



**Institut de Financement
du Développement du
Maghreb Arabe**

**La mise en place d'un système de notation interne :
« Cas de la banque de l'Habitat »**

Elaboré par :

Bchir Nesrine

Encadré par :

Mr.Radhouene Gouja

Table des matières

Introduction générale.....	1
Chapitre 1 : Le risque de crédit et la réglementation bancaire.....	4
Introduction.....	5
Section 1 : Présentation du risque de crédit	6
1.1. Concept du risque.....	6
1.2. Typologie des risques bancaires.....	7
Section 2 : La réglementation prudentielle internationale	11
2.1. Comité de Bâle, présentation et missions.....	11
2.2. L'accord de Bâle I.....	11
2.3. L'accord de Bâle II.....	13
2.4. Bâle III, une suite nécessaire	17
Section 3 : La réglementation prudentielle nationale.....	21
3.1. La division et la couverture du risque	21
3.2. La classification des actifs et la constitution des provisions	22
3.3. La mise en place d'un système de notation interne des contreparties	23
Section 4 : Les paramètres bâlois du risque de crédit.....	24
4.1. La probabilité de défaut	24
4.2. La perte en cas de défaut (Loss Given Default : LGD)	24
4.3. Exposition au moment du défaut (Exposure at Default : EAD)	25
4.4. La maturité	25
4.5. La perte attendue (Expected Loss)	25
Conclusion	25
Chapitre 2 : Méthodologie de notation.....	27
Introduction.....	28
Section 1 : Approche par les notations externes	29
1.1. Définitions	29
1.2. Les agences de notation	29
1.3. Processus de rating et échelle de notation	30

1.4. Notation externe : Méthode de calcul du risque de crédit selon Bâle II.....	32
Section 2 : Approche par les notations internes	33
2.1. Définition et historique de « Credit Scoring »	35
2.2. L'analyse discriminante	36
2.3. La régression logistique	38
2.4. Les réseaux de neurones	42
2.5. Les arbres de décision	45
2.6. Comparaison entre les différentes méthodes statistiques	49
Conclusion	50
Chapitre 3 : Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la Banque de l'Habitat.....	52
Introduction.....	53
Section 1 : Présentation de la BH	54
1.1. La présentation de la BH	54
1.2. Activité et performance de la banque.....	54
Section 2 : Données et méthodologie de l'étude.....	58
2.1. L'échantillon de l'étude	58
2.2. Choix des variables de l'étude	60
Section 3 : Construction de la fonction score.....	66
3.1. Analyse discriminante	66
3.2. La régression logistique	75
3.3. Arbre de décision.....	82
3.4. Réseaux de neurones artificiels.....	84
3.5. La comparaison du pouvoir prédictif des modèles	90
3.6. Transformation des probabilités de défaut en rating	91
Conclusion	93
Conclusion générale	95
Références bibliographiques	98

Liste des figures

Figure 1 : L'évolution des créances impayés et en contentieux du secteur bancaire des crédits aux entreprises et aux professionnels	6
Figure 2 : L'évolution des créances impayés et en contentieux du secteur bancaire des crédits aux particuliers	6
Figure 3 : Les trois piliers de Bâle II	14
Figure 4 Les six parties de l'accord de Bâle III	18
Figure 5. Processus de notation par les agences de rating	30
Figure 6 : Processus du Credit Scoring	36
Figure 7 :Règles de décision pour le modèle d'Altman	37
Figure 9 . Règles de décision pour le modèle Conan et Holder	38
Figure 9 : Schéma général du traitement	43
Figure 10 : Conception d'un réseau de neurones	44
Figure 12 : Arbre de décision	46
Figure 13 La structure du PNB de la BH en 2016/2017	56
Figure 14 La répartition des deux sous-échantillons	60
Figure 15 La courbe ROC analyse discriminante.....	73
Figure 16 Courbe ROC régression logistique	81
Figure 17 L'arbre de décision en utilisant le modèle CHAID	82
Figure 18 La généralisation d'un modèle.....	85
Figure 19 Le réseau de neurone artificiel propre à notre étude	88
Figure 20 :La courbe ROC RNA.....	89
Figure 21 La croissance du taux de défaut	93

Liste des tableaux

Tableau 1 : Pondération selon type de risque de l'actif sous Bâle I.....	12
Tableau 2 Correspondance entre les notes attribuées par les agences de notations	31
Tableau 3 : Pondérations proposées par le Comité de Bâle par nature de contrepartie et par note.	32
Tableau 4 : Les taux de bon classement trouvés par F. Mraïhi (2015).....	42
Tableau 5 : Comparaison entre l'approche neuronale et l'analyse discriminante	45
Tableau 6 : Tableau des effectifs lors du croisement des deux variables	47
Tableau 7 : Comparaison entre les différentes méthodes statistiques (adapté de Tufféry, 2012)	49
Tableau 8 Les principaux indicateurs de la BH en 2016, 2017 et le premier trimestre de l'année 2018	54
Tableau 9 Les ratios de la banque de l'Habitat en 2016/2017	56
Tableau 10 Les créances classées de la BH en 2016 et 2017	58
Tableau 11 Répartition de l'échantillon initial.....	59
Tableau 12 La répartition de l'échantillon de construction entre les deux groupes	59
Tableau 13 La répartition de l'échantillon test entre les deux groupes.....	59
Tableau 14 La liste des ratios	62
Tableau 16 Statistiques descriptives des ratios	65
Tableau 17 Vérification de l'existence de différences entre les sous-groupes.....	67
Tableau 18 Le pouvoir discriminant des ratios	68
Tableau 19 Résultat du Test de Box.....	70
Tableau 20 Valeurs propres.....	71
Tableau 21 Le Lambda de Wilks.....	71
Tableau 22 Le taux de bon classement de l'analyse discriminante	72
Tableau 23 Le pouvoir discriminant du modèle par le critère AUC	72
Tableau 24 Fonction aux barycentres des groupes.....	74
Tableau 25 Règle de décision.....	74
Tableau 26 La validation de l'analyse discriminante.....	74
Tableau 27 Résultat de la régression logistique	75
Tableau 28 Odds ratios.....	77
Tableau 29 Test du rapport de vraisemblance	78
Tableau 30 : Les tests de Cox et Snell et Nagelkerke	79
Tableau 31 Le taux de bon classement de la régression logistique	79
Tableau 32 Test de Hosmer -Lemeshow	80
Tableau 33 La validation de la régression logistique	81
Tableau 34 : Le taux de bon classement de l'arbre CHAID.....	83
Tableau 35 : Coût d'un arbre de décision	84
Tableau 36 : Récapitulatif du modèle.....	88
Tableau 37 :Zone sous la courbe	88
Tableau 38 : La classification des données	89
Tableau 39 :Comparaison entre les résultats des modèles utilisés	90
Tableau 40 La stabilité de la régression logistique.....	91
Tableau 41 Construction de l'échelle de notation.....	92

Introduction générale

Introduction générale

A partir du XVII^e siècle, depuis que la banque moderne a commencé à évoluer, la sphère économique et internationale a connu de multiples mutations accompagnées par un perpétuel mouvement de l'activité bancaire.

Cet essor n'est cependant pas gratuit. En effet, le revers de la médaille consiste à favoriser l'exposition du système bancaire aux diverses formes de risques (opérationnel, de marché, de crédit, etc....) notamment si on appartient à un environnement caractérisé par un univers financier incertain où l'économie est imparfaite, sans scrupule et la compétitivité est à son apogée.

C'est dans un tel cadre, que fut instauré le comité de Bâle, régulateur suprême de l'activité bancaire à l'échelle internationale dont sa principale mission est de promouvoir la stabilité et la sécurité du système financier. C'est sous l'impulsion de Bâle II, la mesure et la gestion du risque de crédit sont devenues l'affaire centrale de chaque établissement bancaire et ce, par le biais de modèles internes exigés par les autorités visant un calibrage personnalisé de fonds propres dits « économiques ». Il y a eu donc lieu d'amorçage de techniques plus fines et plus sensibles au risque qui procurent aux banques une marge de manœuvre modulable selon le risque réel encouru par la contrepartie.

Ce type de risque est le risque le plus répandu. En effet, l'histoire mondiale est remplie d'exemples de lourdes conséquences de la mauvaise anticipation et gestion de ce type de risque (la crise mondiale des Subprimes de 2008, qui d'après **Couderc et al (2010)**, la crise la plus longue et la plus sévère observée depuis celle des années 30).

A cet effet, les autorités prudentielles s'engagent souvent à réviser leur accord afin de garantir la pérennité des banques et donc la stabilité de l'économie. Les banques à leur tour, doivent se conformer à ces réglementations pour pouvoir quantifier les risques, les prévoir, les contrôler et enfin les mitiger. De même, la Banque Centrale de Tunisie a mis en place un plan d'actions quinquennal à l'horizon 2020 visant la convergence vers le cadre prudentiel de Bâle II et Bâle III.

Dans ce contexte, l'enjeu majeur du banquier réside dans la recherche sans relâche d'un équilibre optimal entre une rentabilité souhaitée et un risque inévitable du fait que le risque zéro est impossible dans le cas où on ambitionne de réaliser la rentabilité. Il faut donc

disposer des outils nécessaires pour mesurer, évaluer et assurer une gestion maîtrisée du risque afin de maximiser le profit.

Dans cet ordre d'idées, le présent travail vise à traiter le sujet suivant deux grandes lignes : réglementation prudentielle et modélisation mathématique. Il tente à répondre à la problématique suivante :

« Dans un environnement financier de plus en plus instable, et dans le but de préserver au mieux sa rentabilité et garantir sa part de marché, comment la banque peut-elle évaluer le risque de crédit en élaborant un modèle de notation interne qui vise la prévision du risque de crédit des entreprises par la méthode du scoring ? »

D'autres interrogations s'adosent à cette problématique et qui s'articulent autour des points suivants :

- ✓ Quelle est la typologie des risques bancaires ? et comment sont-ils perçus par la réglementation prudentielle internationale et nationale ?
- ✓ Quelles sont les méthodes qui permettent la construction d'un modèle de prévision de risque de crédit ?
- ✓ Quelle est la méthode qui permet de prévoir au mieux la défaillance des entreprises et quels sont les principaux indicateurs qui expliquent cette dernière ?

Afin de répondre à tous ces questions, nous procéderons à une étude analytique structurée selon le schéma suivant :

1. Première partie : Cadre théorique

Dans le premier chapitre, nous présenterons la typologie des risques bancaires et la réglementation bancaire en matière de risque de crédit au niveau international et national. Ensuite, nous définirons les paramètres bâlois en termes de risque de crédit.

Dans le deuxième chapitre, nous aborderons les méthodologies de notation externe et interne (une revue de littérature sur les principales méthodes de « Credit Scoring »).

2. Deuxième partie : Mise en pratique

Cette partie sera scindée en deux sections : la première contiendra une présentation de l'organisme d'accueil et la deuxième sera consacrée à une étude empirique sur un échantillon d'entreprises appartenant à la banque de l'Habitat. Cette dernière portera sur une comparaison de la performance entre les différentes méthodes utilisées. Une échelle de notation interne sera par la suite établie en fonction des résultats obtenus.

Nous achèverons notre travail avec une conclusion générale récapitulant l'essentiel des aboutissements de notre recherche.

Chapitre 1 : Le risque de crédit et la réglementation bancaire

Chapitre 1 : Le risque de crédit et la réglementation bancaire

Introduction

Selon **Petit-Dutaillis(1967)**, « *faire crédit, c'est faire une confiance, mais c'est aussi donner librement la disposition effective et immédiate d'un bien réel ou d'un pouvoir d'achat, contre la promesse que le même bien ou un bien équivalent vous sera restitué dans un certain délai, le plus souvent avec rémunération du service rendu et du danger couru, danger de perte partielle ou totale que comporte la nature même du service* ».

Dans le même ordre d'idées, **Bernard & al(1989)** définit le crédit comme « *acte de confiance comportant l'échange de deux prestations dissociées dans le temps, biens ou moyens de paiement contre promesse ou perspective de paiement ou de remboursement* ».

De ces deux définitions, on peut conclure que l'octroi d'un crédit se base sur trois notions principales :

- ✓ La confiance entre les parties contractantes,
- ✓ Le facteur temps,
- ✓ La promesse de remboursement du bien prêté.

La banque, en tant qu'une institution financière, est au cœur du système de financement de l'économie : Tout le monde (particuliers, professionnels, entreprises, etc....) s'adresse à sa banque afin d'acquérir un actif ou un bien qu'il n'a pas les moyens de financer par leur fonds propres. Cette opération implique systématiquement un risque de crédit qui est né dès l'octroi de crédit. Le poids de ce risque sur l'activité bancaire peut être illustré par la Figure.1 et la Figure.2.

Étant donné l'importance de l'activité de l'intermédiation au sein des banques, les autorités à l'échelle internationale (comité de Bâle) et à l'échelle nationale ont mis en place un ensemble de réglementation dans le but d'atténuer le risque lié à cette activité et harmoniser les normes prudentielles.

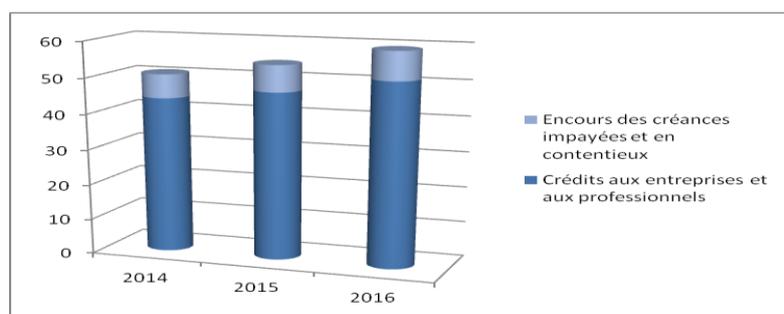


Figure 1 : L'évolution des créances impayées et en contentieux¹ du secteur bancaire des crédits aux entreprises et aux professionnels²

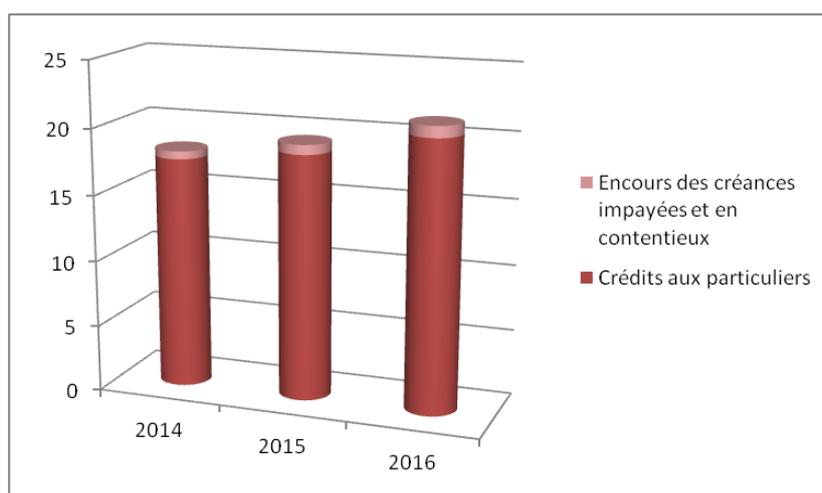


Figure 2 : L'évolution des créances impayées et en contentieux du secteur bancaire des crédits³ aux particuliers

Ce chapitre s'articule comme suit : Dans une première section nous allons présenter la notion de risque de crédit, ses différentes formes et ses caractéristiques, dans une seconde nous allons nous focaliser sur les règles prudentielles à l'échelle internationale. Dans une troisième section, nous allons passer aux règles prudentielles à l'échelle nationale et la dernière sera consacrée à la définition des paramètres bâlois du risque de crédit.

Section 1 : Présentation du risque de crédit

1.1. Concept du risque

Le risque peut être défini comme tout événement ou action qui peut empêcher l'entreprise d'atteindre des objectifs fixés, d'appliquer une stratégie envisagée ou d'assurer sa pérennité. Les déviations positives sont généralement vues comme une opportunité. Toutefois, les déviations négatives sont considérées comme un risque. On adopte les définitions suivantes :

¹ Exprimés en milliard de dinar.

² Comportent les entreprises publiques et privées opérant dans l'industrie, l'agriculture et pêche et le service.

³ Répartis en crédits logement et crédits à la consommation.

« *Le risque est défini comme la mesure d'un ensemble d'éléments de la situation dangereuse qui, combinés à des conditions particulières d'environnement, redoutées ou non, connues ou non, peuvent entraîner des conséquences préjudiciables ou accidentelles* ⁴».

« *Le risque est une situation (ensemble d'événements simultanés ou consécutifs) dont l'occurrence est incertaine et dont la réalisation affecte les objectifs de l'entreprise qui le subit* ».⁵

Quant à eux, **Aubert et Bernard (2004)** définissent le risque comme « *la probabilité d'un événement et l'ampleur de son impact* ».

D'après ces définitions, la notion de risque est étroitement associée, en premier lieu, à celle d'incertitude dans la mesure où elle reste dans le domaine de la probabilité (caractère aléatoire et imprévisible). En deuxième lieu, aux effets négatifs sur l'entité qui subit le risque.

1.2. Typologie des risques bancaires

Dans le système bancaire, **Armand de Servigny** a distingué dans son livre « *Le risque de crédit : Nouveaux enjeux bancaires(2010)* », une panoplie de risques nuisant l'activité de la banque, citons le risque de crédit, le risque de taux, le risque de change, le risque de liquidité, le risque pays, etc....

En d'autres termes, le risque bancaire peut être dû à une insuffisance de diversification, une insuffisance de liquidité, une forte volatilité des marchés financiers ou la volonté même des banques à s'exposer aux risques.

Plusieurs classifications peuvent être retenues. Cependant, on opte la classification proposée par le comité de Bâle II qui distingue à son tour trois grandes catégories : le risque de marché, le risque opérationnel et le risque de crédit.

1.2.1. Le risque de marché

Lors de l'exercice de leurs activités, les banques interviennent sur les marchés financiers en réalisant des opérations d'achat et/ou de vente des devises, actions, obligations, etc. Ces opérations sont exposées à la fluctuation des cours de marchés qui peuvent être favorable ou défavorable. Le risque de marché peut être défini ainsi comme le risque de pertes sur les positions prises (courtes ou longues) suite aux fluctuations des marchés. Ce risque concerne toute une gamme d'instruments financiers : obligations, actions, dérivés de taux, change, matières premières, etc.

⁴ Desroches. A, la gestion des risques, Ed Lavoisier, 2015, P.13.

⁵ Barthélémy, Gestion des risques méthodes d'optimisation globale, Ed d'organisation, 2001, P.13.

Le risque de marché peut être calculé à l'aide de modèles internes, tel que le modèle de valeur à risque (VAR) qui est définie par **Engle et Manganelli (2001)** comme « *la perte maximale potentielle qui ne devrait être atteinte qu'avec une probabilité donnée sur un horizon donné* ».

On distingue généralement trois catégories de risque de marché :

✓ ***Le risque de taux***

Le rôle d'intermédiation au sein des banques les expose structurellement au risque de transformation ou de taux, en finançant des actifs à long terme avec des dépôts ou d'autres formes de financements souvent libellés à court terme. Cette situation s'explique par deux catégories d'arguments (**Maes 2004**) : En premier lieu, cette intermédiation fait subir la banque des impasses de maturité⁶ et/ou de taux⁷. En deuxième lieu, les banques peuvent choisir de s'exposer volontairement à des impasses de taux potentiellement annulées ou réduites par une utilisation d'instruments de couverture, soit parce que leur prévision d'évolution de la courbe des taux d'intérêt diffère de celle du marché, soit parce qu'ils jugent la prime de risque⁸ est attractive.

✓ ***Le risque de change***

C'est le risque de perte ou de gain lié aux variations du cours de change. En d'autres termes, il désigne les pertes ou les gains éventuels réalisés du fait des variations du taux de change entre la monnaie locale et la monnaie étrangère. Le risque de change se mesure par la position de change qui se définit comme le solde net des avoirs dans une devise donnée. En effet, si la banque est en position longue, elle risque la baisse du cours de change. En revanche, si elle est en position courte, une hausse du taux de change serait défavorable.

✓ **Le risque de variation de cours des actions et des matières premières**

Les variations subites et souvent imprévisibles de l'offre ou de la demande sur les marchés au comptant des matières premières et des actifs financiers sont à l'origine des fortes fluctuations qui caractérisent leur prix. Une telle instabilité du prix au comptant représente un risque non négligeable pour tout agent économique dont l'activité principale est reliée à ces biens.

⁶ Pour une échéance donnée, les actifs et les passifs ne sont pas de même montant.

⁷ L'impatte de taux mesure la sensibilité de la marge d'intérêt (solde des intérêts reçus et versés résultant des opérations sur le portefeuille bancaire) aux fluctuations des taux d'intérêt du marché.

⁸ C'est le rendement marginal que le détenteur d'un titre à long terme s'attend à recevoir par rapport au rendement d'un titre à court terme.

1.2.2. Le risque opérationnel

Les accords de Bâle II définissent le risque opérationnel comme « *le risque de pertes résultant de carences ou de défauts attribuables à des procédures, personnels et systèmes internes ou à des événements extérieurs* »⁹. Cette définition inclut le risque juridique (l'exposition à des amendes, pénalités, et dommages pour faute résultant de l'exercice de surveillance prudentielle ainsi que de transactions privées) et exclut les risques stratégiques et d'atteinte à la réputation.

