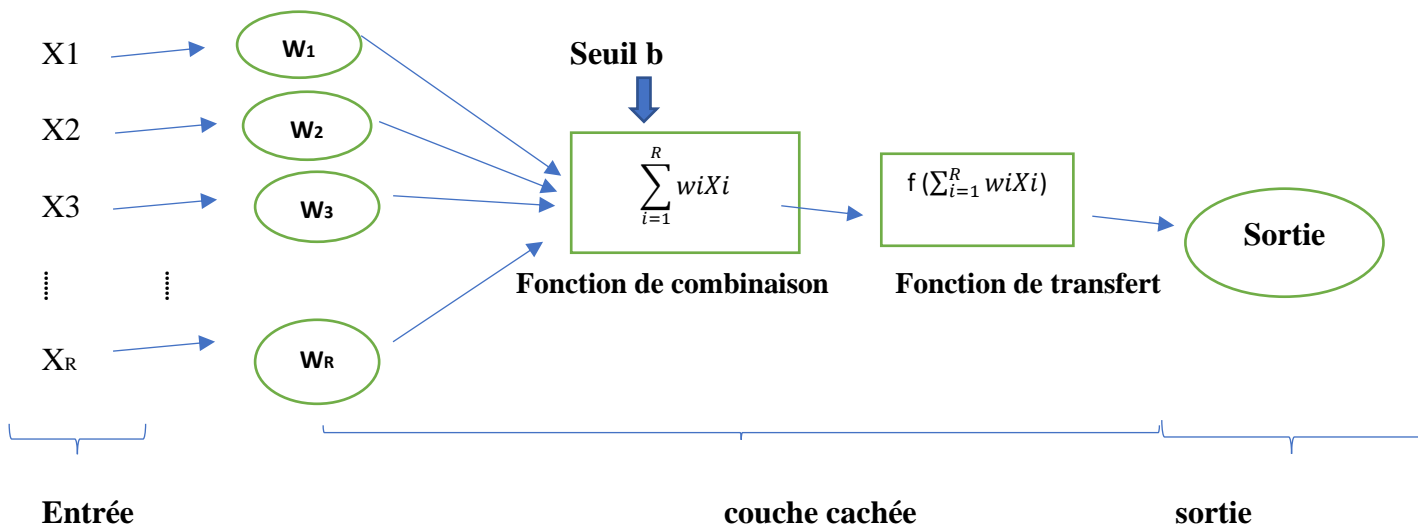
La figure ci-dessous montre bien le fonctionnement et l'architecture d'un réseau de neurones. En effet, Chaque entrée est liée à un poids  $w_i$  appelé aussi **poids synaptique**, ce poids constitue une force de connexion.

---

<sup>5</sup> Thuillier. D (1997), « Principe et applications des réseaux de neurones –deux illustration sur l'habitat au Maroc » ; revue région et développement n°=5-, page3

<sup>6</sup> Jantzen. J (1998), "Introduction to perceptron Networks ", report no 98-H 873 (nnet), 25 Oct, Denmark, page27.

**Figure 7 : Conception d'un réseau de neurones**



Source : Adapté de Tufféry, 2012

Le neurone ne traite pas chaque information reçue unilatéralement, mais effectue une somme pondérée de toutes les entrées. Cette somme est représentée sous la fonction de combinaison suivante :

$$a = \sum_{i=1}^R w_i x_i - b$$

**Avec :**

$b$  : le biais de neurone ou seuil d'activation du neurone.

$x_i$  : l'information qui parvient aux neurones de rang  $i$  de la couche d'entrée

$R$  : le nombre d'informations

$w_i$  : la pondération du signal émis par le neurone de la couche d'entrée vers le neurone de la couche cachée.

$a$  : le niveau d'activation du neurone, qui est le signal total reçu par le neurone de la couche cachée.

### 5.2.2. La phase de la fonction de transfert

Chaque neurone reçoit les informations transformées par les neurones de la couche précédente avec lesquels il se trouve en relation pour calculer son potentiel d'activation. La fonction d'activation, ou fonction de transfert, est la fonction qui détermine l'output d'un neurone à partir de la somme des poids pondérée des entrées du réseau. La fonction d'activation des couches cachées est souvent non linéaire, ce qui confère au RNA la propriété de non-linéarité.