Les travaux de Bâle ont défini une segmentation des risques en huit lignes métiers et sept catégories d'événements qui permettent de constituer une matrice à 56 cases représentatives de l'ensemble des activités bancaires et leurs risques associés.

Les événements compris dans ce risque sont les suivants :

- ✓ Fraude interne,
- ✓ Fraude externe,
- ✓ Insuffisance des pratiques internes concernant les ressources humaines et la sécurité du lieu de travail,
- ✓ Négligences des règles clients, produits et pratiques commerciales,
- ✓ Dommages aux actifs corporels,
- ✓ Interruption d'activité et dysfonctionnement des systèmes,
- ✓ Dysfonctionnement des processus de traitement (exécution, passation, livraison, produit finis).

Compte tenu de la multiplicité des causes relatives au risque opérationnel, les accords de Bâle ont mis en place trois méthodes afin de déterminer les exigences en fonds propres au titre du risque opérationnel, à savoir deux méthodes forfaitaires (approche indicateur de base et approche standardisée) et une de mesure directe (approche de mesure complexe, produite par le système interne de la banque sur la base des critères quantitatifs et qualitatifs propres à l'organisation interne et à l'activité de la banque).

1.2.3. Le risque de crédit

Étant donnée l'importance de l'activité d'intermédiation des banques, le risque de crédit détient la part la plus pesante dans les défaillances bancaires. Selon **Coussergues (2007)**, il s'agit d'un « *risque inhérent à l'activité d'intermédiation traditionnelle et qui correspond à la défaillance de la contrepartie sur laquelle une créance ou un engagement est détenu* ». Dans

⁹ BRI, Comité de Bâle sur le contrôle bancaire, convergences internationales de la mesure et des normes de fonds propres, version compilée, Juin 2006.

ce cas, la dégradation de la situation financière de l'emprunteur représente la principale cause du non respect de ses engagements dans les échéances prévues.

De son côté, **Heem (2000)** le définit pour un banquier comme « *le risque de voir son client ne pas respecter son engagement financier, à savoir, dans la plupart des cas, un remboursement de prêt* ».

Concernant le comité de Bâle, il le définit comme étant « *le risque de non remboursement associé à un prêt accordé par une banque* ». A travers cette définition, **Cohen¹⁰ (1997)** et **Jacob & Sardi¹¹(2001)** ont fait ressortir trois dimensions pour le risque de crédit :

✓ *Le risque de défaut*

C'est le risque que l'emprunteur ne soit pas capable de rembourser le crédit contracté dans les délais prévus.

Calvet (1997) le définit comme « *le risque de perte lié à la défaillance d'un débiteur sur lequel l'établissement de crédit détient un engagement, quelle que soit la nature du débiteur et la forme de cet engagement* ». Cette définition est très proche de celle de Moody's¹² qui définit le risque de défaut comme « *tout manquement ou tout retard sur le paiement du principal ou des intérêts* ».

Selon l'article 414 du nouvel accord de **Bâle II** publié le 23 avril 2003, un défaut est constaté si l'un ou les deux événements ci-dessous se réalisent :

- 1- « *La banque estime improbable que le débiteur rembourse en totalité son crédit au groupe bancaire sans qu'elle ait besoin de prendre des mesures appropriées telles que la réalisation d'une garantie (si elle existe)* ».
- 2- « *L'arriéré du débiteur sur un crédit important dû au groupe bancaire dépasse 90 jours. Les découverts sont considérés comme des créances échues dès que le client a dépassé une limite autorisée ou qu'il a été averti qu'il disposait d'une limite inférieure à l'encours actuel* ».

✓ *Le risque de dégradation de la qualité du crédit (downgrading risk)*

Il se traduit par la dégradation de la situation financière de l'emprunteur. Ceci implique un accroissement de la probabilité de défaut, même si le défaut ne survient pas réellement. En conséquence, la prime de risque (spread de crédit¹³) augmente.

✓ *Le risque de recouvrement*

¹⁰ Cohen, Dictionnaire de gestion, Ed La découverte Paris 1997.

¹¹ H.Jacob & A.Sardi, Management des risques bancaires, Ed AFGES Paris 2001.

¹² Est l'une des agences de notation. Ce point sera détaillé au chapitre suivant.

¹³ Spread = taux de rémunération exigé au débiteur – taux sans risque.

Ce risque représente la probabilité de récupération des créances impayées auprès des débiteurs. Il correspond ainsi à l'incertitude liée au taux de recouvrement/ récupération postérieur à un défaut constaté sur les titres de dette et ce, en entreprenant les procédures judiciaires suite à la faillite de la contrepartie.

Section 2 : La réglementation prudentielle internationale

2.1. Comité de Bâle, présentation et missions

La question de la réglementation bancaire est inséparable de son contexte, celui de la profonde mutation des systèmes financiers.

Selon **Arnaud(2003)**, « *la réglementation bancaire a pour mission de promouvoir la stabilité et la sécurité du système financier en édictant des normes prudentielles s'appliquant aux banques et par des actions de supervision préventives, c'est-à-dire destinées à éviter des crises* ».

Le Comité de Bâle a été institué en 1974 suite à des graves turbulences sur les marchés des changes et dans les secteurs bancaires (faillite de la banque Herstatt) par les gouverneurs des banques centrales du G10¹⁴. Il rassemble de représentants des banques centrales et des autorités de surveillance bancaires des 28 pays suivants : l'Argentine, l'Australie, la Belgique, le Brésil, l'Arabie Saoudite, le Canada, la Chine, la France, l'Allemagne, Hong Kong, l'Inde, l'Indonésie, l'Italie, le Japon, la Corée, le Luxembourg, le Mexique, les Pays-Bas, la Russie, Singapour, l'Afrique du Sud, l'Espagne, la Suède, la Suisse, la Turquie, le Royaume-Uni, les États-Unis et l'Union européenne. Il est hébergé par la Banque des Règlements Internationaux (BRI) à Bâle en Suisse. L'objectif du comité de Bâle peut se résumer ainsi dans la coopération et l'harmonisation internationales en termes de contrôle bancaire. Toutefois, il ne dispose d'aucune autorité, en d'autres termes, ses délibérations n'ont pas la force de la loi.

2.2. L'accord de Bâle I

L'accord de 1988 est centré sur le risque de crédit. Il impose aux banques internationales du G10 un capital réglementaire égal au minimum à 8% du volume des actifs pondérés du risque, le ratio « Cooke ».

$$\text{Ratio Cooke} = \frac{\text{Fonds propres nets}}{\text{Risques pondérés}} \geq 8\%.$$

¹⁴ Le G10 a été élargi et regroupe actuellement 13 pays : Etats-Unis, Canada, France, Allemagne, Royaume-Uni, Italie, Belgique, Pays-Bas, Luxembourg, Japon, Suisse et Espagne.

Chapitre 1 : Le risque de crédit et la réglementation bancaire

L'accord a été amendé afin de s'adopter à l'innovation financière et aux risques qui n'étaient pas couverts dans la configuration initiale. L'amendement de 1996 définissait un tier 3 permettant la couverture des risques de marché.

2.2.1. Les fonds propres nets

Le capital réglementaire = Fonds propres de base + Fonds propres complémentaires et sur complémentaires

Les fonds propres de base comprennent le capital et les réserves, c'est le noyau dur (tier1). Les fonds propres complémentaires sont scindés en deux niveaux upper et lower tier 2 et sont constitués par les produits de dette hybrides assimilables à des quasi-fonds propres (obligations convertibles ou remboursables en actions...). Quant aux fonds propres sur complémentaires (tier3), ils sont composés par les emprunts subordonnés à court terme d'une durée minimale de 2 ans.

2.2.2. Le risque pondéré

Risque pondéré = Σ (Pondérations x Expositions)

Les actifs pondérés du risque comprennent des actifs au bilan et des actifs hors bilan. La pondération des actifs risqués du bilan est répartie en quatre catégories de risque comme l'indique le tableau ci-dessous :

Tableau 1 : Pondération selon type de risque de l'actif sous Bâle I

Source : Arnaud.S. en collaboration avec Zelenko. I., Le risque de crédit, Ed Dunod, Paris, 2003, P.178

Actif	Pondération
Créances sur gouvernements OCDE ¹⁵	0%
Créances sur banques OCDE et gouvernements non OCDE	20%
Créances sur prêts hypothécaires	50%
Autres créances : banques et corporates	100%

¹⁵ Organisation de Coopération et de Développement Economique.

Concernant les actifs hors bilan, ils doivent être convertis en équivalent de crédits, puis insérés dans la catégorie de risque adéquate.

2.2.3. Les insuffisances de l'accord de Bâle I

Certes, l'accord de Bâle I, mis en œuvre en 1992, a marqué un tournant important en termes de réglementation et de stabilité au niveau du système bancaire international. Mais « *sa simplicité qui, dans un premier temps, a constitué sa force, est devenue un handicap par la suite* » affirment **Tiesset et Troussard(2005)**.

En effet, le ratio Cooke ne prend pas en considération de manière pertinente et suffisamment précise le niveau réel du risque de crédit dans la mesure où la définition de catégories de risques est jugée arbitraire. En particulier, le ratio ne tient compte des différences de qualité des emprunteurs puisque quasiment tous les prêts supportent des charges en fonds propres identiques, quelle que soit leur maturité, leur montant ou la solidité financière de leurs bénéficiaires. Ceci signifie une absence de segmentation des risques de crédit selon le degré de séniorité et le niveau de maturité

En outre, « *fixer une borne inférieure du ratio à 8% n'est basé sur aucun fondement théorique ou économique* » (**Roncalli 2004**).

Également, le ratio Cooke tient compte uniquement du risque de crédit en ignorant autres risques fortement présents dans l'activité bancaire tel que le risque opérationnel.

Autres limites peuvent être citées parmi lesquelles :

- ✓ Des mesures trop statiques : Le ratio Cooke ne prend pas en considération ni de l'évolution de risque (changement de la notation), ni du taux de recouvrement (qualité de la garantie).
- ✓ Un développement de l'arbitrage sur le capital, de manière à réduire le niveau de capital réglementaire requis.

La perception sans cesse accrue des faiblesses associées à cet accord, en même temps que l'effort de développement de techniques de gestion des risques du portefeuille de crédit par les institutions financières, rendaient de plus en plus nécessaire la révision de l'accord de 1988 et la naissance d'un nouveau accord qui est celui de Bâle II.

2.3. L'accord de Bâle II

Le comité de Bâle, présidé par William J. Mc Donough, a commencé à élaborer cette réforme en 1998, révisée en avril 2003 et publiée officiellement le 26 juin 2004 puis actualisée en 2005. Cette réforme constitue « le dispositif révisé », selon **Verboomen et al (2011)**.

Chapitre 1 : Le risque de crédit et la réglementation bancaire

Ce nouveau projet d'accord de Bâle est conçu pour permettre de sortir du schéma : « un ratio unique pour tous les engagements ». C'est-à-dire, il vise une couverture plus fine et plus complète des risques en incitant les banques à améliorer leur gestion interne de risque et ce, en éliminant l'approche forfaitaire prise en compte dans l'accord de Bâle I. Ce travail a abouti à l'élaboration du ratio Mc Donough dans la cadre des accords de Bâle II.

L'ambition de la réforme Bâle 2 est de mieux intégrer la complexité accrue de l'activité bancaire tout en laissant aux établissements financiers une certaine autonomie dans le choix des options possibles pour définir le minimum de fonds propres nécessaires à la couverture des pertes potentielles (Nouy, 2003)

Par rapport à l'accord de 1988, le comité de Bâle définit avec clarté un ensemble de déclarations d'intention ¹⁶à savoir :

- ✓ Développer une approche des risques plus globale,
- ✓ Développer une approche qui tienne mieux compte du niveau réel de risque propre à chaque institution,
- ✓ Centrer l'approche sur les banques internationales.

2.3.1. Les trois piliers de Bâle II

Bâle II repose sur trois piliers complémentaires :



Figure 3 : Les trois piliers de Bâle II

Source : www.finmarkets.com

a. Le premier pilier : Exigence en fonds propres

En matière de capital minimum, le niveau de 8% correspond toujours au plancher de sécurité. Mais ce pourcentage est déterminé en incluant les risques de crédit, les risques de marché et les risques opérationnels.

¹⁶ Arnaud Servigny, Le risque de crédit, Ed Dunod, P.183.

Chapitre 1 : Le risque de crédit et la réglementation bancaire

Le nouveau ratio Mc Donough prend la forme suivante :

$$\text{Ratio Mc Donough} = \frac{\text{Fonds propres nets}}{\text{Risques de crédit} + \text{Risques de marché} + \text{Risques opérationnel}} \geq 8\%.$$

La méthode de pondération des risques de crédit est profondément **modifiée**
 La méthode de pondération des risques de marché est **inchangée**.
 La méthode de Pondération des risques opérationnels est **créée**.

En outre, le pilier 1 de Bâle II a présenté les différentes méthodes d'estimation du risque :

Tableau 3: Les différentes approches d'évaluation des différents risques selon Bâle II

Nature de risque	Méthodes présentées
Le risque de crédit	Approche standard
	Approche IRB de base (Foundation Internal Rating Based)
	Approche IRB avancée (Advanced Internal Rating Based)
Le risque opérationnel	Approche de base (% du PNB global).
	Approche standard (% du PNB différencié suivant les lignes de métiers).
	Approche AMA (Advanced Measurement Approaches) basée sur un modèle interne).
Le risque de marché	Modèle forfaitaire calibré par le régulateur.
	Modèle interne de VAR.

Dans ce qui suit, nous allons nous focaliser sur les approches de mesures de risque de crédit.

b. Le deuxième pilier : Le processus de surveillance prudentielle

Il s'agit du processus de surveillance prudentielle qui doit être effectué par chaque banque au niveau interne afin de permettre de maintenir un capital réglementaire à un niveau acceptable, quelle que soit l'évolution de la qualité des engagements. Le pilier 2 institue ainsi le principe d'un dialogue structuré entre établissements et superviseurs.

De manière synthétique, quatre grands principes de supervision sont définis :

- ✓ Les banques doivent disposer d'un système de mesure interne afin de déterminer leur niveau global de capital réglementaire, en relation avec leur profil de risque. En outre, elles doivent avoir une stratégie permettant un maintien de leur niveau de capital.

- ✓ Les régulateurs devront contrôler, analyser et évaluer la mise en place de ce système et la stratégie de maintien pour assurer le respect du niveau de capital réglementaire.
- ✓ Les régulateurs sont en droit d'exiger aux banques un niveau de fonds propres supérieur au niveau des exigences minimums, qui leur paraît nécessaire.¹⁷
- ✓ Les régulateurs doivent veiller que le niveau des fonds propres de la banque ne tombe pas en dessous du capital réglementaire. A défaut, ils doivent demander des mesures rapides de restauration du capital.

c. Le troisième pilier : La discipline de marché

Le pilier 3 vise à instaurer des règles de transparence financière en améliorant la communication d'informations à destination du grand public sur les actifs, les risques et leur gestion. Le comité a convenu de cinq principes directeurs pour cette communication afin de garantir la transparence et la grande qualité nécessaires qui aident les utilisateurs à mieux comprendre et comparer l'activité et les risques des banques. Ces principes sont autour de la qualité des informations qui doivent être :

- **Clares** : Compréhensibles par les parties prenantes (investisseurs, analystes, client...) et communiquées par des moyens accessibles.
- **Exhaustives** : Elles doivent décrire les principales activités et tous les grands risques de la banque. Ceci peut être consolidé par des données et des informations sous-jacentes pertinentes.
- **Pertinentes** : Les informations doivent souligner les principaux risques, actuels et nouveaux, de la banque ainsi que la manière dont ils sont gérés.
- **Cohérentes dans le temps** : Permettent aux parties prenantes de cerner l'évolution du profil de risque de la banque selon tous les grands aspects de ses activités.
- **Comparables d'une banque à l'autre** : Le niveau de détail et le format de présentation des informations doivent permettre aux parties prenantes d'établir des comparaisons utiles entre les banques et les juridictions(en termes d'activités, de mesures prudentielles, de risques et de gestion de ces derniers).

¹⁷ Pour les activités dont la dépendance au cycle économique est marquée, le comité de Bâle invite à prévoir un coussin de complémentaire.

2.3.2. Les limites de Bâle II

La crise financière de 2007 (Subprime) a mis en évidence les lacunes de Bâle II. Cet événement nous a montré que la gestion des risques bancaires est loin d'être résolue. Plusieurs critiques ont été constatées à savoir :

- ✓ Le dispositif de Bâle II est jugé trop compliqué dans la mesure où beaucoup de banques sont incapables de mettre en place les techniques avancées de mesure du risque. D'autres, s'interrogent sur le coût de la mise en œuvre de la supervision. A cet égard, on peut douter de la capacité des pays en voie de développement à contribuer à la mise en place effective et rapide du nouvel accord.
- ✓ Le caractère défavorable pour les pays émergents de Bâle II : fragilité de ces économies qui a accrue en raison du caractère pro-cyclique¹⁸ du nouvel accord. En effet, les exigences en fonds propres sur-réagissent à la hausse en cas de ralentissement de la croissance ou de récession de l'activité ce qui engendre un risque du rationnement de crédit ce qu'on l'appelle « credit crunch », à la baisse en période d'euphorie financière. Ces événements accentuent l'amplitude du mouvement conjoncturel et exercent une influence déstabilisante sur l'activité bancaire.
- ✓ La défaillance du contrôle interne et de la gouvernance des établissements bancaires comme le cas de la défaillance de la Société Générale en 2008.
- ✓ Les risques liés aux instruments financiers complexes n'ont pas été suffisamment pris en compte dans les modalités de calcul du ratio. Dans ce cas, les banques n'ont pas été en mesure d'apprécier correctement les risques qu'elles prenaient en initiant des opérations de titrisation, de retitrisation ou de dérivés de crédit.

2.4. Bâle III, une suite nécessaire

Suite à la crise 2007-2009, le Comité de Bâle a décidé qu'une refonte en profondeur des accords de Bâle II était nécessaire. Le conseil de stabilité financière et le G20¹⁹ à Séoul ont contribué à la mise au point de nouvelles mesures de stabilité pour le système bancaire mondial.

Une première version des propositions de Bâle III a été publiée en décembre 2009. Suite aux commentaires des banques, à une étude d'impact quantitative et à plusieurs sommets

¹⁸ La pro cyclicité est définie par la Fédération Bancaire Française comme « la variabilité excessive d'un facteur aux fluctuations du cycle économique ».

¹⁹ Le groupe des vingt est un groupe composé de dix-neuf pays et de l'Union Européenne dont les ministres des Finances et les gouverneurs des banques centrales se réunissent annuellement.

internationaux, la version finale a été publiée en décembre 2010²⁰ et il n'a été mis en vigueur qu'au cours de l'année 2011 (Bousslama et al. 2009). Le nouvel accord vise d'une part, à renforcer globalement en quantité et en qualité le capital prudentiel mobilisé par les banques pour faire face à des situations adverses et, d'autre part, à garantir leur liquidité en cas de tensions monétaires.

Il comporte six parties :

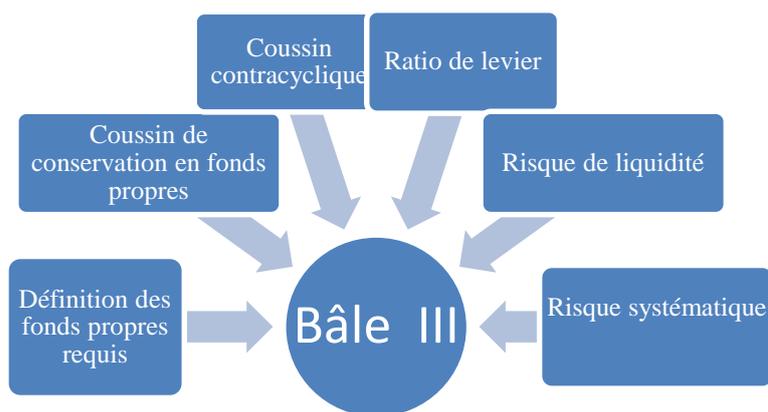


Figure 4 Les six parties de l'accord de Bâle III

Bâle III doit être mis en place entre 2013 et 2019.

Les principales modifications par rapport à Bâle II sont les suivantes :

2.4.1. Instauration de ratio de liquidité

Bâle III instaure un ensemble de dispositions pour faire en sorte que les banques disposent de suffisamment de liquidités, ou de quasi-liquidités pour surmonter un choc financier très sévère à court terme et moins sévère à moyen et long termes. Ces risques proviennent de la tendance des banques à financer des actifs à long terme par des ressources à court terme.

Cet accord a introduit des règles via deux ratios de liquidité :

- **Ratio de liquidité à court terme (Liquidity Coverage Ratio-LCR)**

Le LCR porte sur la capacité d'une banque à survivre à une période de perte de liquidité sur 30 jours et se définit comme suit :

$$\frac{\text{Actifs liquides de hautes qualité}}{\text{Sorties nettes de trésorerie sur 30 jours}} \geq 100\%$$

²⁰ Basel Committee for Bank Supervision, Basel III: A Global Regulatory Framework for More Resilient Bank and Banking Systems, décembre 2010; Basel III: International Framework for Liquidity Risk Measurement Standards and Monitoring, décembre 2010.

Les liquidités de haute qualité comprennent les espèces encaisse, les réserves des banques centrales, les titres émis par des pays souverains, banques centrales, institutions financières et les titres de créances émises par les sociétés industrielles et commerciales. Concernant les sorties de liquidité, ils sont relativement détaillées et on constate une différenciation par type de clients (les clients de détail étant considérés comme moins volatils par le régulateur et les dépôts des institutions financières étant les plus volatils puisqu'ils sont pondérés à 100%).

- **Ratio structurel de liquidité à long terme (Net Stable Funding Ratio-NSFR)**

Le NSFR s'inscrit dans une optique de moyen et long terme. L'objectif affiché par le Comité de Bâle est d'éviter les impasses de maturités entre les actifs et les passifs d'une banque. Le ratio est présenté ainsi :

$$\frac{\text{Financement stable disponible}}{\text{Financement stable requis}} \geq 100\%$$

La conséquence de l'application du NSFR est que :

- Les prêts aux entreprises à plus d'un an devront être couverts par un montant équivalent de ressources à plus d'un an ;
- Les prêts aux entreprises à moins d'un an devront être couverts à hauteur de 50% par des ressources à plus d'un an ;
- Les prêts immobiliers aux ménages devront être couverts à hauteur de 65% par des ressources à plus d'un an.

2.4.2. Maitrise de l'effet de levier

Bâle III vise à instaurer un ratio de levier maximal (égal au rapport entre les fonds propres et le total bilan), qui vise à plafonner l'accumulation de l'endettement dans le secteur bancaire à un niveau mondial. Un ratio représentant 3% du Tier 1 est mis en essai du 1^{er} janvier 2013 jusqu'à 1^{er} janvier 2017 pour opérer des ajustements sur la définition et le calibrage du ratio. A partir du janvier 2018, l'application de ce ratio est devenue obligatoire.

2.4.3. Amélioration de la qualité des fonds propres

La définition du capital réglementaire, a été précisée par le Comité de Bâle du 26 juillet 2010 : Le tiers 3 de Bâle II a disparu et le ratio « core tier1 », formellement introduit. Ce dernier est constitué uniquement du capital social et des réserves non distribuées, il exclut toutes les formes de capitaux hybrides (obligations convertibles, obligations perpétuelles...). Le core tier 1 s'élevait au minimum à 2% des capitaux propres dans Bâle II et il passera progressivement à 4.5% dans Bâle III.

Les fonds propres de base ou le ratio « tier 1 » d'un minimum de 4% dans Bâle II, passera à 4.5% en 2013 puis à 6% progressivement jusqu'au 1^{er} janvier 2015. A ce ratio, se rajoutera un coussin de conservation de 2.5%. En conséquence, les fonds propres de base totaux (tier 1+fonds propres de base additionnels²¹) et les fonds propres totaux (tier 1 et tier 2) doivent être au moins égaux à 8.5% et 10.5% des actifs pondérés au risque respectivement. Ce dernier doit être constitué par les banques en période économique normale pour pouvoir absorber des pertes en cas de difficultés financières. En cas d'utilisation complète ou partielle de ce matelas de sécurité, les banques sont tenues de réduire leurs dividendes jusqu'à ce qu'elles soient adéquatement recapitalisées.

2.4.4. Réduction de la pro cyclicité

Bâle III exige également la détention d'un coussin contra cyclique, similaire au coussin de conservation, néanmoins la mise en place est laissée à la discrétion des autorités nationales. Ce dernier vise la protection contre la cyclicité des résultats bancaires. Il doit représenter de 0% à 2.5% des actifs pondérés au risque et se composer de fonds propres de base (tier1).

2.4.5. Traitement du risque systématique

Le G20 de novembre 2010 a validé la proposition de créer deux catégories d'institutions financières nationales et internationales d'importance systémique: les SIFIs (Systemically Important Financial Institutions) et les Global SIFIs. En cas de faillite, ces dernières présentent un risque majeur de perturbation pour l'ensemble du système financier. Et par conséquent, elles seraient soumises à une régulation plus renforcée.

L'accord de Bâle III présenté en décembre 2017 complète les réformes prudentielles internationales et remède aux faiblesses mises au jour par la crise financière mondiale. Selon le mandat donné par le G20, la finalisation de Bâle III a pour objectif « une harmonisation mondiale des méthodes de calcul du risque, sans augmentation significative des fonds propres ni discrimination entre modèles bancaires ».

Cette finalisation comporte principalement trois dispositions et elle n'est pas encore mise en place :

- La révision des exigences en matière de **risque de marché** : Ce volet de la réforme vise à pondérer certaines expositions de marché prises par les banques. Initialement, l'entrée en vigueur du FRTB (Fundamental Review of the Trading Book, ou revue fondamentale du portefeuille de marché) était prévue pour 2019, mais elle est

²¹ Incluent des composantes comme les actions privilégiées à dividende non cumulatif.

repoussée à 2022 pour permettre de finaliser des travaux complémentaires et faire de sorte que l'ensemble des pays mettront en œuvre ces règles.

- La révision des exigences en matière de **risque opérationnel** : Une nouvelle approche standard, plus développée et plus sensible aux risques doit être retenue par les banques avec un niveau d'exigences revu afin de s'adapter aux évolutions observées de ce risque. Les modèles internes ne sont plus autorisés pour ce risque, car ils se sont avérés trop peu robustes.
- Amélioration du traitement du risque : la principale modification apportée à l'approche IRB pour la gestion du risque de crédit est la suppression de la possibilité d'utiliser l'approche avancée (A-IRB) pour les expositions aux banques et autres établissements financiers (Les expositions aux actions ne pourront faire l'objet d'aucune approche IRB). Toutefois, l'utilisation de cette approche est **introduite** pour les entreprises appartenant à des groupes dont le chiffre d'affaires consolidé dépasse 500 millions EUR.
- Introduction d'un plancher en capital « **output floor** » plus solide et plus sensible au risque : Il prévoit qu'à partir de 2027, le résultat obtenu par la méthode interne ne pourra pas être inférieur à 72.5% de ce qu'aurait donné la méthode standard. Le plancher sera introduit à partir de 2022 et réévalué chaque année.

Section 3 : La réglementation prudentielle nationale

En s'inspirant de la réglementation internationale, l'autorité de contrôle représentée par la Banque Centrale de Tunisie(BCT) a mis en place une réglementation nationale afin d'assurer la sécurité et la pérennité du système bancaire tunisien.

En effet, elle assure l'organisation du secteur via la publication des circulaires. Parmi les principaux textes réglementaires en matière de gestion de risque bancaire publiés, on distingue la circulaire **n°91-24** du 17 décembre 1991, modifiée par la circulaire aux établissements de crédits n°2012-09 du 29 juin 2012 et qui traite la division, la couverture des risques et le suivi des engagements. Également, on trouve la circulaire **n°2016-06** relative au système de notation des contreparties qui a été publiée le 11 octobre 2016.

3.1. La division et la couverture du risque

L'article 1 de la circulaire **n°91-24** stipule que le montant total des risques encourus ne doit pas excéder :

- 3 fois les fonds propres nets de la banque, pour les bénéficiaires dont le risque encouru de chacun est 5% ou plus des fonds propres nets.
- 1.5 fois les fonds propres nets de la banque, pour les bénéficiaires dont le risque encouru de chacun est 15% ou plus des fonds propres nets.

D'après l'**article 2** de la même circulaire, les risques encourus sur un même bénéficiaire ²²ne doivent pas dépasser 25% des fonds propres nets de la banque.

Finalement, l'article 3 renforce la division du risque par le fait que le montant total des risques encourus sur les personnes ayant des liens avec l'établissement de crédit, au sens de l'**article 23 de la loi n°2001-65 du 10 juillet 2001** relative aux établissements de crédit, ne doit pas dépasser une seule fois les fonds propres nets.

Concernant la couverture du risque, les banques doivent respecter en permanence un ratio de solvabilité de 10%, calculé par le rapport entre les fonds propres nets (fonds propres nets de base + fonds propres complémentaires) et les risques encourus (montant des risques de crédit pondérés+ montant des risques opérationnels). Le même article porte aussi sur les fonds propres nets de base qui ne doivent pas être inférieurs en permanence à 7% de la somme des risques encourus.

3.2. La classification des actifs et la constitution des provisions

Selon l'**article 8** de la dite circulaire, les banques sont tenues de classer leurs actifs indépendamment de leurs formes (au bilan ou hors bilan, en dinars ou en devises). Ces derniers seront par conséquent classés en actifs courants et actifs classés.

Les actifs détenus sur la BCT ou sur l'Etat ne font pas l'objet de cette classification.

- **Actifs courants** : Ce sont les actifs dont le recouvrement intégral dans les délais paraît assuré et qui sont détenus sur des entreprises dont la situation financière est équilibrée.
- **Actifs classés** : Ce sont les actifs dont le remboursement est incertain. L'aléa se matérialise par un risque d'impayé dont le degré de gravité est classé en quatre classes :
 - *Actif nécessitant un suivi particulier* : Ces actifs présentent un impayé < 3 mois dont le recouvrement intégral dans les délais est assuré, mais détenus sur des entreprises ayant un secteur d'activité qui connaît des difficultés ou ayant une situation financière qui se dégrade.

²² Sont considérés comme "même bénéficiaire" les emprunteurs affiliés à un même groupe.

- *Actifs incertains* : Ce sont les actifs dont l'impayé couvre une période comprise entre [3mois ; 6mois] et le recouvrement intégral dans les délais est incertain. Ils sont détenus généralement sur des entreprises qui connaissent des difficultés financières. Les provisions à constituer sont au moins égales à 20% pour les actifs de cette classe.
- *Actifs préoccupant* : sont concernés par cette classe les actifs détenus sur des entreprises qui présentent un retard de paiement supérieur 6 mois sans excéder une année. Les banques doivent constituer des provisions d'au moins 50% pour les actifs appartenant à cette classe.
- *Actifs compromis* : Ce sont des actifs détenus sur des entreprises qui présentent un retard de paiement supérieur à une année. L'ensemble des créances relatives à ces actifs nécessitent un provisionnement intégral (100%).

Il est à noter que le montant de ces provisions doit être affecté spécifiquement à tout actif classé ≥ 50 milles dinars et ceci compte tenu des garanties reçues par l'Etat, des assurances, des Etablissements de crédit ainsi que des garanties sous forme de dépôts ou d'actifs financiers susceptibles d'être liquidés sans que leur valeur soit affectée.

3.3. La mise en place d'un système de notation interne des contreparties

Dix ans après la publication de la **circulaire n°2006-19** relative au contrôle interne qui a chargé le comité de risque de crédit de procéder à l'élaboration d'un système de notation et d'en assurer la révision permanente, la BCT a publié le 11 octobre, une nouvelle **circulaire 2016-06**.

Cette dernière énonce des dispositions qui obligent les banques à mettre en place un système de notation lors de l'octroi de crédits. Elle permet d'édicter certains principes inspirés du cadre bâlois relatifs à la conception, à la structure, à la mise à jour, à l'utilisation et au contrôle du système de notation qui constituent des exigences minimales à respecter par les établissements afin de pouvoir attribuer une note à chaque contrepartie.

Cette notation doit jouer un rôle principal dans le processus d'octroi des crédits, la politique de tarification appliquée aux clients, la politique de gestion des risques et dans l'allocation interne des fonds propres en préparation de l'adoption de l'approche basée sur les notations internes de l'accord de Bâle II.

Pour ce faire, les banques ont adressé avant décembre 2016 une feuille de route pour l'implémentation du système de notation des contreparties. Ensuite, elles sont tenues de respecter les exigences minimales prévues par la dite circulaire avant décembre 2017. Ces

exigences s'appliquent pour la notation de contreparties souveraines, banques, établissements financiers et entreprises. Elles concernent :

- **Les paramètres et la structure du système de notation interne :** Le système de notation interne doit comprendre deux paramètres distincts, le premier sera porter sur l'appréciation du risque de la contrepartie et le deuxième va concerner la qualité des garanties reçues en couverture. Il est à noter que l'échelle de notation comporte au moins sept notes pour les contreparties qui ne sont pas en défaut et une note pour les contreparties en défaut. En outre, un tel système ne doit pas être figé, sa revue doit être faite au moins une fois par an.
- **La documentation du système de notation :** Les données sur les contreparties doivent être collectées et suffisamment détaillées (les dates d'attribution des notes, la méthode et les principales données utilisées pour établir les notations, l'identité de la personne qui a attribué les notes....).
- **La gouvernance et le contrôle système de notation :** La structure chargée de la gestion du risque de crédit au sein des établissements est responsable de la conception ou la sélection du système de notation, de sa mise en œuvre, de sa surveillance et de son efficacité. Le conseil doit le valider et la structure d'audit interne est tenue de le revoir, au moins une fois par an, contrôler son fonctionnement, et de s'assurer du respect des exigences minimales définies dans la présente circulaire.

La mise en place d'un modèle de notation interne repose initialement sur la définition de certaines variables clés, appelés également les paramètres de risques. Ce qui fera l'objet de la prochaine section.

Section 4 : Les paramètres bâlois du risque de crédit

Dans ce paragraphe, nous allons définir les paramètres du risque tels que la probabilité de défaut (PD)²³, la perte en cas de défaut (LGD), l'exposition au moment de défaut(EAD) et la perte attendue (PA).

4.1. La probabilité de défaut

Exprimée en pourcentage, elle correspond à la probabilité qu'une contrepartie soit défailante sur un horizon.

4.2. La perte en cas de défaut (Loss Given Default : LGD)

Exprimée en pourcentage, elle correspond au taux de perte constaté en cas de défaillance.

²³ Notion développée à la page 6. Nous nous contenterons ici d'un simple rappel.

La LGD peut être estimée à partir des taux moyens de recouvrement selon la formule :

$$\text{LGD} = 1 - \text{taux de recouvrement}$$

L'analyse de la LGD permet d'orienter la banque vers le renforcement de la politique de prévention contre les pertes par la mise en place des couvertures appropriées tel que le recours à des nouvelles sûretés, garanties ou couvertures. Toutefois, ce paramètre constitue une source d'incertitude pour la banque dans la mesure où il demande une information riche et complexe comme la durée des procédures judiciaires, la valeur réelle des garanties,...

4.3. Exposition au moment du défaut (Exposure at Default : EAD)

Elle correspond à l'encours de crédit et autres engagements (y compris les engagements hors bilan) en cas de défaillance de la contrepartie. En d'autres termes, l'EAD correspond au risque brut anticipé sur l'engagement à la suite du défaut du débiteur.

4.4. La maturité

Dans l'approche standard, le comité de Bâle a fixé l'échéance effective à 2 ans et demi. Néanmoins, l'approche avancée introduit la notion de l'échéance effective qui se définit comme suit :

$$EE = \frac{\sum t \times Ft}{\sum Ft}$$

Où F_t indique les flux de trésorerie (paiements du principal, des intérêts et des commissions) remboursables par contrat pendant la période t .

4.5. La perte attendue (Expected Loss)

La banque est en mesure de prévoir le niveau moyen de perte future auquel elle peut raisonnablement s'attendre, ce niveau est la perte attendue. Elle le considère comme le coût moyen du risque de crédit lié à son activité de financement qui sera comblé par le capital réglementaire.

L'Expected Loss (espérance mathématique de la distribution de perte) peut être simplifiée comme suit :

$$\text{EL} = \text{EAD} * \text{LGD} * \text{PD}$$

Conclusion

Plusieurs types de risques peuvent affecter la survie d'une banque. Parmi ces risques on trouve le risque de marché, le risque opérationnel, le risque de crédit.... En conséquence, la

Chapitre 1 : Le risque de crédit et la réglementation bancaire

mesure et la gestion de ces risques demeurent la préoccupation majeure de chaque banquier conscient des enjeux économiques et financiers actuel.

Afin d'organiser et déterminer un cadre bien précis dans lequel les marchés financiers peuvent opérer et évoluer, un cadre réglementaire prudentielle à l'échelle mondiale est instauré.

Le premier fruit de cette réglementation fût le ratio Cooke mais avec l'évolution perpétuelle de l'activité bancaire, le comité de Bâle essaye dans chaque accord de remédier aux lacunes du précédent. C'est ainsi que sont venus les accords de Bâle II puis Bâle III dont la détermination des exigences minimales en fonds propre reste le noyau de toute réglementation, puisque grâce à eux, une banque peut surmonter ces différents risques.

Parallèlement, la Banque Centrale de Tunisie vise à s'aligner progressivement avec les recommandations édictées par les autorités de régulation internationale afin d'assurer la stabilité financière dans le système bancaire tunisien et protéger en premier lieu les déposants par l'allocation optimale des fonds propres. A l'horizon de 2020, elle ambitionne que toutes les banques de place utiliseront la notation des contreparties pour modéliser leur risque de crédit et à ajuster leurs fonds propres en fonction du risque encouru. Cela nécessite le recours aux modèles spécifiques à la quantification de ce risque. Ce qui fera l'objet dans le chapitre qui suit.

Chapitre 2 : Méthodologie de notation

Chapitre 2 : Méthodologie de notation

Introduction

Apparue au début du 20^{ème} siècle aux États-Unis afin de faciliter la gestion du risque de crédit, la notation financière occupe aujourd'hui une place centrale dans le fonctionnement des marchés financiers. En effet, elle fait partie des approches traditionnelles de mesure du risque selon **Saunders(1999)**. En outre, **Rouges(2003)** affirme qu'elle est un instrument de gestion des risques de premier ordre.

D'autre part, le **règlement 1060/2009 du Parlement Européen et du Conseil** du 16 septembre 2009 sur les agences de notation de crédit, définit la notation comme étant « *un avis émis par application d'un système de classification bien défini et bien établi prévoyant différentes catégories de notation, concernant la qualité de crédit d'une entité, d'une dette ou obligation financière, d'un titre de créance, d'actions privilégiées ou autres instruments financiers, ou d'un émetteur d'une telle dette ou obligation financière, d'un tel titre de créance, de telles actions privilégiées ou d'un tel instrument financier* ».

La notation permet de mesurer ainsi le risque lié à un client ou un crédit. Suite à cette mesure, la banque détermine la décision de poursuite de la relation bancaire, la tarification des services, l'intensité de la surveillance....Elle est généralement attribuée lors de l'entrée en relation du client avec la banque et révisée à des intervalles réguliers notamment lors de renouvellements d'octrois ou suite à des événements susceptibles de modifier significativement le risque d'un client.

Face aux réglementations prudentielles internationale (nouvelles réformes de Bâle III) et nationale (circulaire 2016-06), les banques sont tenues d'attribuer une grande importance à la maîtrise du risque de crédit. Pour ce faire, elles doivent être dotées d'un mécanisme pour l'évaluer et pouvoir par la suite le minimiser. Une telle évaluation peut être effectuée à l'aide de :

- ✓ Approche par les notations externes,
- ✓ Approches par la notation interne (modèles statistiques).

Dans ce chapitre, nous allons dans une première section présenter l'approche par les notations externes et dans une seconde étape nous allons nous focaliser sur les principales méthodes statistiques utilisées dans le « Credit Scoring ».

Section 1 : Approche par les notations externes

1.1. Définitions

Selon **Zak(2002)**, « *la notation est l'évaluation de la solvabilité des entités par les agences de notation qui attribuent une note selon une échelle prédéfinie. La note exprime la probabilité que l'entité notée respecte ses engagements* ».

En outre, **Weber et Krahnén (2001)** ont indiqué que « *le but d'une notation de crédit est de classer les demandeurs de crédit en fonction de leur qualité, c'est-à-dire de leur probabilité de défaut sur un horizon de temps donné* ».

Aujourd'hui, la notation de crédit reflète l'opinion d'une agence à partir d'une analyse financière et opérationnelle. Elle se fonde sur l'analyse des éléments quantitatifs et qualitatifs relatifs à la position actuelle et prévisible de l'entreprise (**Ferri et Liu 2005**).

La notation externe facilite ainsi l'évaluation de la solvabilité d'un émetteur ou d'un emprunteur dans le but d'affiner la modélisation de son risque de défaut. En conséquence, cette note de crédit est considérée comme un outil synthétique puissant ayant un impact décisif sur les décisions au sein des marchés financiers. Elle est établie par les agences de notation.

1.2. Les agences de notation

Zelenko et al. (2003) définissent ces agences de notation comme étant « *des entreprises privées chargées d'élaborer des ratings pour les différents actifs financiers sur la base d'opinions indépendantes, objectives, crédibles et transparentes* ».

Egalement, le **règlement 1060/2009 du Parlement Européen et du Conseil** sur les agences de notation de crédit les définit comme étant « *une personne morale dont l'activité inclut l'émission de notations de crédit à titre professionnel* ». L'activité principale de ces agences de notation selon **Cantor et Packer (1994)** consiste à « *donner des avis sur la qualité du crédit, sur la capacité d'honorer les obligations financières d'un émetteur ou d'un instrument financier* ».

Les premiers ratings, qui ne couvrent au début que les entreprises de chemins de fer aux États-Unis, apparaissent en 1909 dans le manuel « *Moody 'Analyses of railroad investments* ». La notation devient alors un réel outil d'information pour les investisseurs en réponse aux émissions obligataires importantes de ce secteur d'activité. Puis en 1941 il y a eu la création de Standard & Poor's (S&P) suite à la fusion de Poor's et Standard statistics.

La plus importante agence de notation européenne IBCA (International Banking and Credit Analysis) a été créée en 1978 comme une institution spécialisée dans les banques dans les îles britanniques plus connue depuis 2000 comme Fitch rating.