Il existe plusieurs types de fonctions de transfert, on trouve : la fonction de Heaviside, La fonction linéaire, La fonction sigmoïde exponentielle, mais la fonction la plus utilisée est la fonction sigmoïde, elle est définie comme une fonction réelle de  $\mathbb{R}$  vers  $\mathbb{R}$ , continue, bornée, sachant que sa dérivée est toujours positive. Elle peut prendre deux formes : la fonction logistique exponentielle, ou la fonction hyperbolique. (Masters (1993)).

### 5.2.3. La phase d'apprentissage

Parmi les caractéristiques les plus importantes des réseaux de neurones est l'apprentissage, cette dernière repose sur le développement du réseau de neurones jusqu'à l'obtention du résultat attendu. Autrement dit, le fonctionnement du réseau de neurones repose sur la fourniture des données en entrée que le neurone doit apprendre à reconnaître dans le but de les classer en sous-groupes homogènes, c'est ce qu'on appelle : apprentissage non supervisé à des fins descriptives,

ou pour les associer à une variable de sortie : apprentissage supervisé à des fins prédictives<sup>7</sup>, ce qui convient aux objectifs de notre étude.

On distingue ainsi deux types d'apprentissages : Apprentissage supervisé (Supervised Learning) et Apprentissage non supervisé (Unsupervised Learning).

Dans l'**apprentissage supervisé**, le réseau s'adapte par comparaison entre le résultat qu'il a calculé, en fonction des entrées fournies, et la réponse attendue en sortie. Le système ajuste les données jusqu'à ce qu'il obtienne la bonne sortie.

En ce qui concerne l'**apprentissage non supervisé**, il est capable de découvrir la forme à partir des données fournies vu que le réseau dans ce cas, doit chercher les ressemblances entre les informations de la base de données, dans ce type d'apprentissage, les données d'entrées existent, mais il n'y a aucune information sur la sortie désirée.

### 5.3. Algorithmes Génétiques

Il s'agit d'une méthode qui appartient à la famille des méthodes d'intelligence artificielle, c'est un sujet qui intéresse bon nombre d'auteurs. En effet, F. Varetto (1998) et D.K. Barney et al. (1999) étaient les premiers à développer les algorithmes génétiques dans le cadre de la prédiction de la faillite. Ce modèle de prédiction façonné est donc très similaire à la méthode des réseaux de neurones.

Cette méthode permet de déduire une trajectoire du cours de la faillite en se basant sur toute une série d'informations et de ratios comptables sur les entreprises ayant connu la faillite ou non et grâce à une succession de récurrence.

Dès lors, cette trajectoire peut être reprise pour d'autres entreprises. Le but est de parvenir, suite aux différentes récurrences, à des populations de solutions de plus en plus homogènes jusqu'à atteindre le seuil d'homogénéité prédéfini (C. Refait, 2004). Il va observer leur qualité à travers les taux de bons classements et enfin, il va maintenir les modes de résolution les plus efficaces et potentiellement les combiner afin d'en obtenir de nouveau de façon aléatoire.

## 6. Les théories mobilisées dans le champ du risque de crédit

### 6.1. La théorie d'intermédiation financière

Dans le cadre de la théorie d'intermédiation financière, les banques jouent un rôle primordial dans la création d'une connexion entre les agents économiques disposant d'une capacité de financement et ceux ayant un besoin de financement<sup>8</sup>. Cette théorie permet de se renseigner sur les techniques utilisées par la banque et sur le fonctionnement du marché de crédit bancaire.

Selon R. Merton (1995) trois types d'intermédiation existent dans le contexte de la relation Banque – entreprise :

- **L'intermédiation de la liquidité** : ce type d'intermédiation postule que la banque collecte des dépôts et, en contrepartie, distribue des crédits.
- **L'intermédiation de l'information** : face à l'existence d'une asymétrie d'information entre l'entreprise propriétaire de projets et les financeurs pouvant inciter l'entrepreneur à dissimuler ou à minimiser les gains réels provenant du projet, la banque doit exercer un rôle de surveillance active (monitoring) selon D. Diamond (1984). Cette action de monitoring consistant à suivre régulièrement la valeur du projet peut dissuader l'entreprise de sous-estimer les cash-flows.

<sup>7</sup> Fred Ntoutoume Obiang-Ndong, 2007 « Scoring du risque de crédit des PME par la modélisation statistique et l'intelligence artificielle chez l'UMECUDEFS : une application comparative de la régression logistique et des réseaux de neurones ». Mémoire de fin d'études, Université Cheikh Anta Diop - Faculté Des Sciences Economiques Et De Gestion.