En 1975, on recensait sept agences globales²⁴. Aujourd'hui, la notation des actifs financiers est une véritable industrie oligopolistique partagée essentiellement entre trois agences :

- **Moody's Corporation** est une entreprise cotée dont l'actionnaire majoritaire est Berkshire Hathaway Inc. (12.9%) et dirigée par le milliardaire Warren Buffett.
- **Fitch rating** appartient à hauteur de 60% par le groupe français FIMALAC et à hauteur de 40% par le groupe américain Hearst Corporation.
- **Standard & Poor's** est détenue à 100% par la maison d'édition McGraw-Hill.

L'oligopole de ces trois grandes agences « détient 97 % de parts de marché dans le secteur des produits structurés et 99 % dans celui de la notation des titres du secteur public et parapublic » (Autorité des Marchés Financier, 2010).

1.3. Processus de rating et échelle de notation

Le processus de notation est complexe et propre à chaque agence de notation, il se base sur plusieurs critères, qui se décomposent eux-mêmes en sous critères, qu'il convient d'estimer et de pondérer à l'aide d'un modèle mathématique de notation. Chaque agence possède son propre système de notation, toutefois on peut voir les grandes lignes : ²⁵

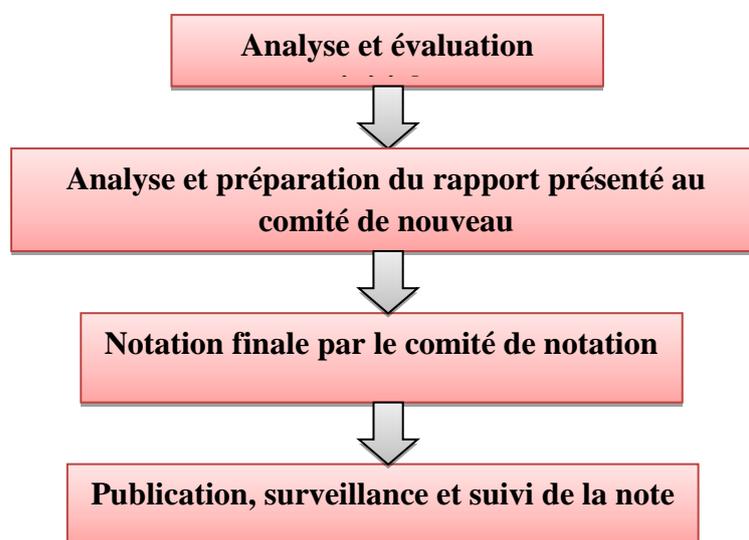


Figure 5. Processus de notation par les agences de rating

²⁴ Gurnot.N, Effets de la notation financière sur les prix des obligations, Master comptabilité contrôle audit, Université Paris Dauphine, 2013, P.6.

²⁵ Ben Ayed.M, le risque de crédit des obligations : analyse de la migration des notes et des effets de contagion, thèse de doctorat en sciences de gestion, Université Cergy Pontoise, 2013, P.17.

Chapitre 2 : Méthodologie de notation

Comme on a évoqué précédemment, chaque agence a sa propre méthode de notation et sa propre échelle de notation. Ainsi, pour un même emprunteur, noté par plusieurs agences peut obtenir des notes différentes. Ces agences distinguent deux types d'échelles, une échelle pour les notations à long terme et une échelle de notation à court terme.

Le tableau qui suit représente une correspondance entre les différentes notes à long terme attribuées par les trois principales agences.

Tableau 2 Correspondance entre les notes attribuées par les agences de notations
Source : Autorité des Marchés Financiers, juillet 2013, France.

Fitch	Moody's	Standard & Poor's	
AAA	Aaa	AAA	Grade Investissement
AA+	Aa1	AA+	
AA	Aa2	AA	
AA-	Aa3	AA-	
A+	A1	A+	
A	A2	A	
A-	A3	A-	
BBB+	Baa1	BBB+	
BBB	Baa2	BBB	
BBB-	Baa3	BBB-	
BB+	Ba1	BB+	Grade Spéculatif
BB	Ba2	BB	
BB-	Ba3	BB-	
B+	B1	B+	
B	B2	B	
B-	B3	B-	
CCC+	Caa1	CCC+	
CCC	Caa2	CCC	
CCC-	Caa3	CCC-	
CC	Ca	CC	
CD	C	CD	

Chapitre 2 : Méthodologie de notation

La catégorie d'investissement dont les notes varient de AAA à BBB - est la moins risquée en termes des perspectives futures par rapport à la catégorie de spéculation. Cette dernière dont les notes varient de BB+ à CD regroupe des émissions dont le remboursement du capital et d'intérêt est plus aléatoire puisqu'il dépende des conditions externes à l'entité notée.

L'échelle de notation de Fitch ne diffère pas de celle de S&P. Elle est affinée en ajoutant des signes + ou - signalant que l'émetteur se trouve plutôt dans le haut ou dans le bas de la classe attribuée. D'une manière semblable, la notation présentée par Moody's attribue des indices numériques (1,2 et 3) à ses notes de Aa à Caa : L'indice 1 indique que l'obligation est située dans le premier cran de la note principale, chose qui stipule la meilleure qualité de crédit au sein de cette classe, l'indice 2 indique que l'obligation est située dans le second cran de sa note principale, chose qui stipule une qualité de crédit moyenne au sein de cette classe. Au final, l'indice 3 indique que l'obligation est située dans le troisième cran de sa note principale, chose qui stipule une qualité de crédit réduite au sein de cette classe.

1.4. Notation externe : Méthode de calcul du risque de crédit selon Bâle II

Étant donné que l'ancienne pondération (ratio du Cooke) n'avait pas de fondements théoriques, la nouvelle pondération a été basée sur la notation externe (agences de notation). En effet, le système de référence utilisé dans l'accord de Bâle II est celui de Standard & Poor's.

Tableau 3 : Pondérations proposées par le Comité de Bâle par nature de contrepartie et par note.
Source : Le comité de Bâle, Banque des Règlements Internationaux, texte de janvier 2001.

Notation	AAA à AA-	A+ à A-	BBB+ à BBB-	BB+ à BB-	B+ à B-	< B-	Non noté
Souverain	0%	20%	50%	100%	100%	150%	100%
Banques²⁶	20%	50%	50%	100%	100%	150%	50 %
Banque – actif à CT	20%	20%	20%	50%	50%	150%	20%
Entreprise	20%	50%	100%	100%	150%	150%	100%

Concernant les créances aux particuliers (Retail Risk), ne sont pas notées mais elles sont soumises à une pondération de 75% (s'il s'agit d'un prêt immobilier, la pondération demeure 35%). La lecture du tableau est la suivante : les risques des crédits à la catégorie de souverain

²⁶ La pondération est en fonction du risque souverain et proposée par la banque.

des pays les mieux notés (de AAA à AA-) n'auraient pas à être provisionnés tandis que ceux des pays les moins bien notés (sous B-) devraient être provisionnés.

On peut conclure que cette approche est plus sensible au risque dans la mesure où les pondérations dépendent de la nature de l'emprunteur (dépassent dans certains cas 100%) mais aussi de la qualité de risque de ce dernier (sa notation).

Le calcul à partir du bilan s'effectue en appliquant aux différentes natures d'engagements un des coefficients de pondération cités précédemment. Le montant du risque pondéré des engagements s'obtient par la formule suivante :

$$\text{Risque pondéré} = \text{Engagements au bilan} \times \text{Taux de pondération du risque}$$

La procédure est la même pour les éléments hors bilan²⁷, à la différence que ces éléments seront auparavant pondérés par des coefficients de conversion en fonction des catégories hors bilan.

Il faut noter que le comité de Bâle a laissé le choix des méthodes aux banques : l'approche standard est relativement simple à mettre en place, a vocation à être utilisée par les banques n'ayant pas les moyens techniques et humains d'utiliser les méthodes fondées sur les notations internes. Cette méthode accorde un rôle clef aux agences de notation, ce qui constitue l'une des principales critiques adossées par les professionnels. Une autre critique est avancée concernant les erreurs commises par ces agences de notation.

Dans ce qui suit, on va se focaliser sur la deuxième approche présentée par le comité de Bâle : la méthode de notation interne.

Section 2 : Approche par les notations internes

Contrairement à l'approche standard, cette approche est basée sur l'estimation de la probabilité de défaut. Il s'agit donc d'une méthode plus sensible au risque selon **Roncalli(2004)**.

Dans cette approche, « *les banques pourront utiliser leurs estimations internes sur la solvabilité de leurs emprunteurs pour évaluer le risque de crédit inhérent à leur portefeuille, à condition qu'elles respectent des critères stricts en matière de méthodologie et de communication financière* ». ²⁸

²⁷ Parmi les éléments hors bilan, on note : Obligations cautionnées, Crédits documentaires ou les marchandises servent de garantie, Cautions, avals et autres garanties accordés à la clientèle ou à des établissements de crédit.

²⁸ Eric- PAGET BLANC., Le rôle informationnel des ratios de fonds propres des banques, Document de travail, Université d'Evry- Val D'Essonne.

Chapitre 2 : Méthodologie de notation

De ces définitions on peut conclure que les banques concernées par cette approche s'appuient sur leurs estimations internes des composantes du risque afin de déterminer l'exigence en fonds propres en regard d'une exposition donnée. **Dumontier et al. (2008)**²⁹ affirment que les pondérations des actifs sont déterminées en fonction de quatre variables : la probabilité de défaut des emprunteurs, le taux de recouvrement des pertes, le montant exposé au risque et la maturité des engagements.

$$\text{RW} = f(\text{PD}, \text{LGD}, \text{EAD}, \text{M})$$

Avec RW = Pondération de risque (Risk Weight)

PD = Probabilité de défaut (Probability of Default).

LGD = Perte en cas de défaut (Loss Given Default).

EAD = Exposition au moment de défaut (Exposure At Default).

M = Maturité de l'actif.

Les techniques de notation interne permettent de fournir une notation et une probabilité de défaut pour chaque contrepartie. Il existe deux approches pour déterminer ces différents paramètres :

- **IRB de base** : Les banques estiment elles-mêmes la probabilité de défaut de chaque client et les régulateurs fournissent les autres éléments (LGD et EAD).
- **IRB avancée** : Les banques calculent elles-mêmes tous les des paramètres nécessaires à l'évaluation du risque de crédit. Cette méthode est plutôt appliquée par les plus grands établissements bancaires ayant des systèmes sophistiqués d'allocation de capital.

La notation interne a donc pour but de prédire la probabilité de défaillance d'un client et en conséquence de réduire les créances non performantes et de piloter leurs engagements. Comme toute méthode de prévision, il s'agit d'utiliser des modèles statistiques en s'appuyant sur l'historique des résultats de remboursement, des caractéristiques des clients et du type du crédit pour construire une fonction de score qui sera utilisée pour la discrimination (bon payeur / mauvais payeur) des nouveaux demandeurs de crédits. Dans ce cas, on parle d'un crédit scoring qui fera l'objet de la partie suivante.

²⁹ Dumontier, P. Dupre, D. et Martin C., **2008** : «Gestion et contrôle des risques bancaires : L'apport des IFRS et de Bâle II». Paris : Revue Banque Edition.

2.1. Définition et historique de « Credit Scoring »

En 1936, Fisher a mis en place une technique de discrimination intitulée « l'analyse discriminante » pour différencier trois types de variétés d'iris. Durand(1941) était le premier à utiliser ces techniques de discrimination afin de départager les bons et les mauvais demandeurs de crédit voiture en utilisant certaines caractéristiques de ces derniers. En 1958, le cabinet « Fair et Isaac » fût le pionnier dans l'automatisation des décisions d'accord de crédit. Cette firme continue d'être leader dans l'industrie du Credit Scoring (Thomas et al.2002) et à partir des années soixante l'industrie continue de croître et le traitement de masse des dossiers des demandeurs de crédit est devenu possible.

Le crédit scoring fût ensuite appliqué par Beaver (1968) aux entreprises pour prévoir leur défaillance. Il a utilisé dans son étude un échantillon de 79 entreprises ayant fait faillite entre 1954 et 1964 et il a conclu qu'à partir du ratio Cash flow/ Total dette le taux de bon classement a atteint 77%. Altman (1968) s'est aussi intéressé à l'activité d'octroi de crédit aux entreprises. Il a développé une technique d'analyse discriminante que nous présentons plus loin.

A partir des années 1980, le crédit scoring a connu un grand essor dans le domaine du crédit à la consommation et l'application de ces techniques de scoring n'a cessé d'augmenter (Malhotra & Malhotra, 2003³⁰, Sustersic et al. 2009³¹).

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

Cette technique « ni approuve, ni rejette une demande de prêt, elle peut plutôt prédire la probabilité d'occurrence de mauvaise performance (défaut) telle que définie par le prêteur » (Caire et Kossmann, 2003).

De leur part, Thomas et al. (2002) ont affirmé que « Le Credit Scoring constitue un ensemble de modèles de décision et les techniques sous-jacentes qui aident dans la décision d'octroi des crédits de consommation ».

Plusieurs explications peuvent être ainsi fournies sur la définition et le rôle du Credit Scoring. Elles peuvent être résumées sur la Figure 6.

³⁰ Malhotra, R. & Malhotra, D. K., 2003, «Evaluating consumer loans using Neural Networks». Omega the International Journal of Management Science, 31, 2, 83-96.

³¹ Sustersic, M., Mramor, D. & Zupan J., 2009, «Consumer credit scoring models with limited data». Expert Systems with Applications, 36, 3, 4736-4744.

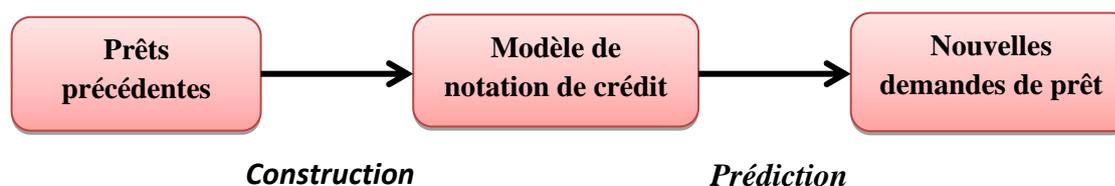


Figure 6 : Processus du Credit Scoring³²

Pour conclure, le Credit Scoring est un outil de gestion de risques qui vise à prédire la probabilité de défaut d'un nouveau prêt en utilisant les prêts précédents. Il existe une panoplie de méthodes statistiques de cette technique qui peuvent être utilisées au sein d'une banque. Nous présentons ci-après les modèles les plus courants.

2.2. L'analyse discriminante

L'analyse discriminante est privilégiée par les constructeurs de scores. C'est une technique statistique qui consiste à reclasser les emprunteurs en deux groupes : défaut et non défaut et à rechercher l'ensemble des variables quantitatives qui permettent de prévoir au mieux la défaillance des emprunteurs (**Beaver (1968)**).³³

Comme son nom l'indique, cette méthode cherche à discriminer entre deux classes d'individus ou d'entreprises et de tirer par la suite une conclusion concernant l'affectation d'une nouvelle relation dans le groupe qui lui convient. Elle a été proposée pour la première fois par Fisher (1936) puis elle a été appliquée par Altman (1968) pour pouvoir prévoir la défaillance des entreprises.

Sa formule générale est comme suit :

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

Où Z représente le score de l'entreprise, α est la constante et β_i représente les coefficients de la combinaison linéaire des variables explicatives X_i , avec $i = 1 \dots n$

2.2.1. Le modèle Altman (1968)

Il s'agit d'une application célèbre de l'analyse discriminante multivariée. Ce modèle est valable pour les entreprises de taille grande.

Altman a étudié 66 entreprises industrielles entre 1946 et 1965 dont la moitié a fait défaut. Il a considéré 5 facteurs (liquidité, solvabilité, rentabilité, structure financière et activité) et 22 ratios afin de sélectionner les ratios les plus explicatifs. Après avoir essayé toutes les

³² Adapté de Yang Liu(2000).

³³ Beaver, W.H., 1967, « Financial Ratios as Predictors of Failure », Journal of Accounting Research, Vol. 4, p. 71-111.

Chapitre 2 : Méthodologie de notation

combinaisons possibles de ces ratios, évalué les corrélations et mesuré la précision de la capacité prédictive, Altman a fini avec une fonction discriminante plus précise qui regroupe seulement 5 ratios.

La fonction de score d'Altman, couramment nommée Z-score, s'exprime par la relation suivante :

$$Z = 1.2 \times R_1 + 1.4 \times R_2 + 3.3 \times R_3 + 0.6 \times R_4 + 0.9 \times R_5$$

Avec $R_1 = \frac{\text{Fonds de roulement}}{\text{Total des actifs}}$; $R_2 = \frac{\text{Réserves}}{\text{Total des actifs}}$; $R_3 = \frac{\text{Excédent Brut d'exploitation}}{\text{Total des actifs}}$
 $R_4 = \frac{\text{Fonds propres}}{\text{Dettes totales}}$; $R_5 = \frac{\text{Chiffre d'affaires HT}}{\text{Total des actifs}}$

Le risque encouru par la banque varie dans le sens contraire de Z, avec 3 comme valeur critique tel que résumé par la figure ci-dessous présentée par **Hull et al. (2007)**.

Si le score d'une entreprise est inférieur à 3 elle est jugée défailante. Autrement elle serait considérée comme saine.

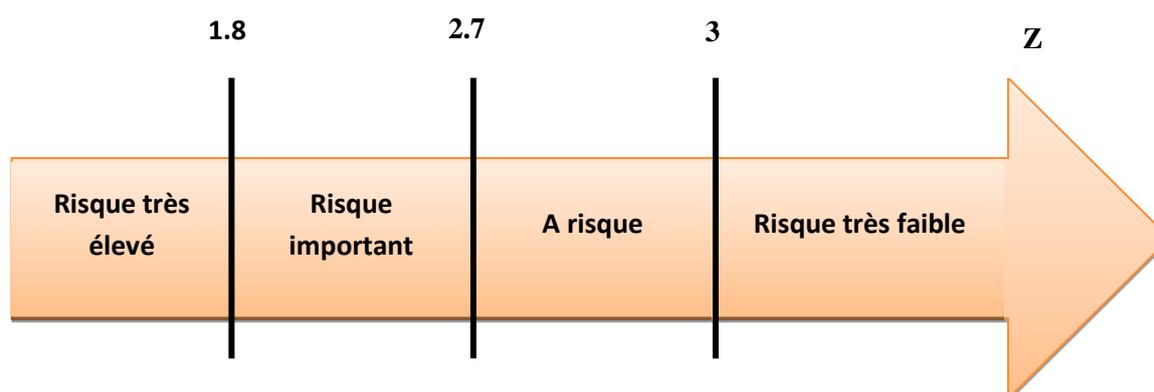


Figure 7 : Règles de décision pour le modèle d'Altman

Le taux de bon classement réalisé par Altman sur cet échantillon était de 95% et 82% sur un échantillon de validation.

Outre que le modèle d'Altman, Conan et Holder ont établi un autre modèle en 1979.

2.2.2. Le modèle de Conan et Holder (1979)

En se basant sur un échantillon de 190 petites et moyennes entreprises industrielles équitablement réparties entre deux classes (95 saines et 95 défailtantes), Conan et Holder ont commencé leur analyse avec 31 ratios dont seule 5 ratios étaient plus pertinents et plus significatifs et ont abouti à la formalisation de la fonction score (notée Z) par l'équation suivante :

$$Z = 0.24 \times R_1 + 0.22 \times R_2 + 0.16 \times R_3 - 0.87 \times R_4 - 0.1 \times R_5$$

Avec $R_1 = \frac{\text{Excédent Brut d'Exploitation}}{\text{Total des dettes}}$; $R_2 = \frac{\text{Capitaux permanents}}{\text{Total des actifs}}$;

$R_3 = \frac{\text{Valeur réalisable et disponible}}{\text{Total des actifs}}$; $R_4 = \frac{\text{Charges financières}}{\text{Chiffre d'affaires HT}}$; $R_5 = \frac{\text{Charges du personnel}}{\text{Valeur ajoutée}}$

Pour affiner le Z-score et perfectionner la règle de décision des banquiers, les deux auteurs proposent une probabilité de défaillance selon la valeur du score Z. En fonction de ce score, l'entreprise est considérée comme saine ou défaillante. Le calcul de cette loi de probabilité de défaillance permet de classer les entreprises selon leur niveau de risque présumé comme la montre la figure ci-dessous :

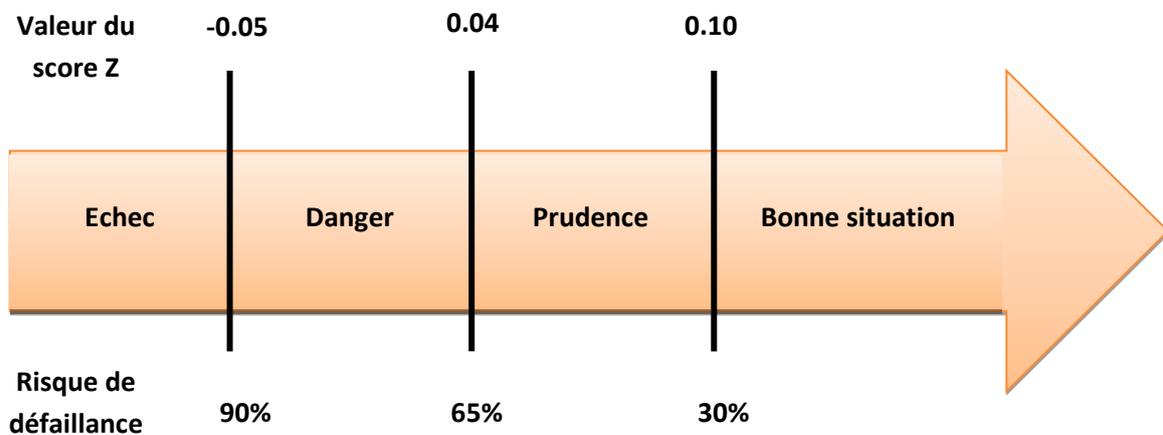


Figure 8 . Règles de décision pour le modèle Conan et Holder

Source : Sadi, 2009

- Si le score Z est > 0.1 : La situation financière de l'entreprise est bonne avec une probabilité de défaillance inférieure à 30%.
- Si $0.04 < Z < 0.10$: l'entreprise est dans une zone d'alerte avec une probabilité de défaillance de 30% à 65%.
- Si $-0.05 < Z < 0.04$: l'entreprise est dans une zone de danger avec une probabilité de défaillance de 65% à 90%.
- Et finalement, Si $Z < -0.05$: l'entreprise est classée comme défaillante avec une probabilité de défaillance qui dépasse 90%.

2.3. La régression logistique

2.3.1. La définition de la régression logistique

La régression logistique est une technique statistique qui permet d'expliquer une variable qualitative à deux modalités ou plus (dans notre cas : « défaillantes » et « saines ») appelée

aussi variable dichotomique. La spécificité de ce modèle consiste à la possibilité d'en introduire des variables explicatives discrètes ou continues, quantitatives ou qualitatives. Elle permet de prévoir la probabilité de l'occurrence de l'événement, soit la probabilité de défaillance.

Ceci est affirmé par **Tenenhaus (2007)** qui définit ce modèle comme étant « un modèle multivarié qui permet d'expliquer, sous forme de probabilité, la relation entre une variable dépendante Y qualitative le plus souvent binaire, $Y \in \{0,1\}$, et une ou plusieurs variables indépendantes X qui peuvent être quantitatives ou qualitatives.

2.3.2. Les propriétés mathématiques de la régression logistique³⁴

Soient :

Y_i = La variable à expliquer

- $Y_i = 1$ si la firme i est défaillante
- $Y_i = 0$ si la firme i est saine

X = la matrice des variables explicatives

Pour une entreprise i , on suppose que :

$$Y_i = \begin{cases} 1 & \text{si } X_i \beta + U_i > 0 \\ 0 & \text{si } X_i \beta + U_i \leq 0 \end{cases}$$

Avec β est le vecteur des coefficients d'une combinaison linéaire à estimer.

U_i est un terme d'erreur de moyenne nulle et de variance 1.

Le choix du modèle est lié au choix de la loi de U_i . Une hypothèse classique est de supposer qu'elle suit une loi logistique de fonction de répartition $F(x) = \frac{e^x}{1+e^x}$, il s'agit du modèle LOGIT. Une autre hypothèse consiste à supposer que les U_i suivent une loi normale, on parle dans ce cas d'un modèle PROBIT.

La méthode d'estimation des coefficients β est le maximum de vraisemblance.

Soit p_i la probabilité de non défaillance de l'entreprise i

$$p_i = P(Y_i = 0 / X_i) = P(X_i \beta + U_i \leq 0) = P(U_i \leq -X_i \beta) = F(-X_i \beta) = \frac{e^{-X_i \beta}}{1 + e^{-X_i \beta}} = F(-score)$$

La probabilité de défaut de l'entreprise i :

$$p_i = P(Y_i = 1 / X_i) = P(X_i \beta + U_i > 0) = P(U_i > -X_i \beta) = 1 - P(U_i \leq -X_i \beta) = 1 - F(-X_i \beta)$$

$$p_i = 1 - \frac{e^{-X_i \beta}}{1 + e^{-X_i \beta}} = \frac{e^{X_i \beta}}{1 + e^{X_i \beta}} = F(score)$$

Soit L : la vraisemblance de l'échantillon

³⁴ La référence de ce développement est le cours de Monsieur « Mouhamed Hlél »

$$L = \prod_{i=1}^n F[(-X_i\beta)]^{y_i} \times [F(X_i\beta)]^{(1-y_i)}$$

En admettant l'indépendance des observations y_1, y_2, \dots, y_n , relatives à la variable endogène, le logarithme de la vraisemblance s'écrit sous la forme :

$$\text{Log}(L(\beta)) = \sum_{i=1}^n y_i \times F(-X_i\beta) + (1 - y_i)[F(-X_i\beta)]$$

Le problème est d'estimer le vecteur des β qui maximise le logarithme de la vraisemblance.

2.3.3. Les principaux travaux empiriques

- **Le modèle de « Ohlson » 1980**

Parmi les premières études ayant utilisé l'analyse logit comme moyen de prévision de la défaillance des entreprises nous citons l'étude d'Ohlson. Cette étude a porté sur un échantillon de 105 entreprises cotées qui ont fait faillite sur la période 1970-1976 et un échantillon de 2058 entreprises saines choisies de façon aléatoire. Les ratios financiers introduits par Ohlson sont les suivants :

$$R_1 = \text{ratio taille} = \log \left(\frac{\text{Actif total}}{\text{PNB ajusté}} \right); R_2 = \frac{\text{Dettes totales}}{\text{Actif total}}; R_3 = \frac{\text{Fonds de roulement}}{\text{Actif total}}$$

$$R_4 = \frac{\text{Dettes à court terme}}{\text{Actif circulant}}; R_5 = 1 \text{ si Dettes totales} > \text{total actif}; 0 \text{ sinon}$$

$$R_6 = \frac{\text{Résultat net}}{\text{Actif total}}; R_7 = \frac{\text{Fonds générés par l'exploitation}}{\text{Dettes totales}}$$

$$R_8 = 1 \text{ si le résultat net est} < 0 \text{ pour les deux dernières années}; 0 \text{ sinon}$$

$$R_9 = \frac{\text{Résultat net } N - \text{résultat net } N-1}{|\text{Résultat net } N| + |\text{Résultat net } N-1|}$$

La fonction score obtenue appelée « Oscore » est la suivante :

$$\text{Oscore} = -1.32 - 0.407 R_1 + 6.03 R_2 - 1.43 R_3 + 0.0757 R_4 - 2.37 R_5 - 1.83 R_6 - 1.72 R_7 + 0.285 R_8 - 0.521 R_9$$

$$\text{La probabilité de défaillance} = \pi_i = \frac{e^{\text{Oscore}}}{1 + e^{\text{Oscore}}}$$

Les règles de décisions sont les suivantes :

- Si $\pi_i < 0.5$, l'entreprise est défaillante.
- Si $\pi_i \geq 0.5$, l'entreprise est saine.

- **Les modèles dans le contexte tunisien**

Matoussi(2010) a mené une analyse de la défaillance des entreprises tunisiennes, intitulée « la prédiction de faillite des entreprises tunisiennes par la régression logistique ». Cette étude est basée sur une base de données composée de 1435 dossiers de crédit octroyés aux entreprises industrielles tunisiennes durant la période 2003-2006.

Chapitre 2 : Méthodologie de notation

Au début, il a commencé l'analyse avec 26 ratios mais avec le problème de multicollinéarité 8 ratios sont retenus :

$$R_1 : \text{Liquidité des comptes clients} = \frac{\text{Clients-clients nets}}{\text{Clients bruts}} ; R_2 : \text{Couverture BFR} = \frac{\text{BFR}}{\text{FR}}$$

$$R_3 : \text{Couverture des dettes} = \frac{\text{Total passifs}}{\text{Chiffre d'affaires}} ; R_4 : \text{Valeur liquidative} = \frac{\text{Total passifs}}{\text{Total actifs}}$$

$$R_5 : \text{Structure financière} = \frac{\text{Endettement global}}{\text{Fonds propres nets}} ; R_6 : \text{Rotation des stocks} = \frac{\text{Chiffre d'affaires}}{\text{Stocks nets}}$$

$$R_7 : \text{Taille de l'entreprise} = \text{Log (total actifs)} ; R_8 : \text{Garanties} = \text{Log (garanties)}$$

La fonction score obtenue est la suivante :

$$Z = -12.122 R_1 + 0.501 R_2 + 22.9965 R_3 - 0.843 R_4 + 7.6334 R_5 + 0.096 R_6 + 3.5208 R_7 - 0.2477 R_8$$

Ce modèle a abouti à un taux de bon classement global de 88.7% sur l'échantillon initial et un taux de bon classement de 81.89% sur l'échantillon de validation.

Dans le même ordre des idées, **Mraïhi (2015)** a mené une analyse intitulée « Distressed Company Prediction Using Logistic Regression : Tunisian's Case ». Cette dernière est basée sur un échantillon de 212 entreprises équitablement sur les deux groupes (saines et défaillantes), sur la période 2005-2010. Il a commencé son analyse avec une batterie de 87 ratios pour finir avec 12 ratios significatifs à savoir :

$$R_1 : \text{Ratio de liquidité immédiate} = \frac{\text{Liquidité et équivalent de liquidité}}{\text{Passifs courants}}$$

$$R_2 : \text{Ratio de solvabilité} = \frac{\text{Capitaux permanents}}{\text{Total bilan}}$$

$$R_3 : \text{Ratio du degré de liquidité} = \frac{\text{Actifs courants}}{\text{Total actif}}$$

$$R_4 : \text{Ratio d'autonomie financière} = \frac{\text{Fonds propres}}{\text{Total actif}}$$

$$R_5 : \text{Ratio de structure d'endettement} = \frac{\text{Dettes à court terme}}{\text{Total passif}}$$

$$R_6 : \text{Ratio du degré de renouvellement des immobilisations} = \frac{\text{Amortissement des immobilisations}}{\text{Immobilisations brutes}}$$

$$R_7 : \text{Ratio du degré de liquidité} = \frac{\text{FR}}{\text{Total actif}}$$

$$R_8 : \text{Ratio de liquidité réduite} = \frac{\text{Actifs Courants hors stocks}}{\text{Passifs courants}}$$

$$R_9 : \text{Ratio de liquidité} = \frac{\text{Actifs circulant hors stocks}}{\text{Total actif}}$$

$$R_{10} : \text{Ratio d'endettement 1} = \frac{\text{Dettes à moyen et long terme}}{\text{Flux de trésorerie}}$$

$$R_{11} = \text{Ratio de rentabilité} = \frac{\text{Bénéfice net}}{\text{Total passif}} ; R_{12} = \text{Ratio d'endettement 2} = \frac{\text{Total passif}}{\text{Total actif}}$$

La fonction score obtenue est la suivante :

$$Z = 14,057 R_1 - 131,311 R_2 - 272,144 R_3 + 10,482 R_4 - 23,350 R_5 + 66,129 R_6 + 178,682 R_7 - 13,401 R_8 + 87,654 R_9 - 0,501 R_{10} - 15,515 R_{11} + 52,925 R_{12} + 126,426$$

Le taux de bon classement de ce modèle est le suivant :

Tableau 4 : Les taux de bon classement trouvés par F. Mraïhi (2015)

<i>Le nombre d'années avant la défaillance</i>	<i>Taux de bon classement</i>
1 année	100%
2 années	99.34%
3 années	96.71%

2.4. Les réseaux de neurones

Il s'agit d'une technique de « Credit Scoring » non paramétrique, c'est-à-dire qui ne prévoit aucune hypothèse quant à la distribution des données (ni la loi normale, ni la loi logistique...). La première utilisation des réseaux de neurones dans le domaine de prévision de défaillance, qui ait eu un certain retentissement, remonte aux travaux **d'Odom et Sharda (1990)**. Cette étude consiste à estimer la défaillance des entreprises à partir d'un petit échantillon (129 entreprises) et elle a abouti à un taux de bon classement de 81%.

Cette nouvelle technique a ouvert la voie à tout un courant de recherche laissant supposer qu'elle résoudrait notamment la question liée à l'absence de prise en compte, par les méthodes habituelles, du non linéarité des relations entre les ratios financiers.

Les réseaux de neurones ne constituent pas la seule famille de méthodes non linéaires, on trouve également les arbres de décision. En effet, les réseaux de neurones et les arbres de décision, sont préconisées de plus en plus (**Armingler, Enache et Bonne 1997, Hand et Henley 1997, Henley et Hand 1996 et Hand 2001**).

Les réseaux peuvent être classés de différentes manières : selon leur architecture (bouclé ou non), leur technique d'apprentissage (supervisée ou non) ou l'algorithme employé au cours de la procédure d'apprentissage.

Le réseau utilisé par la majorité des études sur la prévision de défaillance est le perceptron multicouche, qui fait partie de la famille des réseaux non-bouclés à apprentissage supervisé. Ce réseau (**Stéphane Tufféry, 2010**) se décompose en plusieurs couches : les variables en entrée, la ou les variables en sortie, et un ou plusieurs niveaux cachés. Chaque nœud d'un niveau est connecté à l'ensemble des nœuds du niveau précédent.

Chapitre 2 : Méthodologie de notation

Les réseaux de neurones se basent sur l'apprentissage, c'est-à-dire, que ces systèmes apprennent par eux-mêmes les relations entre les différentes variables et ce, à partir d'un échantillon de données et en simulant le raisonnement humain. Au final, ils nous permettent de mettre en relation les inputs (variables explicatives) et les outputs (le résultat) sous la supposition que cette relation est non linéaire.

Dans notre cas, le résultat peut être modélisé conformément au schéma ci-dessous :



Figure 9 : Schéma général du traitement

Un réseau de neurones est généralement formé de :

- Couche d'entrée représentant les neurones d'entrées (variables d'input).
- Couche de sortie représentant le vecteur des variables d'outputs (dans notre cas, la réponse est binaire : 0 ou 1)
- Couches cachées présentant l'ensemble des nœuds cachés des connexions entrantes qui proviennent des neurones d'entrée.

Ce réseau peut être synthétisé dans le schéma illustré par la Figure 10.

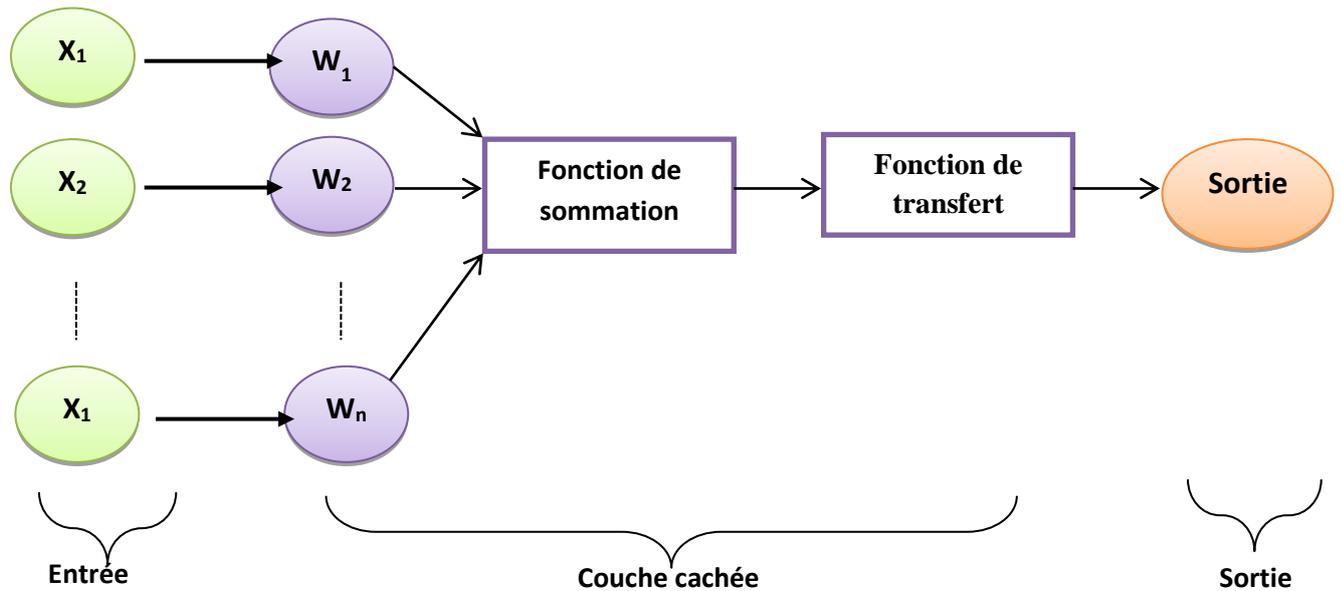


Figure 10 : Conception d'un réseau de neurones³⁵

La figure ci-dessus possède n entrées, à chaque entrée X_i est affecté un poids synaptique W_i représentatif de la force de connexion. Le neurone ne traite pas chaque information reçue unilatéralement, mais effectue une somme pondérée de toutes les entrées. Cette somme représente la fonction de sommation ou la fonction de combinaison : $a = \sum_{i=1}^n W_i X_i - b$ avec a est le niveau d'activation, qui est le signal total reçu par le neurone de la couche cachée est b est le biais de neurone ou le seuil d'activation du neurone.

Le résultat a de cette somme est transformée par une fonction de transfert (appelée aussi fonction d'activation), elle est généralement de type sigmoïde : $f(a) = \frac{1}{1+e^{-a}}$

Comparaison entre l'analyse discriminante et l'approche neuronale

La comparaison de ces deux méthodes est présentée à travers les résultats de plusieurs articles. De nombreuses études ont tenté de démontrer l'efficacité de l'une rapport à l'autre, les résultats obtenus sont présentés par Tableau 5.

Ces études ont affirmé la dominance des réseaux de neurones, ceci est confirmé par **Oreski et al (2012)** qui les considèrent comme « *un outil de prévision plus précis, adaptable et robuste que les méthodes statistiques classiques pour l'évaluation du risque de crédit* ».

Une autre méthode non paramétrique peut être utilisée pour prévoir la défaillance des entreprises est celle de l'arbre de décision, ce qui fera l'objet de la partie suivante.

³⁵ Adapté de : Tufféry, 2012

Tableau 5 : Comparaison entre l'approche neuronale et l'analyse discriminante

Auteur	Résultats obtenus
Odom et Sharda(1990)	Les réseaux de neurones sont plus performants que la méthode d'analyse discriminante : un taux de bon classement obtenu sur l'échantillon test est de l'ordre de 81.81% contre 74.28% pour l'analyse discriminante.
Coats et Fant (1993)	Les réseaux de neurones sont plus performants que la méthode d'analyse discriminante si l'on tient compte du type d'erreur commise et du coût associé.
Altman et al (1994)	L'obtention des meilleurs résultats nécessiterait l'utilisation à la fois des méthodes de réseaux de neurones et la technique de l'analyse discriminante.
Paquet(1997)	Importance de l'approche neuronale : Aucune hypothèse n'est nécessaire concernant la relation entre les variables et la distribution des termes d'erreur et des variables.
Abdou et al (2008)	L'approche neuronale domine la technique de l'analyse discriminante dans la mesure où les réseaux de neurones présentent un taux de bon classement de l'ordre de 94.84% contre 86.75% pour l'analyse discriminante.
Boujelbène.Y et Khemakhem.S (2013)	Une étude comparative entre ces deux approches présentée au sein de leur thèse a apparue l'approche neuronale comme un outil de prévision puissant avec un taux de bon classement de l'ordre de 80.23% contre 74.4% pour l'analyse discriminante.

2.5. Les arbres de décision

Les arbres de décision ont connu à partir de 1984 un grand essor avec les travaux de **Breiman et al** sous l'acronyme de CART (Classification and Regression Tree).

Egalement, **Frydman et al (1985)** les utilisent dans le domaine « Credit Scoring » et il a abouti à un taux de bon classement de l'ordre de 89%.

Russell et Norvig(2003) considèrent que « *l'induction avec des arbres de décision est l'une des formes d'algorithme d'apprentissage les plus simples et pourtant les plus efficaces* ».

Selon **Catherine (2004)**, cette méthode consiste à mettre en place un arbre décisionnel en se basant sur un échantillon de firmes (défaillantes et saines) qui seront décomposées en différents sous-groupes en se basant sur différents critères. Cette opération sera répétée jusqu'à ce qu'on obtient des sous-groupes composés que d'entreprises défaillantes ou saines. On obtient alors des nœuds dits « purs ».

De ce fait, il s'agit de construire un arbre de décision à l'aide de divisions successives des individus d'un échantillon en deux, ou plus segments (appelés également nœuds) homogènes par rapport à une variable dépendante Y de nature binaire dans notre cas.

Il est à noter que cette méthode ne requière pas d'hypothèses sur la distribution des variables. Son principe de fonctionnement est le suivant : pour expliquer une variable, le système recherche le critère le plus déterminant et découpe la population en sous populations possédant la même entité de ce critère. Chaque sous population est ensuite analysée comme la population initiale.