<sup>8</sup> Rim BOUSSAADA, 2012 « L'impact de la gouvernance bancaire et de la relation bancaire sur le risque de crédit : Cas des banques Tunisiennes ».

- **L'intermédiation du risque**, selon laquelle, devant un risque, l'établissement de crédit a le choix entre deux comportements : soit l'acceptation de la gestion du risque à travers la diversification et surveillance des clients emprunteurs, soit le transférer un tiers.

Dans ce contexte, la revue de littérature présente les différents chercheurs ayant placé la problématique du risque de crédit dans la théorie de l'intermédiation financière présentés dans le tableau ci-dessous.

**Tableau 1 : Revue de littérature sur la théorie d'intermédiation financière**

<u>Chercheurs</u>	<u>Revue de littérature</u>
J. Gurley et E. Shaw (1960)	Ont souligné que le rôle d'intermédiation financière consiste à produire de la liquidité à travers la transformation des actifs financiers. Cette transformation constitue l'origine de risque du fait que les établissements bancaires financent les emplois longs par des ressources à plus courte durée.
R. Merton (1995)	En tant qu'intermédiaire financier, la banque reçoit des dépôts en contrepartie, elle les prête aux agents qui ont un besoin de financement.
De Servigny et Zelenko, (2003), p. 34	Ils définissent la banque comme un intermédiaire financier monétaire, son activité réside dans l'intermédiation entre les agents à excédent de ressources et les agents à déficit de financements.
Diamond (1984)	Selon la théorie de l'intermédiation financière, la banque est considérée comme un surveillant qui exerce une surveillance active sur les crédits (monitoring).
(Leland et Pyle, 1977 ; Campbellet Kracaw, 1980).	Cette théorie de l'intermédiation financière permet d'analyser le fonctionnement du marché de crédit bancaire et la politique de crédit adoptée par l'établissement bancaire.
Plihon (1998)	Souligne que la prise de risque est primordiale pour la réalisation des profits. La rémunération des risques représente des profits pour la banque.
Ramakrishman et Thakor (1984) et Millon et Thakor (1985),	Dans la théorie de l'intermédiation financière, les banques sont représentées comme des institutions depositaires disposent de la capacité à offrir de la liquidité et améliorer des services de surveillance.
Diamond et Dybvig (2002),	L'offre de l'assurance et de la sécurité aux déposants est un avantage qui attribue à la banque de la valeur. Grâce à la collecte des dépôts, les banques peuvent assurer une meilleure distribution de risque entre les actionnaires.

Source : *Amal Ben Hassena*<sup>9</sup>, 2006

## 6.2. La théorie d'asymétrie d'information

Le risque de crédit peut trouver une explication dans l'asymétrie informationnelle existante entre des agents n'étant pas identiquement dotés d'un même niveau d'information. En effet, l'asymétrie informationnelle peut conduire à une allocation inefficace du crédit et donc à un important risque de crédit.

<sup>9</sup> Amal Ben Hassena, 2006 « L'impact de la libéralisation financière sur l'intermédiation bancaire » Université de Sfax : Ecole Supérieure de Commerce de Sfax - Diplôme de maîtrise en Hautes Etudes commerciales 2006

Le phénomène qualifié « **d'asymétrie d'information** » a fait l'objet d'études de plusieurs auteurs à travers différentes théories permettant de comprendre le comportement des banques dans le processus de financement.

**Tableau 2 : Les principales études sur la théorie d'asymétrie d'information**

<u>Auteur</u>	<u>Revue de littérature</u>
Akerlof (1970), Diamand et Dybvig (1975)	Le phénomène d'asymétrie d'information conduit à des comportements opportunistes de la part de l'emprunteur.
Jaffée et Russel (1976)	Ils distinguent dans leurs modèles les emprunteurs honnêtes et ceux malhonnêtes. Ils assimilent les emprunteurs honnêtes aux entreprises qui, estimant les coûts d'une éventuelle défaillance de leurs projets élevés, décident de rembourser. Quant aux emprunteurs malhonnêtes, ils préfèrent faire faillite en proposant des taux d'intérêt plus élevés.
Calomiris et Hubbard, (1990)	Montrent que les firmes ont plus d'information sur le niveau de risque de leurs projets, elles sont dotées de plus d'informations sur l'utilisation des fonds empruntés.
(Fraser et al, 2001)	L'asymétrie d'information se manifeste lorsque les emprunteurs doivent disposer de plus d'information sur leurs entreprises que les bailleurs de fonds.
(Sylvie Cieply et Bernard Parañaque, 1997)	Les banques ne disposent pas de toutes les caractéristiques des projets des entreprises. En même temps, ces firmes n'ont aucune idée sur les stratégies envisagées par les banques par rapport à la décision d'octroi des crédits.

Source : NEZIEN<sup>10</sup>, 2010

## 7. Analyse comparative des techniques du crédit scoring

De toutes ces techniques de scoring présentées précédemment, on peut assister à des avantages et des limites propres à chaque méthode. Le tableau suivant présente les différentes méthodes de scoring ainsi que leurs avantages et leurs inconvénients :

**Tableau 3 : Tableau récapitulatif des techniques de scoring**

<u>Technique de scoring</u>	<u>Avantages</u>	<u>Inconvénients</u>	<u>Règle classification</u>
<u>Analyse discriminante</u>	Des prédictions explicites.	Variables explicatives continues et sans valeurs manquantes.	Score d'appartenance à une classe.
	Un résultat analytique direct.	Sensible aux individus hors norme.	
	Des calculs très rapides	Absence de tests statistiques de significativités des coefficients.	

<sup>10</sup> Théophile Fabrice NEZIEN, 2010 « Problématique du financement des PME par les établissements bancaires : cas de Coris bank ». Mémoire pour l'obtention du diplôme de Maîtrise en Economie et Gestion des Entreprises et des Organisations (EGEO)- Université Saint Thomas d'Aquin.

	Ne nécessite pas un échantillon de grande taille pour l'apprentissage.		
	Tient compte des variables qualitatives (procédure DISQUAL)		
<b><u>Régression logistique</u></b>	Variables explicatives discrètes, qualitatives ou continues.	Les variables explicatives doivent être non colinéaires.	Probabilité que l'évènement de défaut se produise
	Variables à expliquer ordinale ou nominale.	Calcul itératif plus long qu'une analyse discriminante de Fisher.	
	Pas d'hypothèses de multi normalités, ni d'homoscédasticités pour les variables explicatives.	La précision est moindre que celle de l'analyse discriminante.	
	Possibilité de prise en compte les interactions entre variables.	La régression logistique ne converge pas toujours vers une solution optimale.	
	Résultats faciles à interpréter.	Ne traite pas les valeurs manquantes. Sensible aux valeurs hors norme	
<b><u>Réseaux de neurones</u></b>	Modéliser des relations non linéaires entre les données.	Les résultats ne sont pas explicites et sont difficile à comprendre par les utilisateurs	Affecter l'appartenance des individus aux classes définies
	Modéliser des problèmes de différents types.	Le risque de sur-apprentissage.	
	Résiste aux données défectueuses.	Ne traite pas un grand nombre de variables. La convergence vers la meilleure solution globale n'est pas toujours garantie	
<b><u>Arbres de décision</u></b>	Les résultats sont exprimés sous forme de condition explicites sur les variables d'origine.	La détermination des nœuds du niveau (n+1) dépend fortement du nœud précédent (n).	Associer une observation à l'attribut attaché à la feuille à laquelle il appartient.
	Compréhensibilité des résultats pour les utilisateurs.	L'apprentissage d'un arbre de décision nécessite un nombre assez grand d'individus.	
	Les variables explicatives peuvent ne pas suivre des lois probabilistes particulières	Le score d'un individu dépend de la feuille à laquelle le conduisent les valeurs de ses prédicteurs.	
	Les arbres ne sont pas affectés par les individus hors norme.		
	Traite les données manquantes.		



	Tous types de variables : continues, discrètes et qualitatives.		
	Simple à utiliser.		

Source : *Adapté de Tufféry, 2012*

Le pouvoir discriminant du modèle de scoring et sa capacité de mesurer la détresse financière des emprunteurs représentent des critères déterminants dans la performance et la qualité du modèle de scoring.

## 8. Conclusion

La maîtrise du risque de crédit demeure une préoccupation majeure et un objectif recherché pour les banques qui ont pour but la détention des modèles de scoring efficaces capables de classer les firmes dans des classes de risques afin de distinguer entre les bons et les mauvais emprunteurs. En effet, la mise en place de ces techniques permet de d'obtenir et d'une manière rapide et efficace les risques potentiels sur chaque ligne de crédit.