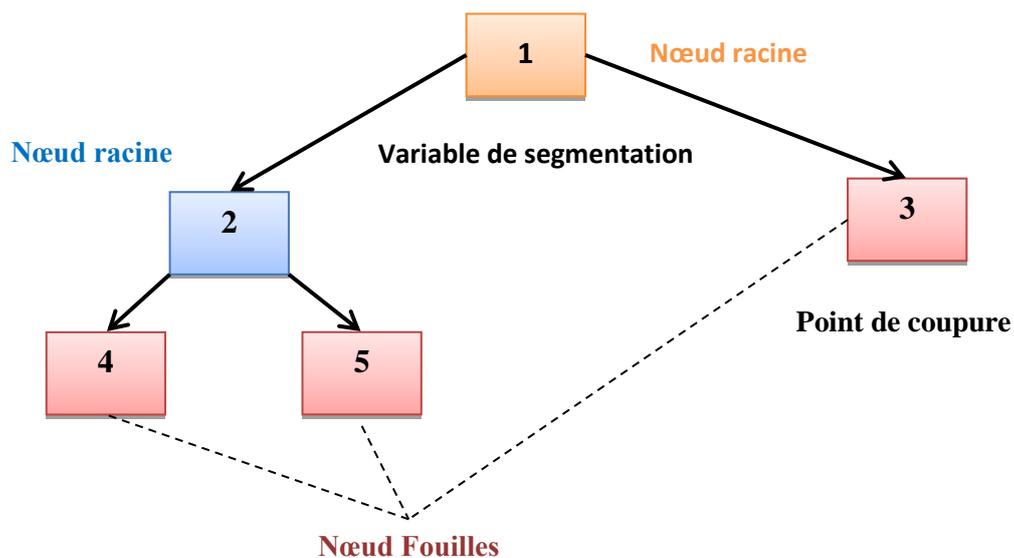


Figure 11 : Arbre de décision

La construction de l'arbre de décision repose essentiellement sur trois points à savoir :

- Le choix de la variable de segmentation la plus discriminante
- Le choix du point de coupure pour les variables continues
- Le choix de la taille de l'arbre ou le nombre de feuilles

Afin de définir ces critères, il existe plusieurs méthodes d'induction d'arbres tels que l'arbre CHAID, CART, ID3, C4, 5

2.5.1. Arbre CART

L'acronyme CART signifie Classification And Regression Tree. Il désigne une méthode statistique, introduite par **Breiman et al(1984)**. Bâtir un arbre CART se fait en deux étapes :

Une première phase consiste à la construction d'un arbre maximal, qui permet de définir la famille de modèles à l'intérieur de laquelle on cherchera à sélectionner le meilleur.

Une deuxième phase est l'élagage : C'est le fait de chercher le meilleur sous-arbre élagué de l'arbre maximal dans le sens où il minimise le taux d'erreur (réduction du critère d'impureté) à l'aide d'un échantillon test.

2.5.2. Arbre CHAID

L'acronyme CHAID signifie Chi-squared Automatic Interaction Detector, initialement proposée par **Kass (1980)**. Dans ce cas, la segmentation s'effectue selon la variable (X_i) qui a le plus fort lien statistique avec la variable à estimer (Y_i). Selon cette méthode, le choix de la variable de segmentation, du point de coupure et de taille de l'arbre s'effectue comme suit :

- **Le choix de la variable de segmentation**

Selon **Rakotomalala(2005)**, il s'agit de choisir « X » telle qu'elle est la plus corrélée avec « Y ». Autrement dit, il faut calculer un indicateur de qualité et la variable retenue est celle qui optimise cet indicateur. Ce dernier est l'écart à l'indépendance (khi deux : x^2), il se calcule comme suit :

Tableau 6 : Tableau des effectifs lors du croisement des deux variables

X/ Y	Y= 1 Défaillante	Y=2 Saine	Σ
X_1	n_{11}	n_{12}	L_1
X_2	n_{21}	n_{22}	L_2
X_i	n_{i1}	n_{i2}	L_i
X_n	n_{n1}	n_{n2}	L_n
Σ	C_1	C_2	n

$$X^2 = \sum_{i=1}^n \frac{(n_{i1} - e_{i1})^2}{e_{i1}} + \sum_{i=1}^n \frac{(n_{i2} - e_{i2})^2}{e_{i2}}$$

Avec : $e_{i1} = \frac{L_i \times C_1}{n}$; $e_{i2} = \frac{L_i \times C_2}{n}$

On choisit la variable qui est la plus corrélée avec Y, autrement dit, qui a la valeur de X^2 la plus élevée.

▪ *Le choix de points de coupure pour les variables continues*

Selon **Rakotomalala(2005)**, le choix du seuil de discrétisation doit être cohérent avec le choix de la variable de segmentation. On doit donc choisir le seuil qui maximise le X^2 .

En effet, nous allons prendre quelques valeurs possibles de cette variable et calculer le X^2 pour chaque valeur et choisir celle qui offre un X^2 plus important.

Rakotomalala(2005) a proposé autre méthode afin de déterminer le meilleur découpage et ce en normalisant cet indicateur en calculant le Tschuprow tel que : $t = \frac{X^2}{n\sqrt{L}-1}$ dont le domaine de définition est $[0 ; 1]$. Le meilleur découpage produit un Tschuprow est égale à 1.

▪ *Le choix de la taille de l'arbre*

Il a été démontré également que la taille des arbres a tendance à croître avec le nombre d'observations dans la base d'apprentissage (**Oates et Jensen, 1997**). L'arbre CHAID est basé sur le **pré-élagage** : C'est le fait de fixer une règle d'arrêt qui permet de stopper la construction de l'arbre lors de la phase de construction.

Selon la méthode CHAID, on accepte la segmentation si le khi deux calculé est significatif au seuil de 1% ou 5% ou 10% autrement dit si la p-value est < au seuil choisi : Il s'agit du test d'indépendance de khi deux :

$$\left\{ \begin{array}{l} H_0 : Y \text{ et } X \text{ sont indépendants} \\ H_1 : Y \text{ et } X \text{ sont dépendants} \end{array} \right.$$

Rakotomalala (2005) a indiqué l'existence d'autres critères plus empiriques relatifs au choix de la taille :

- Éviter d'avoir des sommets avec un effectif trop faible afin de rendre la prévision plus fiable.
- Fixer le nombre de feuilles à partir duquel on stoppe la segmentation ou fixer l'effectif d'admissibilité : si une feuille contient un effectif inférieur au seuil on arrête la segmentation.

Une fois ces critères sont choisis et l'arbre est construit, on passe à la dernière étape qui est celle de décision. Lorsque la feuille est pure (c'est-à-dire qui ne contient que des entreprises

saines ou défaillantes), lui attribuer la conclusion correspond à la seule modalité présente semble naturel. En revanche, lorsque plusieurs modalités sont présentes dans la feuille, il faut utiliser une règle d'attribution efficace. La règle la plus souvent utilisée est la règle de la majorité : on affecte à la feuille la modalité de la variable à prédire qui présente l'effectif le plus grand.

2.6. Comparaison entre les différentes méthodes statistiques

Tableau 7 : Comparaison entre les différentes méthodes statistiques (adapté de Tufféry, 2012)

Technique statistique	Avantages	Inconvénients
Analyse discriminante	<ul style="list-style-type: none"> - Possède une solution analytique directe, les calculs sont rapides, -Les coefficients des combinaisons linéaires constituent des résultats relativement interprétables. 	<ul style="list-style-type: none"> -Variables explicatives continues et sans valeurs manquantes, -Sensible aux individus hors norme, -Ne détecte que les phénomènes linéaires.
Régression logistique	<ul style="list-style-type: none"> -Pas d'hypothèses de multinormalité, ni d'homoscédasticités pour les variables explicatives, -Modélise directement une probabilité de défaut, -De nombreux tests statistiques, notamment la significativité des coefficients, sont disponibles, -Fournit des intervalles de confiance pour les résultats. 	<ul style="list-style-type: none"> -Les variables explicatives doivent être non colinéaires, -Ne traite pas les valeurs manquantes, -Sensible aux valeurs hors normes, -Sa précision est moindre que celle de l'analyse discriminante lorsque les hypothèses de cette dernière sont remplies.
Réseaux de neurones	<ul style="list-style-type: none"> -Modéliser des relations non linéaires entre les données, -Résiste aux données défectueuses (si une variable en entrée est trop bruitée, le nœud correspondant verra son 	<ul style="list-style-type: none"> -La convergence vers la meilleure solution globale n'est pas toujours garantie, -Les résultats ne sont pas explicites et sont difficile à comprendre par les utilisateurs,

	poids diminuer jusqu'à zéro).	-Le risque de sur-apprentissage.
Arbre de décision	-Simple à utiliser, -Traite les données manquantes, -Compréhensibilité des résultats pour les utilisateurs, -Les résultats sont exprimés sous forme de conditions explicites sur les variables d'origine.	-La détermination des nœuds du niveau (n+1) dépend fortement du nœud précédent (n), -L'apprentissage d'un arbre de décision nécessite un nombre assez grand d'individus.

Conclusion

Selon **Weber et Krahn**(2000), il existe une variété de procédures pour arriver à une note : la procédure typique utilisée est le scoring. Cette procédure comprend des techniques statistiques classiques (analyse discriminante, régression logistique...) et autres techniques dans le cadre des approches d'apprentissage automatique en Intelligence Artificielle (réseaux de neurones, arbre de décision...).

Selon **Koutanaei**(2015), les méthodes de notation de crédit ont plusieurs avantages :

- Une baisse des coûts de l'analyse du dossier de crédit,
- Un processus décisionnel efficace et rapide (lors de l'octroi du crédit),
- Une forte probabilité de remboursement du crédit,
- Un niveau faible de risque.

En utilisant ces méthodes, la banque renforce le contrôle du risque de crédit et pourrait par la suite calculer la provision nécessaire sur la base d'un niveau réel du risque qu'elle courue et non pas d'une manière forfaitaire ce qui conduirai à une amélioration de sa productivité ainsi qu'à sa rentabilité. Selon **S.Tuffery**(2007), ces méthodes permettent à la banque de calculer différents scores qui sont au nombre de cinq :

- Le **score d'appétence** : Il permet d'estimer l'intérêt qu'un client porte à un produit bancaire.
- Le score **comportemental** : Il permet d'estimer le risque de non remboursement tout au long de la durée d'emprunt.
- Le score **d'octroi** : Il permet d'estimer le risque d'un nouveau dossier de crédit.

Chapitre 2 : Méthodologie de notation

- Le score de **recouvrement** : Il permet d'estimer le montant susceptible d'être récupéré dans un cas de non remboursement.
- Le score d'**attrition** : Il permet d'estimer la probabilité qu'un client quitte la banque.

Dans ce qui suit, nous allons focaliser sur le score comportemental qui sera déterminé par l'utilisation de différentes techniques statistiques.

**Chapitre 3 : Étude empirique des
méthodes du « Credit Scoring » :
Application sur un échantillon de la
Banque de l'Habitat**

Chapitre 3 : Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Introduction

Au vue de ce qui a été exposé précédemment, nous entamons dans ce dernier chapitre une analyse empirique. Cette dernière consiste à la mise en place d'un système de « Credit scoring », un tel système s'avère indispensable afin de pouvoir gérer le risque de crédit de la manière la plus efficiente et par conséquent se conformer à la réglementation prudentielle internationale (comité de Bâle) et la réglementation nationale (la circulaire de la BCT 2016-06 publiée le 11 octobre 2016).

L'objectif de notre étude sera donc le développement d'un modèle de prévision de la défaillance des entreprises. Pour ce faire, nous allons développer des modèles statistiques (analyse discriminante, régression logistique) et autres techniques d'Intelligence Artificielle (arbre de décision, réseaux de neurones...). Nous allons par la suite comparer leur performance et retenir le modèle le plus approprié à notre étude qui sera utilisé pour l'instauration du système de notation désiré.

Avant de procéder à la construction des modèles, une étude statistique exploratrice des données s'impose afin de mieux comprendre leur structure. Cette étude va nous permettre de présenter le choix des variables de l'étude avec lesquels nous allons modéliser le défaut de remboursement ainsi que la démarche de la construction des modèles appropriés.

Une fois que le modèle le plus adéquat est construit, nous allons procéder à une classification des entreprises selon leurs scores en plusieurs classes. Ces classes vont correspondre à des niveaux différents de risque.

A cet effet, ce chapitre est scindé en deux sections. Dans la première, nous présenterons la structure de l'organisme d'accueil, puis dans la deuxième section, nous construirons le crédit scoring à base des clients de la banque de l'Habitat.

Section 1 : Présentation de la BH

1.1. La présentation de la BH

La banque de l'Habitat a été créée en mai 1989 par la transformation de la CNEL (Caisse Nationale d'Épargne Logement) en banque commerciale. C'est une société anonyme régie par la *loi du 10 juillet 2001* relative aux établissements de crédit.

Le capital social de la banque s'élève actuellement à 238 millions de dinars.

Avant de présenter les principaux indicateurs et ratios de l'activité, nous présentons les faits marquants des années 2017 et 2018.

✓ *Les faits marquants de l'année 2017*

- Augmentation du capital social pour le porter de 170 MD à 238 MD.
- Émission d'un emprunt obligataire subordonné d'un montant de 70 MD

D'où le : **Renforcement des quasi-fonds propres.**

- Signature d'une convention de prêt avec la BEI (la Banque Européenne d'Investissement) pour 120 M€.

D'où la **Consolidation des ressources longues.**

✓ *Les faits marquants de l'année 2018*

- Emprunt obligataire subordonné « *BH-subordonné 2018-1* » d'un montant de 50MD, divisé en 500 000 obligations subordonnées de nominal 100 dinars. L'emprunt est susceptible d'être porté à 70 MD, divisé en 700 000 obligations subordonnées de nominal 100 dinars.

D'où le **Renforcement des ressources stables de la banque ainsi que ses fonds propres et amélioration de son ratio de solvabilité.**

1.2. Activité et performance de la banque

1.2.1. Les principaux indicateurs

Tableau 8 Les principaux indicateurs de la BH en 2016, 2017 et le premier trimestre de l'année 2018

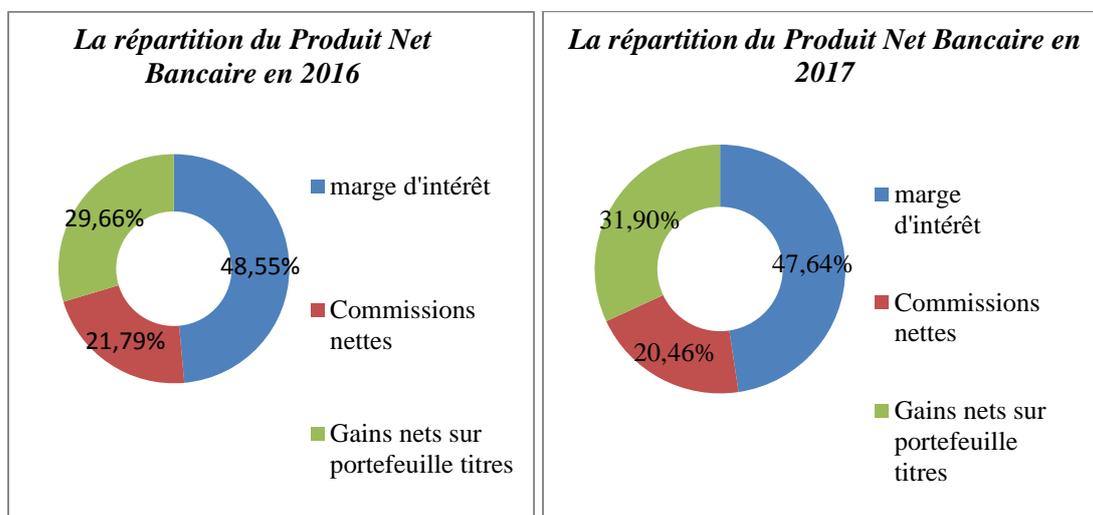
Indicateur	Le montant en Mille Dinars au 31/12/2016	Le montant en Mille Dinars au 31/12/2017	Le montant en Mille Dinars au 30/06/2018
Total bilan	8 240 102	9 991 217	10 653 466
Dépôts et avoirs	5 194 082	5 941 916	6 196 590

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

de la clientèle			
Créances sur la clientèle	6 272 654	7 736 647	8 396 771
Capitaux propres	540 954	758 992	788 454
Produit net bancaire	319 627	385 620	204 526
La marge d'intérêt (Intérêts reçus- Intérêts payés)	147 605	183 652	95 390
Commissions nettes	75 470	78 924	45 360
Résultat net de l'exercice	102 548	115 528	57 738

La banque de l'Habitat gère des capitaux moyens de 8 240 102 mille dinars au 31/12/2016, ces derniers ont augmenté de 21.25% en 2017. Ces capitaux sont employés essentiellement dans l'activité de crédit à la clientèle (77.44% en 2017) avec un rendement moyen de 6.36%. Ces emplois sont financés par des ressources qui proviennent essentiellement des dépôts clients qui représentent 59.47% des capitaux en 2017 ainsi que des capitaux propres (7.6%). Le coût moyen de ces ressources est de 3.73% ; **Soit un écart de 2.63%**.

La répartition du produit net bancaire durant les années 2016 et 2017 est donnée par la figure 13 :



Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Figure 12 La structure du PNB de la BH en 2016/2017

La banque réalise une marge d'intérêt de 149.3 MDT en 2016 contre 183.7MDT en 2017, soit une hausse de 23%. En effet, l'activité d'intermédiation représente 47.64% de son PNB en 2017 alors que l'activité de service (commissions nettes) et celle du portefeuille représentent respectivement 20.46% et 31.90% du PNB en 2017.

1.2.2. Les principaux ratios

Tableau 9 Les ratios de la banque de l'Habitat en 2016/2017

Ratio	Mesure au 31/12/2016	Mesure au 31/12/2017
ROE (Bénéfice net/Capitaux propres)	16.71%	15.22%
ROA (Bénéfice net/ Total actif)	1.12%	1.16%
Coefficient d'exploitation (Frais généraux+Dotations aux amortissements/ PNB)	45.94%	42.98%
Ratio de solvabilité	10.03%	10.94%
Ratio de liquidité	82.59%	87.35%
Ratio de couverture des créances classées	76.2%	78%
NPL (créances classées/total engagements)	15.1%	13%

La rentabilité économique s'est améliorée en 2017, en revanche, la rentabilité financière a diminué de 1.49%. Le ratio de solvabilité a augmenté de 0.91% et il reste toujours conforme aux règles prudentielles (>10%, le minimum prévu par l'article 9 de la circulaire n°2018-06), ceci s'explique par la consolidation des fonds propres qui provient de l'amélioration du résultat net de la banque, l'émission d'un emprunt obligataire subordonné de 70 MD et l'augmentation du capital social.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Également, le ratio de liquidité s'est élevé de 4.76% en 2017. Ce dernier est conforme à la réglementation prudentielle (>80% à compter du 1^{er} janvier 2017) dictée par *le premier article de la circulaire aux banques n°2014-14* relative au ratio de liquidité.

L'évolution du PNB à une cadence plus élevée que celle des frais généraux a amélioré le coefficient d'exploitation qui s'est situé à 42.9% en 2017 contre 45.9% en 2016.

Dans le même ordre des idées, le ratio de couverture des créances classées s'est amélioré en passant de 15.1% en 2016 à 13% en 2017. Ceci est expliqué par l'amélioration de la qualité d'actifs de la banque.

1.2.3. La position de la BH dans le secteur

La BH s'est positionnée au **31/12/2017** :

- ✓ À la **3^{ème} place** en termes des **crédits à la clientèle** gagnant près d'un point de pourcentage de part de marché atteignant 12.91% au 31/12/2017.
- ✓ À la **3^{ème} place** en termes du **PNB** avec une part de marché de 10.5% en 2017 contre 9.9% en 2016.
- ✓ À la **3^{ème} place** en termes de résultat brut d'exploitation avec une part de marché de 11.62% en 2017 contre 11.08 en 2016.
- ✓ À la **4^{ème} place** en termes d'encours des **dépôts clientèles** avec une part de marché en amélioration, soit 10.54% en 2017 contre 10.17 en 2016.

Passant au taux de crédits non performants, ce dernier a atteint 13% en 2017 ce qui représente 10.3% du total des crédits non performants dans le secteur. Les créances douteuses de la BH au cours de cette année proviennent essentiellement de 3 secteurs d'activité :

- Service (32.54%)
- Tourisme (21.43%)
- Industrie (16.34%)

1.2.4. Le portefeuille des créances de la BH

La banque de l'Habitat procède à l'évaluation de ses engagements et au calcul des provisions conformément à *l'article 10 de la circulaire 91-24 de la BCT* (voir section 3 du premier chapitre).

Les créances brutes à la clientèle, hors engagements par signatures, par classe de risque se présentent comme suit :

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Tableau 10 Les créances classées de la BH en 2016 et 2017

Désignation	Le montant en Mille Dinars au 31/12/2016	Le montant en Mille Dinars au 31/12/2017
Classe 0 et 1	6 026 484	7 584 532
Classe 2	51 735	42 571
Classe 3	105 508	67 428
Classe 4	619 828	593 650
Classe 5	525 870	537 158
Provisions et agios réservés	1 237 716	1 284 115

La proportion des créances appartenant aux classes 0 et 1 a augmenté de 3.72% en 2017 par rapport au 2016 ce qui explique la baisse du taux des créances classées qui a passé de 15.1% à 13%.

Le reste des créances (classe 2, 3, 4 et 5) représente 14.06% de la totalité des créances classées en 2017 contre 17.78% en 2016.