En outre, ces modèles de scoring nécessitent la présence des informations historiques sur l'emprunteur, ces informations représentent un ensemble de caractéristiques permettant de prévoir si celui-ci aura une solvabilité future.

Dans notre étude, nous avons analysé les différentes techniques de prédiction du risque de crédit, particulièrement leurs caractéristiques, leur fonctionnement, leurs avantages et leurs limites.

En effet, en plus des méthodes traditionnelles du crédit scoring (comme la régression logistique et l'analyse discriminante), on trouve les techniques intelligentes (comme les réseaux de neurones et les arbres de décision, systèmes experts), récemment utilisées dans plusieurs domaines comme la médecine, le marketing, etc..., ils ont aussi démontré leurs capacités d'évaluer le risque de crédit des emprunteurs.

Il serait donc intéressant de reprendre cette étude en utilisant un échantillon des entreprises saines et défaillantes et appliquer deux techniques de crédit scoring une traditionnelle et une autre appartenant aux techniques d'intelligence artificielle et ainsi comparer les résultats obtenus.

## Références

- (1) Abdeljelil Farhat, Sami Mestiri, 2008. « Détection de la défaillance des entreprises Tunisiennes par la régression logistique semi paramétrique et les réseaux de neurones. ». P :19
- (2) AKERLOF G., 1970, "The Market for "Lemons": Quality, Uncertainty and the Market Mechanism ", Quaterly Journal of Economics, vol.84 n°3, August.
- (3) ALTMAN E.I. [1968], "Financial ratios, discreminant analysis and the prediction of corporate bankruptcy", in Journal of Finance, vol. 23, p. 589609.
- (4) Altman, E. I., Marco, G., Varetteo, F. (1994), Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Network – The Italian Experience, Journal of Banking and Finance, vol. 18, n° 3, pp. 505-529.
- (5) Amal Ben Hassena, 2006 « L'impact de la libéralisation financière sur l'intermédiation bancaire » Université de Sfax : Ecole Supérieure de Commerce de Sfax - Diplôme de maîtrise en Hautes Etudes commerciales 2006
- (6) Anandarajan, Lee et Anandarajan, (2001) Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks June 2001 Intelligent Systems in Accounting Finance & Management 10(2) :69-81

- (7) Aomar Ibourk & Abdelkrim Aazzab, 2016 « Modélisation Des Défaillances D'entreprises Au Maroc : Une Étude Économétrique ». Publié dans *Global Journal of Management and Business Research: B Economics and Commerce*.
- (8) Anaclet Ananga Onana, Joseph Bikay Bi Batoum et Robert Wanda, 2013 « Les déterminants de la compétitivité des franchises internationales : une analyse empirique du cas du Cameroun » Dans *Revue Congolaise de Gestion* 2013/2 (Numéro 18), pages 9 à 44
- (9) Anderson, R. (2007). "The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation". New York: Oxford University Press.
- (10) Aubier, M. (2007), « Impact de Bâle II sur l'offre de crédit aux PME. », *Economie & prévision* 2/2007 (n° 178-179), p. 141-148. Pour plus de détail consulter : URL: [www.cairn.info/revue-economie-et-prevision](http://www.cairn.info/revue-economie-et-prevision).
- (11) BANQUE DE FRANCE (2010), « Crise financière : origines et dynamiques », Documents et Débats, Chapitre 1, n° 3, janvier, p. 9.
- (12) Bardos M. (1998), Detecting the risk of company failure at the Banque of France, *Journal of Banking and Finance*, 22, 1405-1419.
- (13) Bardos M., Zhu W.H. (1997), Comparaison de l'analyse discriminante linéaire et des réseaux neuronaux : application à la détection de défaillance d'entreprises, *Revue Statistique Appliquée*, XLV (4), 65-92.
- (14) BARDOS Mireille, 2008, « Scoring sur données d'entreprises : instrument de diagnostic individuel et outil d'analyse de portefeuille d'une clientèle », *Revue MODULAD*, N°38.
- (15) BEAVER, W. H. (1966), «Financial Ratios as Predictors of Failure », *Empirical Research in Accounting, Selected Studies, Journal of Accounting Research, Supplement*, vol. 4, pp. 71-111.
- (16) Benoit Crabbé, (2011) « Notions de modélisation statistique : Predicting the Dative Alternation »
- (17) Berg Eugène. Antony Sampson, 1982, « Les banquiers dans un monde dangereux ». In: *Politique étrangère*, n°4 - 1982 - 47<sup>e</sup>année. p. 1051.
- (18) Bonfim, D., (2009), «Credit Risk Drivers: Evaluating the Contribution of Firm Level Information and Macroeconomic Dynamics», *Journal of Banking and Finance*, Vol. 33, No. 2.
- (19) Boubacar Diallo, 2006 « Un modèle de crédit scoring pour une institution de micro finance Africaine : Le cas de NYESIGISO AU MALI ». HAL Id : halshs-00069163. P : 4
- (20) Catherine Refait-Alexandre, 2004 « La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux Dans *Économie & prévision* 2004/1 (no 162), pages 129 à 147
- (21) Catherine Refait-Alexandre, 2004 « La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux Dans *Économie & prévision* 2004/1 (no 162), pages 129 à 147.
- (22) Cécile Kharoubi, Philippe Thomas : « Gestion de risque de crédit : Banque & marché », Edition 2016
- (23) Charalambous C, Charitou A et Kaourou F. (2000), Comparative analysis of artificial neural network models: application in bankruptcy prediction, *Annals of operations research*, Octobre, university of Cyprus.
- (24) Cristián Bravo, Thomas, Lyn C, Richard Weber, (2017 ) Improving credit scoring by differentiating defaulter behaviour, University of Southampton Highfield Southampton UK & Richard Weber Pages 771-781 *Journal of the Operational Research Society*. Volume 66, 2015 - Issue 5