Dans ce cas, nous pouvons affirmer que l'exercice 2017 s'est caractérisé par une meilleure maîtrise des risques se traduisant par l'amélioration du ratio des créances non performantes et celui de couverture des créances classées. Ceci est dû au projet lancé par la banque lié au système de notation interne qui permettra une meilleure sélectivité de la clientèle et constituera un pas vers l'adoption des normes bâloises en matière prudentielle. L'avancement de ce projet va être accéléré en 2018 afin d'atteindre une meilleure gestion des risques.

En s'appuyant sur cet aspect, nous venons de présenter dans la section suivante une application de la « notation interne» sur un échantillon d'entreprises de la BH.

Section 2 : Données et méthodologie de l'étude

2.1. L'échantillon de l'étude

Notre étude est centrée sur un échantillon de base constitué de 634 petites et moyennes entreprises domiciliées auprès de la BH dont les états financiers datent de 2015. Cet échantillon est réparti comme l'indique le Tableau 11.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Tableau 11 Répartition de l'échantillon initial

Échantillon initial	Entreprises saines (0)	Entreprises défaillantes (1)
634	620	14

Nous remarquons que les firmes saines sont beaucoup plus nombreuses que celles défaillantes c'est pour cela nous avons procédé à un tirage aléatoire avec remise de façon à constituer un ensemble représentatif et de taille comparable au groupe des firmes non défaillantes. Nous n'avons pas ciblé un secteur d'activité bien déterminé, les entreprises appartiennent à des secteurs différents (industriel, commercial, service...).

20.8% de l'échantillon initial a été réservé pour estimer le taux d'erreur et donc valider le modèle. L'échantillon de construction comporte par la suite 588 observations qui sont réparties ainsi :

Tableau 12 La répartition de l'échantillon de construction entre les deux groupes

Variable	Modalité	Effectifs	%
Classe	0 Saine	491	83.503
	1 Défaillante	97	16.497

Le reste des observations vont être consacrées à l'étape de validation, Ces dernières sont détaillées dans le tableau 13 :

Tableau 13 La répartition de l'échantillon test entre les deux groupes

Variable	Modalité	Effectifs	%
Classe	0 Saine	129	83.766
	1 Défaillante	25	16.233

L'entreprise est considérée défaillante si elle a un impayé de 3 mois et plus ce qui correspond à la définition bâloise du défaut.

Au final, les entreprises sont réparties en deux échantillons : un échantillon de construction et un échantillon de validation, comme l'indique le graphique ci-dessous :

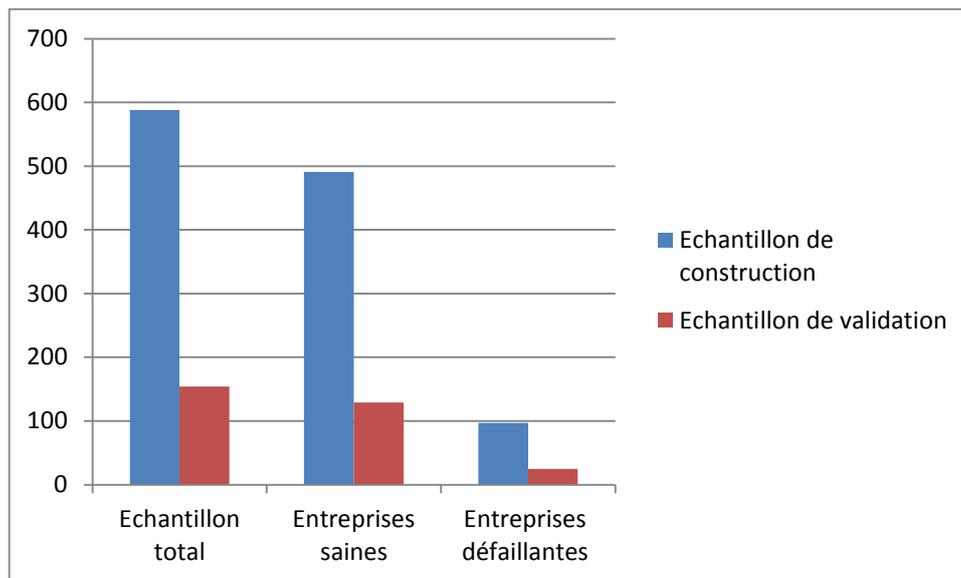


Figure 13 La répartition des deux sous-échantillons

2.2. Choix des variables de l'étude

2.2.1. Justification du choix des variables

Les ratios financiers sont les variables le plus souvent utilisées dans les modèles de prévision du risque. Les ratios retenus ainsi que leur nombre varient fortement d'une recherche à l'autre. A titre d'exemple, il existe des recherches qui ont été menées avec un seul ratio **Beaver (1966)**, 6 ratios **Bardos(1989)** et **Laitinen(1991)**, 5 ratios **Altman (1968)** et 7 ratios **Zavgren (1985)**.

La plupart des études menées sur la défaillance ont montré que les dimensions endettement, rentabilité et la liquidité sont les plus explicatives de la défaillance, **Dumontier (2001)**. Dans le contexte tunisien, l'étude de **Matoussi, Mouelhi et Sayah (1999)** a permis de conclure que les dimensions liquidité, solvabilité, équilibre financier, rentabilité et autonomie financière sont les plus explicatives.

En effet, les études de **St-Cyr et Pinsonneault(1997)** ont confirmé le pouvoir discriminant de l'endettement : Plus une entreprise est endettée, plus il y a de risque qu'elle éprouve des problèmes de solvabilité un jour ou l'autre. L'utilisation de la dette induit des risques relatifs à la variabilité du rendement et augmente la probabilité d'insolvabilité.

En outre, **Mraïhi (2015)** a mené une analyse sur des entreprises tunisiennes et il a conclu que la variable liquidité a un impact positif sur la survie des firmes. En effet, les ratios de liquidité doivent être suffisamment importants afin de refléter la capacité de la firme à honorer ses engagements donc plus le ratio est important plus l'entreprises est solvable, par conséquent plus sa probabilité de défaillance est faible.

Concernant la rentabilité économique (ROA), la littérature a prouvé qu'il s'agit d'une variable significative : **Altman (1968)**, **Altman, Haldeman and Narayanan (1977)**, **Izan (1984)**, **Mcgurr and DeVaney(1998)**, **Laitinen and Laitinen (2000)**, **Zapranis and Ginoglou (2000)**, **Ginoglou, Agorastos and Hatzigagios (2002)** and **Beaver, McNichols and Rhie (2005)**, ont approuvé la pertinence de ce ratio dans l'explication de la défaillance des entreprises.

Nous avons introduit aussi les ratios portant sur le **cash flow** du fait que **Hol et al. (2002)** décrivent la défaillance comme étant la situation où la valeur du cash flow est insuffisante pour couvrir les dettes. En outre, **Liang et Wu (2003)** considèrent qu'il y a situation de défaillance lorsque l'état d'un cash flow est non efficient. Dans le même ordre des idées, **Casey et Bartczak (1985)** ont abouti à ce que l'information du cash flow augmente la capacité prédictive de l'analyse discriminante et le modèle Logit dans la distinction entre les entreprises saines et défaillantes.

À côté de ces ratios, nous trouvons autres formes qui jouent des rôles secondaires, sans être totalement négligeables. Elles désignent toutes les transformations appliquées à des ratios ou à des variables financières à l'aide de fonctions statistiques ou mathématiques. D'ailleurs, depuis **Altman (1968)**, le logarithme de la variable « total actifs » fait partie de ces variables qui présentent un réel pouvoir discriminant : Il s'agit de distinguer entre les grandes entreprises et les PME.

L'étude de la défaillance des entreprises est basée essentiellement sur les variables financières. Toutefois, il existe autre type de variables dites comportementales qui peuvent être ajoutées et qui contribuent à la prévention de la défaillance des entreprises. Ces dernières sont plus spécifiques dans le sens où elles indiquent le comportement propre de chaque entreprise vis-à-vis sa banque. Il est à noter que ce type de variable ne peut être fourni que par la banque concernée (mouvements créditeurs par rapport au chiffre d'affaires, nombre de jours débiteurs, ancienneté de la relation, solde des impayés par rapport au chiffre d'affaires...).

Une multitude de ratios sera donc organisée par thème (solvabilité, liquidité, structure financière...). Par la suite, seules les variables significatives et peu corrélées entre elles seront finalement retenues.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

2.2.2. Présentation des variables

Nous avons commencé l'étude avec une batterie de 22 ratios dont 18 sont financiers et 4 sont comportementaux. Ces derniers reflètent les aspects les plus importants de l'entreprise et sont susceptibles d'expliquer la défaillance.

La base de données clients a été obtenue auprès de la direction de l'évaluation et le suivi des risques de la banque l'habitat.

Les ratios choisis sont présentés dans le Tableau 14.

Tableau 14 La liste des ratios

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

La variable	La mesure
Ratios financiers	
Ratios de structure	
Indicateur de taille	$R1 = \log(\text{actifs})$
Autonomie financière	$R2 = \text{Capitaux propres} / \text{total passif}$
Ratio d'endettement à MLT	$R3 = \text{Dettes long et moyen terme (DLMT)} / \text{total passif}$
Ratios de rentabilité	
Rentabilité financière	$R4 = \text{Résultat net} / \text{Capitaux propres}$
Rentabilité économique 1	$R5 = \text{Résultat net} / \text{Total actif}$
Rentabilité économique 2	$R6 = \text{Cash flow} / \text{Total actif}$
Marge brute d'autofinancement	$R7 = \text{Cash flow} / \text{Chiffre d'affaires}$
Marge nette	$R8 = \text{Résultat net} / \text{Chiffre d'affaires}$
Ratios de rotation et de gestion	
Rotation de l'actif	$R9 = \text{Chiffre d'affaires} / \text{Total actif}$
Poids des frais de personnels	$R10 = \text{Charge salariale} / \text{Chiffre d'affaires}$
Délai de paiement accordé à la clientèle	$R11 = \text{Clients et comptes rattachés} / \text{Chiffre d'affaires}$
Poids des charges financières	$R12 = \text{Frais financiers} / \text{Chiffre d'affaires}$
Couverture BFR	$R13 = \text{BFR} / \text{FR}$
Couverture des dettes à CT par le chiffre d'affaires	$R14 = \text{Passifs courants} / \text{Chiffre d'affaires}$
Ratios de liquidité et de financement	
Liquidité 1	$R15 = \text{FR} / \text{Total actif}$
Liquidité 2	$R16 = \text{Actifs courants} / \text{Passifs courants}$

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Liquidité 3	R17 = Actifs courants/ Total actif
Structure de financement	R18 = Dettes long et moyen terme (DLMT) / Capitaux propres
Ratios comportementaux	
	R19 = Mouvements créditeurs/Chiffre d'affaires
	R20 = Nombre de jours débiteurs
	R21 = Ancienneté de la relation ³⁶
	R22=Solde des impayés/ Chiffre d'affaires

2.2.3. Analyse de la corrélation

Avant de procéder à l'estimation des modèles, nous avons commencé par l'analyse de la corrélation entre les différents ratios. Cette étape reste indispensable **DeLaurentis(2010)** car :

- Une corrélation trop importante peut conduire à un problème dans l'estimation des coefficients de la fonction score,
- Des coefficients de corrélation importants illustrent une quantité d'informations similaires contenues dans les variables, ce qui peut être source de confusion pour l'analyse,
- Certaines techniques, comme la régression logistique sont sensibles à la multicolinéarité des variables.

Nous tenons à ce que les ratios corrélés entre eux ne figurent pas dans la même équation.

Hypothèse : La présence de corrélation est détectée lorsque le coefficient de corrélation est supérieur en valeur absolue à 0.7.

L'analyse fait apparaître des fortes corrélations entre les ratios ci-dessous. Ces corrélations sont significatives ($p\text{-value} < 0.05$) :

- R_5 et R_6 : 0.954 ; $p\text{-value} = < 0.0001$
- R_5 et R_{15} : 0.741 ; $p\text{-value} = < 0.0001$
- R_6 et R_{15} : 0.793 ; $p\text{-value} = < 0.0001$

³⁶L'ancienneté de la relation est la différence entre la date d'observation (2015) et la date de création de l'entreprise dans la base de donnée client.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

- R₇ et R₈: 0.989 ; p-value = <0.0001
- R₇ et R₁₀: -0.868 ; p-value = <0.0001
- R₇ et R₁₂: -0.754 ; p-value = <0.0001
- R₇ et R₁₄: -0.960 ; p-value = <0.0001
- R₈ et R₁₀: -0.929 ; p-value = <0.0001
- R₈ et R₁₂: -0.834 ; p-value = <0.0001
- R₈ et R₁₄: -0.916 ; p-value = <0.0001
- R₈ et R₁₉: -0.753 ; p-value = <0.0001
- R₁₀ et R₁₂: 0.973 ; p-value = <0.0001
- R₁₀ et R₁₄: 0.719 ; p-value = <0.0001
- R₁₀ et R₁₉: 0.877 ; p-value = <0.0001
- R₁₂ et R₁₉: 0.910 ; p-value = <0.0001
- R₁₄ et R₂₂: 0.999 ; p-value = <0.0001

Après avoir analysé ces corrélations, nous avons éliminé 7 ratios qui sont les plus corrélés entre eux et les moins corrélés avec la variable expliquée (dans ce cas, on parle d'une faible intensité de liaison linéaire entre la variable à expliquer et le ratio concerné et donc une faible capacité à prédire le défaut en fonction de ce ratio) :

R5, R6, R7, R8, R10, R12, R14

Suite à cette analyse, notre étude portera seulement sur les 17 ratios restants.

2.2.4. Statistiques descriptives

Tableau 15 Statistiques descriptives des ratios

Ratio	Nombre d'observation	Médiane	Moyenne	Ecart type
R1	588	9.635	9.622	0.885
R2	588	0.185	0.179	0.420
R3	588	0.051	0.115	0.150
R4	588	0.159	0.269	2.298
R9	588	0.833	1.760	5.643
R11	588	0.180	0.298	0.441

R13	588	0.949	0.332	22.475
R15	588	0.081	0.039	0.507
R16	588	1.110	2.227	8.366
R17	588	0.744	0.680	0.283
R18	588	0.091	0.978	17.127
R19	588	0.001	0.007	0.104
R20	588	0.937	0.717	0.361
R21	588	9	8.680	4.693
R22	588	0.000	0.000	0.002

Section 3 : Construction de la fonction score

3.1. Analyse discriminante

L'objectif de l'analyse discriminante est de prédire l'appartenance à des groupes prédéfinis (0 et 1 dans notre cas), à partir de certaines variables prédictives (les ratios financiers et comportementaux). Dans sa mise en œuvre, l'analyse discriminante va rechercher les combinaisons linéaires des variables discriminantes qui permettent de séparer le mieux possible les groupes considérés, c'est-à-dire avec un pouvoir discriminant maximum.

L'échantillon contiendra 50% d'entreprises saines et 50% d'entreprises défailtantes.

Selon **Ludovic LE MOAL(2002)**, l'analyse discriminante se déroule en trois étapes :

- Vérification de l'existence de différences entre les groupes,
- Validation de l'étude,
- Estimation des coefficients de la fonction discriminante et évaluer la qualité de la représentation.

3.1.1. Vérification de l'existence de différences entre les sous-groupes

Cette première analyse permet de déterminer quelles sont les variables qui sont les plus discriminantes entre les groupes. Ceci peut être réalisé grâce à trois indicateurs : la moyenne ou la variance, le test de Fisher et le lambda de Wilks. Ils s'interprètent de la façon suivante :

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Tableau 16 Vérification de l'existence de différences entre les sous-groupes
Source : Ludovic LE MOAL (2002)

	En cas d'influence	En absence d'influence
Moyenne ou variance	Différence	Similitude
Test de F	F est élevé et sa significativité tend vers 0.	F es faible et sa significativité ≥ 0.01 ou 0.05.
Lambda de Wilks	≤ 0.9	Tend vers 1

Rappelons que pour les deux groupes, la statistique F est le rapport de la variance intergroupe sur la variance intragroupe et le lambda de Wilks, mesurant le pouvoir discriminant de chacun des ratios, est une fonction inverse de la statistique F, soit :

$$\Delta = \frac{1}{[1 + \frac{F}{n-2}]}$$

Avec n est le nombre d'observation dans le modèle.

Ces deux indicateurs peuvent être observés grâce aux tests d'égalité des moyennes des groupes (voir annexe 1).

Ce tableau montre la pertinence du ratio R2 (autonomie financière) ayant un pouvoir discriminant le plus élevé (F de Fisher 319.752), ce qui montre que cette variable influence bien la situation de l'entreprise et permet par la suite de différencier entre les deux catégories d'entreprises (saines et défaillantes).

Pour mieux révéler les ratios les plus pertinents dans l'explication et la prévision de la défaillance des entreprises, nous avons eu recours à la procédure pas à pas (*Stepwise Discriminant Analysis*) qui fournit un moyen économique, informatiquement parlant, pour optimiser séquentiellement un choix plus limité des ratios dont nous disposons.

Cette méthode consiste à sélectionner progressivement la meilleure variable, puis les deux meilleurs ratios, jusqu'à ce que l'on juge que l'ajout de nouvelles variables n'apporte plus d'information nouvelle pertinente.

Au bout du septième pas (voir annexe 2), le programme arrive à discriminer les

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

982 entreprises de l'échantillon d'apprentissage et il a retenu 7 ratios qui vont être inclus dans la fonction score à savoir : **R2, R3, R9, R11, R15, R20 et R21**.

La disposition des différents ratios dans la fonction n'est pas aléatoire, chaque ratio qui précède l'autre signifie qu'il a été retenu un pas avant et qu'il est plus discriminant que le suivant. Avant l'analyse de ces ratios, une description de leur pouvoir discriminant s'est avérée nécessaire afin de dégager l'ordre d'importance et de discrimination entre les deux groupes d'entreprises.

Selon **Altman(1968)**, si Z désigne la fonction de discrimination entre les deux groupes d'entreprises, a représente le coefficient de pondération du ratio R dans la fonction de discrimination et δ désigne l'écart type du ratio R pour l'ensemble des entreprises de l'échantillon.

Le pouvoir discriminant du ratio R dans la fonction Z est mesuré par le coefficient de pondération de celui-ci multiplié par son écart-type et le pourcentage P de discrimination du ratio R_j dans la fonction Z est défini par le rapport :

$$P = \frac{\delta_j^2 \cdot a_j^2}{\sum \delta_j^2 \cdot a_j^2}$$

Le tableau suivant montre le pouvoir discriminant de chaque ratio dans la fonction score :

Tableau 17 Le pouvoir discriminant des ratios

Ratios (R_j)	Coefficient (a_j)	Ecart-type (δ_j)	Pouvoir discriminant	% de discrimination
R2=Capitaux propres/Total passif	1.032	0.586	0.605	25.02%
R15= FR/ Total actif	0.817	0.694	0.566	22%
R11 =Clients et comptes rattachés/ Chiffre d'affaires	-0.931	0.592	-0.551	20.78%
R21 = Ancienneté de la relation	0.094	4.764	0.447	13.72%
R3= DLMT/ Total passif	-2.415	0.15	-0.362	8.98%
R20 = Nombre de jours débiteurs	-1.031	0.339	-0.349	8.36%
R9 = Chiffre d'affaires / Total actif	0.029	4.412	0.128	1.12%

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Le tableau ci-dessus montre la participation de 7 ratios dans la discrimination entre les deux groupes. Nous remarquons que le ratio R2 (autonomie financière), a été sélectionné dans les premiers pas avec un pouvoir de l'ordre de 25%, ce qui prouve le rôle primordial joué par l'autonomie financière dans la différenciation entre les deux catégories d'entreprises. Cette dernière participe ainsi avec une pondération positive dans la fonction discriminante. En effet, l'accroissement de ce ratio signifie que l'entreprise a un part de financement assez important dans l'ensemble des financements de l'entreprise. Dans ce cas, nous pouvons affirmer sa bonne santé financière d'où la forte participation dans la fonction de discrimination entre les deux catégories d'entreprises.

La liquidité (R15) a été sélectionnée aussi dans les premiers pas avec une pondération positive dans la fonction discriminante. Ce constat rejoint ceux de **Back et al (1996)** et **Chralambous et al(2000)** qui estiment de leur tour que la mesure pertinente de la défaillance un, deux et trois ans avant, est la liquidité puisque, si elle s'avère insuffisante, elle génère au fil du temps, l'incapacité de l'entreprise à régler ses créanciers, ce qui provoque une baisse évidente de sa rentabilité et une augmentation de son risque de défaut. Ceci explique bien le pouvoir discriminant de R15 obtenu ci-dessus.

Le ratio R11 (Délai de paiement accordé à la clientèle) occupe la troisième place avec un pouvoir discriminant de l'ordre de 20.78%, ce dernier participe négativement dans la fonction de discriminante. Ceci peut être expliqué par le fait que le recours à ce type de financement est trop risqué dans la mesure où l'entreprise a peu de contrôle sur cette variable. Cela a été confirmé par **Dietsch (1997)** : En accordant un délai de paiement à un client, toute entreprise s'expose au risque de subir les effets d'un comportement de « mauvais payeur » de celui-ci, ce qui engendre un allongement de la durée du crédit et une augmentation des coûts de surveillance.

De plus, le gonflement des créances traduit une augmentation du besoin en fonds de roulement et des difficultés de paiement des clients notamment des créances douteuses qui forment des risques financiers énormes pour les entreprises.