- (25) Desjardins Julie (2015) L'analyse de régression logistique September 2005 Tutorials in Quantitative Methods for Psychology 1(1) DOI: 10.20982/tqmp.01.1.p035
- (26) Douglas W. Diamond, (1984) Financial Intermediation and Delegated Monitoring. The Review of Economic Studies. Vol. 51, No. 3 (Jul.,1984), pp. 393-414 (22 pages)
- (27) Fred Ntoutoume Obiang-Ndong, 2007 « Scoring du risque de crédit des PME par la modélisation statistique et l'intelligence artificielle chez l'UMECUDEFS: une application comparative de la régression logistique et des réseaux de neurones ». Mémoire de fin d'études, Université Cheikh Anta Diop - Faculté Des Sciences Economiques Et De Gestion.
- (28) Gilbert Saporta (2015) Classification supervisée et credit scoring Conservatoire National des Arts et Métiers, Paris.
- (29) Hand, D.J. and Henley, W.E. (1997) Statistical Classification Methods in Consumer Credit Scoring: A Review. Journal of Royal Statistical Society, 160, 523-541.
- (30) Jantzen. J (1998), "Introduction to perceptron Networks ", report no 98-H 873 (nnet), 25 Oct, Denmark, page27.
- (31) Julie DESJARDINS, « L'analyse de la régression logistique », Université de Montréal, Tutorial in Quantitative Methods for Psychology, 2005, Vol. 1(1), p. 35-41.
- (32) M. BARDOS W. H. ZHU (1997), Comparaison de l'analyse discriminante linéaire et des réseaux de neurones. Application à la detection de défaillance d'entreprises "Revue de statistique appliquée, tome 45, no 4 (1997), p. 65-92".
- (33) R. Anderson (2007) The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation Illustrated Edition.
- (34) Rim BOUSSAADA, 2012 « L'impact de la gouvernance bancaire et de la relation bancaire sur le risque de crédit: cas des banques tunisiennes ». Thèse de doctorat en sciences économiques, Université Montesquieu - bordeaux iv, université de Tunis institut supérieur de gestion de Tunis.
- (35) Stéphane Tufféry, « Data Mining et Statistique décisionnelle, l'intelligence des données », Editions Technip (2e Ed), 2007, p.25.
- (36) Théophile Fabrice NEZIEN, 2010 « Problématique du financement des PME par les établissements bancaires : cas de Coris bank ». Mémoire pour l'obtention du diplôme de Maîtrise en Economie et Gestion des Entreprises et des Organisations (EGEO)- Université Saint Thomas d'Aquin.
- (37) Thuillier. D (1997), « Principe et applications des réseaux de neurones –deux illustration sur l'habitat au Maroc »; revue région et développement n°=5-, page3
- (38) (Varetto, 1998; Barney et al., 1999, Shin et Lee, 2002; Chen et Du, 2009; Ravisankar et Ravi, 2010), les techniques de support vector machine.  
Yang Liu, (2001), New Issues in Credit Scoring Application, Work report N°16, Institut für Wirtschaftsinformatik, 2001