L'apparition du ratio R22 (ancienneté de la relation) dans la discrimination entre les deux groupes d'entreprises a été accompagnée par un pouvoir discriminant de l'ordre de 13.7%. Ce dernier participe avec un signe positif dans la fonction discriminante : Une relation durable entre l'entreprise et sa banque peut être interprété comme un signe de la bonne santé financière de la firme ce qui réduit par la suite sa vulnérabilité et donc sa probabilité de défaillance.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Les ratios d'endettement à moyen et long terme (R3) ainsi que le ratio R21 (nombre de jours débiteurs) sont déclarés significatifs. Toutefois, ils participent négativement dans la constitution de la fonction discriminante avec un pouvoir discriminant respectivement de 8.98% et 8.36%. En effet, selon **Finet (2001)**, l'octroi de la dette résulte des difficultés subies par l'entreprise, il devient un élément de sa fragilisation des performances d'exploitation, de son capital comme de sa réputation.

Dans le même sens des idées, l'augmentation du nombre de jours débiteurs renseigne sur les difficultés financières supportées par l'entreprise dans le sens où elle se déclare incapable de rembourser à une échéance prédéterminée son engagement.

Arrivant au dernier pas, le ratio R9 (rotation de l'actif) participe dans la fonction de discriminante avec un pouvoir discriminant faible. Ce dernier renseigne sur l'efficacité d'exploitation de l'actif. Et par conséquent, plus il augmente plus l'entreprise est capable de générer des revenus en immobilisant ses actifs donc le signe positif dans la constitution de la fonction de discrimination.

Nous avons constaté que les quatre premiers ratios qui possèdent le pouvoir les plus élevés (R2, R15, R11 et R21) expliquent en moyenne près de 80% du pouvoir total de discrimination. Ce qui dévoile que la discrimination entre les deux types d'entreprises est fortement liée à l'importance de l'autonomie financière, la liquidité et la durée de crédit clients et l'ancienneté de la relation banque-entreprise.

Pour conclure, l'analyse discriminante arrive à distinguer entre les deux groupes d'entreprises en accordant une importance inégale aux ratios financiers. De ce fait, la faillite des entreprises est en partie, liée à l'autonomie financière, à la liquidité, à la durée de crédit clients, à l'ancienneté de la relation banque-entreprise, à l'endettement, au nombre de jours débiteurs et à la rotation de l'actif.

1.3.2. Validation de l'étude

La validité de l'analyse discriminante peut être déterminée par quatre indicateurs :

- **Le test de Box**

Le test de Box permet de vérifier l'homogénéité des matrices de covariances. Le résultat de ce dernier est récapitulé dans le tableau ci-dessous :

Tableau 18 Résultat du Test de Box

F	M de Box	3648.568
	Approximativement	129.324

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

	Signification	<0.0001
--	---------------	---------

Pour que l'analyse soit valide, le M doit être le plus élevé possible et la significativité de F doit tendre vers 0. Ce qui est vérifié dans notre étude d'où le rejet de l'hypothèse nulle d'égalité des matrices de variance-covariance intragroupe.

▪ La corrélation globale

Une autre mesure du pouvoir discriminant de la fonction linéaire discriminante est le coefficient de corrélation canonique. Elle se mesure à travers le tableau « Valeurs propres ». Plus elle est proche de 1, plus le modèle est meilleur. Dans notre cas elle est de **0.703** ce qui atteste la performance du modèle.

Également, les valeurs propres associées aux fonctions linéaires discriminantes permettent de juger le pouvoir discriminant de ces fonctions, en effet, chaque valeur propre μ_h de rang h est égale à la variance interclasse de la fonction discriminante de même rang. Dans notre cas, la fonction discriminante explique 100% de la variabilité interclasse.

Tableau 19 Valeurs propres

Valeur propre	% de la variance	Corrélation canonique
0.975	100	0.703

▪ Le Lambda de Wilks

La statistique associée au lambda de Wilks suit une distribution Khi deux à $p(K-1)=7$ degrés de libertés sous l'hypothèse nulle d'égalité des moyennes des $K=2$ groupes pour les $p=8$ variables introduites dans le modèle. Etant donnée que la p-value est inférieure à 5%, on doit alors rejeter l'hypothèse nulle.

Le modèle est jugé bon lorsque la valeur du Lambda Wilks est faible et sa significativité tend vers 0 car plus il se rapproche de 0, plus les barycentres sont distincts. Ceci est bien vérifié dans notre modèle, donc le modèle est bon.

Tableau 20 Le Lambda de Wilks

Lambda de Wilks	Khi deux	Ddl	Signification
0.506	664.723	7	<0.0001

3.1.3. Estimation des coefficients de la fonction discriminante

Au final, il est possible d'estimer les coefficients de la fonction discriminante et déduire par la suite la fonction score :

$$Z = 0.874 + 1.032R_2 - 2.415R_3 + 0.029R_9 - 0.931R_{11} + 0.817R_{15} - 1.031R_{20} + 0.094R_{21}$$

3.1.4. Qualité de la représentation

Apprécier la qualité de la représentation revient à s'assurer que la fonction discriminante classe bien les entreprises en sous-groupes. Pour ce faire, nous avons recouru à l'analyse de la matrice de confusion qui regroupe les entreprises bien classées et les mal classées. Il en sort que :

Tableau 21 Le taux de bon classement de l'analyse discriminante

	0	1	Total	% correct
0	434	57	491	88.39%
1	111	380	491	77.39%
Total	545	437	982	82.89%

Le taux de bon classement global est de 82.89% avec 88.39% pour les entreprises saines contre 77.39% pour les entreprises en défaut. Concernant l'erreur de type I³⁷, elle est de l'ordre de 11.61% contre 22.60% pour l'erreur de type II³⁸.

La courbe ROC (Receiver Operating Characteristic) est une représentation graphique de la relation existante entre la sensibilité et la spécificité d'un test, calculée pour toutes les valeurs seuils possibles. Elle permet de comparer les performances diagnostiques de plusieurs tests à l'aide de l'évaluation de l'aire sous la courbe appelé AUC (Area Under the Curve). Celle-ci informe sur la probabilité que le résultat du test permette de poser le diagnostic correct. Autrement dit, le test est parfaitement discriminant lorsque l'aire sous la courbe (AUC) vaut 1. La courbe des points (1-spécificité, sensibilité) est la courbe ROC. On désigne par sensibilité la proportion d'événements positifs bien classés. Toutefois 1- spécificité correspond à la proportion d'événements positifs mal classés. Le meilleur modèle est celui qui permet de capturer le plus possible de vrais positifs avec le moins possible de faux positifs.

Le tableau suivant explique la nature de la discrimination selon l'aire sous la courbe (AUC) :

Tableau 22 Le pouvoir discriminant du modèle par le critère AUC

³⁷ Le pourcentage des entreprises saines considérées comme en défaut : Taux de faux positifs

³⁸ Le pourcentage des entreprises en défaut considérées comme saines : Taux de faux négatifs

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Source : Cours Michael Genin, Université de Lille

AUC	Discrimination
0.5	Nulle
[0.7-0.8]	Acceptable
[0.8-0.9]	Excellente
>0.9	Exceptionnelle

Si l'aire sous la courbe(AUC) est égale à 0.5 alors le modèle classe de manière complètement aléatoire les observations. Toutefois si ce dernier est supérieur à 0.9, le modèle est jugé très bon voire trop bon.

La première bissectrice correspond à ce que l'on obtiendrait avec un modèle aléatoire de Bernoulli avec une probabilité de réponse égale à celle observée sur l'échantillon étudiée. Un modèle proche de cette droite est donc inefficace puisqu'il n'est pas meilleur qu'un simple tirage au hasard. Dans notre cas, la courbe ROC se présente comme suit :

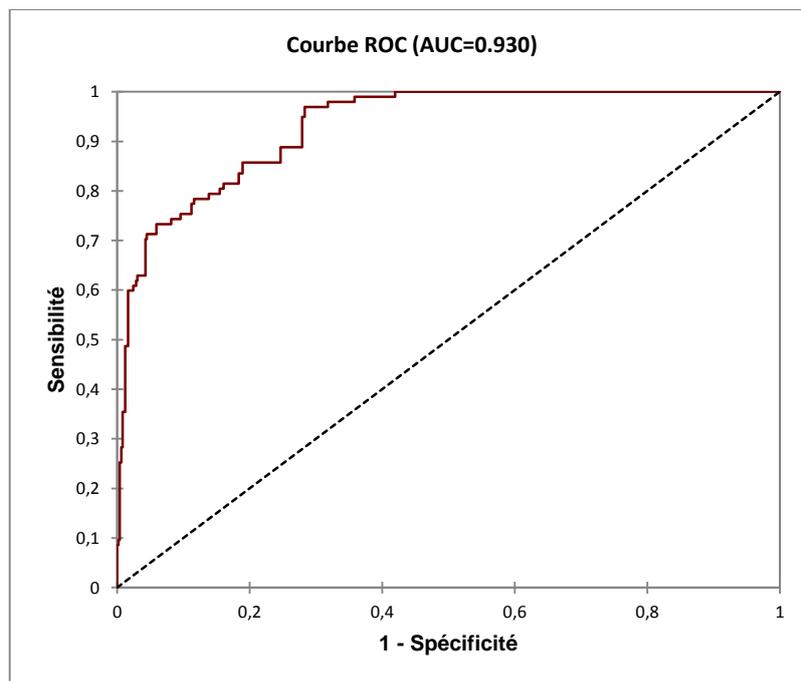


Figure 14 La courbe ROC analyse discriminante

La courbe est au dessus de la bissectrice et l'AUC 0.930 indique une discrimination parfaite.

3.1.4. Validation du modèle

L'affectation aux groupes est faite en fonction des centroides de ces derniers, c'est-à-dire, par comparaison avec un score discriminant moyen pour chaque groupe. Ce score moyen est

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

calculé à partir de la fonction discriminante, où l'on remplace les valeurs individuelles par les moyens des variables indépendantes pour le groupe dont on s'occupe. Les scores discriminants moyens pour les deux groupes sont donnés ainsi :

Tableau 23 Fonction aux barycentres des groupes

Appartenance	Scores moyens
0 (entreprises saines)	0.987
1 (entreprises défailtantes)	-0.987

Chaque score individuel discriminant est ensuite comparé aux deux scores moyens et affecté au groupe dont-il est le plus proche. Mais la question qui se pose est la suivante : *à partir de quel score peut-on affecter les individus au groupe 0 et non pas au groupe 1 ?* .Pour ce faire, nous devons déterminer un score qui joue le rôle de frontière entre les deux groupes. Si ces groupes sont de dimensions égales, le score critique est égal à la moyenne des moyennes des scores des groupes. Dans notre cas, ce score est égale à $0.987 - 0.987 / 2 = 0$.

Cette situation nous emmène à constater que chaque entreprise peut se classer selon la règle de décision suivante :

Tableau 24 Règle de décision

Valeur du score	Affectation selon notre modèle
$Z \geq 0$	Groupe 0 (Saine)
$Z < 0$	Groupe 1 (Défaillante)

Afin de valider la performance de notre modèle, nous avons pris un échantillon de test de l'ordre de 20% de l'échantillon d'apprentissage. Les résultats sont présentés dans le tableau ci-dessous :

Tableau 25 La validation de l'analyse discriminante

	0	1	Total	% correct
0	76	23	99	76.76 %
1	31	68	99	68.68%
Total	107	91	198	72.72%

Le taux de bon classement de l'échantillon de validation est de l'ordre de 72.72%. Concernant l'erreur de type I est 23.23% et celui de type II est 31.31%.

3.2. La régression logistique

3.2.2. Résultats des traitements

Après avoir écarté les ratios qui posent le problème de corrélation (multicolinéarité), nous avons introduits le reste des ratios afin d'expliquer la défaillance des firmes. Nous avons procédé à la modélisation avec 15 ratios, cependant 11 ratios s'avèrent significatifs et ont été par la suite retenu dans le modèle final. (Voir annexe 1).

Au final, Le modèle a été construit avec 10 ratios significatifs, les résultats trouvés sont les suivants :

Tableau 26 Résultat de la régression logistique

Ratios	Signe attendu	β	Probabilité
Constante		15.131	<0.0001
R1= log (actifs)	Négatif	-1.706	<0.0001
R2 = Capitaux propres/ Total passif	Négatif	-3.168	<0.0001
R3 = DLMT/ Total passif	Positif	6.940	<0.0001
R9 = Chiffre d'affaires /total actif	Négatif	-0.746	0.032
R11 = Clients et comptes rattachés/ Chiffre d'affaires	Positif	2.815	<0.0001
R15 = FR / Total actif	Négatif	-1.730	<0.0001
R16 = Actifs courants/ Passifs courants	Négatif	-0.225	0.011
R19= Mouvements créditeurs/ Chiffre d'affaires	Négatif	-4142.985	<0.0001
R20 = Nombre de jours débiteurs	Positif	3.096	0.001
R21 = Ancienneté de la relation	Négatif	-0.235	<0.0001
R22= Solde des impayés/ Chiffre d'affaires	Positif	881.509	0.002

Premièrement, d'un point de vue statistique, les résultats sont satisfaisants dans le sens où tous les coefficients sont significatifs à 1% (sauf pour le R9 et R16 qui sont significatif à 5%).

Deuxièmement, le signe attendu de tous les ratios est confirmé. Il est cohérent avec l'analyse financière classique et donc avec la logique économique. En effet, l'autonomie financière(R2), la liquidité 1 (R15) et la liquidité 2 (R16) contribuent négativement et celui le ratio d'endettement à MLT(R3) et le délai de paiement accordé à la clientèle (R11) positivement à la formation de la probabilité de défaillance (lorsque ce ratio est élevé, la part des dettes clients non réglées par rapport au chiffre d'affaires s'élève aussi, dans ce cas le

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

risque auquel s'expose la firme augmente dans le sens où elle va faire face à des problèmes de liquidité si ses clients n'honorent pas leurs engagements au délai communiqué.

Concernant le ratio R2, plus l'entreprise possède une certaine autonomie financière, plus elle est indépendante vis-à-vis de ses créanciers. Autrement dit, l'autonomie financière constitue une marge de sécurité à la disposition de l'entreprise, plus elle s'élève plus sa situation financière est bonne et plus sa probabilité de défaillance est faible. En outre, R16 exprime le degré de liquidité de l'entreprise : plus il est important, plus l'entreprise dispose suffisamment de liquidité pour honorer ses engagements à court terme donc moins exposée au risque de faillite.

Selon Altman(1968), R15 (ratio de liquidité 1) est le ratio le plus significatif pour prédire la faillite de l'entreprise par rapport aux autres ratios de liquidité. Donc plus ce ratio augmente, moins la firme est menacée de faillite.

Troisièmement, l'augmentation du ratio R9 (rotation de l'actif) indique que l'entreprise est capable de générer un chiffre d'affaires en utilisant ses actifs d'où la contribution négative à la formation de la probabilité de défaillance. La taille de l'entreprise(R1) s'avère significative aussi: Ce constat est conforme à celui de **Krichène et al(2010)** qui ont utilisé la variable « taille » afin de discriminer entre les firmes saines et défaillantes, ils ont trouvé que cette dernière permet d'accroître la probabilité de non défaillance.

Finalement, on remarque que toutes les variables comportementales sont intégrées dans notre modèle, autrement dit, la totalité de ces ratios sont significatifs .En effet, l'augmentation du ratio R19 (Mouvements créditeurs/Chiffre d'affaires) implique une diminution de la probabilité de défaut dans la mesure où l'entreprise réalise des mouvements créditeurs assez importants par rapport à son chiffre d'affaires ce qui forme une indication sur la bonne santé financière de cette dernière.

En outre, le ratio R20 (nombre de jours débiteurs) contribue positivement à la formation de la probabilité de défaillance ce qui constitue une source de risque dans la logique où l'accroissement du nombre de jours débiteurs signale un problème de trésorerie au niveau de l'entreprise (incapacité à honorer ses engagements) et donc elle demeure de plus en plus susceptible de se retrouver en situation de défaut. Le signe négatif du ratio R21 (ancienneté de la relation) signifie qu'un accroissement de la durée de la relation banque-entreprise renseigne sur la bonne situation de l'entreprise et réduit par la suite sa vulnérabilité financière. Toutefois, le ratio R22 (Solde des impayés/Chiffre d'affaires) contribue positivement à la

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

formation de la probabilité de défaillance, ceci est attendu puisque si l'entreprise a un solde des impayés important par rapport à son chiffre d'affaires, elle va être exposée à des difficultés financières ce qui augmente sa vulnérabilité et par la suite sa probabilité de défaut.

Ainsi, la fonction score obtenue est la suivante :

$$Z = 15.131 - 1.706R1 - 3.168R2 + 6.940R3 - 0.746R9 + 2.815R11 - 1.730R15 - 0.225R16 - 4142.985R19 + 3.096R20 - 0.235R21 + 881.509R22$$

En se basant sur cette fonction on peut estimer la probabilité de défaut des entreprises :

$$P(Y_i = \text{défaut}) =$$

$$\frac{1}{1 + \exp(-(15.131 - 1.706R1 - 3.168R2 + 6.940R3 - 0.746R9 + 2.815R11 - 1.730R15 - 0.225R16 - 4142.985R19 + 3.096R20 - 0.235R21 + 881.509R22))}$$

La régression logistique nous permet de mesurer directement le surcroît de risque associé à un facteur explicatif binaire par le biais de l'odds ratio (ou le rapport des cotes). C'est l'exponentiel du coefficient β estimé par le modèle. En pratique, si ce dernier est supérieur à 1, on parle de facteur de risque. Toutefois s'il est inférieur à 1, on parle d'un facteur protecteur. Les résultats obtenus dans notre cas sont les suivants :

Tableau 27 Odds ratios

Ratio	Exp(β)
R1	0.182
R2	0.042
R3	1032.902
R9	0.474
R11	16.686
R15	0.177
R16	0.799
R19	0
R20	22.116
R21	0.791
R22	$+\infty$

On remarque que les ratios R3, R11, R20 et R22 possèdent un odd ratio supérieur à 1, cela signifie que ces derniers présentent des facteurs de risque ce qui confirme notre analyse

économique. Autrement dit, si R3, R11, R20 ou R22 augmente d'une unité, le risque de tomber en défaut pour l'entreprise augmente.

En outre, R3 et R22 ont l'odd ratio le plus élevé, c.à.d. les firmes ayant un endettement à MLT, le nombre de jours débiteurs et le solde des impayés par rapport au chiffre d'affaires les plus élevés sont plus risqués de tomber en faillite par rapport aux autres firmes.

3.2.2. Significativité globale du modèle

➤ Test du rapport de vraisemblance

Le test de rapport de vraisemblance consiste à confronter deux déviations, celle du modèle étudié et du modèle trivial **Rakotomalala (2015)**. La statistique du test se calcule comme suit : $LR = 2 \log(L(\beta)) - 2 \log(L(0))$. Elle suit une loi khi-deux de degré de liberté équivalent au nombre de variables introduits dans le modèle. Il s'agit de comparer la valeur critique à celle calculée (LR) ou de vérifier la significativité de la probabilité à un seuil fixé, généralement 5%.

Avec :

$-2 L(\beta) = C$ 'est la valeur de $-2 \cdot \ln$ (vraisemblance) lorsque tous les paramètres sont estimés.

$-2 L(0) = C$ 'est la valeur de $-2 \cdot \ln$ (vraisemblance) lorsque tous les paramètres égalent à zéro sauf la constante.

Tableau 28 Test du rapport de vraisemblance

Statistique	DDL	Khi ²	Probabilité
LR= -2log (vraisemblance)	11	348.167	<0.0001

Le modèle est globalement significatif (probabilité <0.05).

➤ Le coefficient de détermination de Mc Fadden

Le R^2 de Mc Fadden est le plus adapté à la régression logistique **Rakotomalala (2015)**, il est le plus proche conceptuellement du coefficient de détermination de la régression linéaire multiple et admet une interprétation similaire : Plus il s'approche de 1, plus le pouvoir explicatif du modèle est bon. Ce dernier se calcule comme suit :

$$R^2 = 1 - \frac{L(\beta)}{L(0)}$$

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Dans notre modèle il est égal à 66.1% (supérieur à 50%). Autrement dit, le pouvoir explicatif des ratios introduits (c'est-à-dire la probabilité de défaillance expliquée) est de l'ordre de 66.1%

Il existe autres indicateurs qui peuvent être aussi utilisés comme le **R² de Cox et Snell(1989)** et le **R² de Nagelkerke(1991)**. Ils permettent d'expliquer le pourcentage de la variable dépendante expliquée par les ratios retenus. Le premier n'atteint jamais le maximum théorique de 1 et varie en fonction de la taille de l'échantillon. Le second est une modification du premier afin d'obtenir une valeur théorique plus près de 1 et donc plus près de la réalité. Ils représentent un estimé de la variance expliquée par le modèle. Plus leur valeur est élevée, plus la probabilité prédite par le modèle s'approche de la valeur observée.

Dans notre modèle, ces indicateurs prennent les valeurs suivantes :

Tableau 29 : Les tests de Cox et Snell et Nagelkerke

R² de Cox et Snell	R² de Nagelkerke
44.7%	75.5%

Les R2 de Nagelkerke est supérieur à 50% ce qui indique la bonne qualité prédictive du modèle.

➤ **Le taux de bon classement**

La matrice de classement de la régression logistique (matrice de confusion) prend la forme suivante :

Tableau 30 Le taux de bon classement de la régression logistique

	0	1	Total	% correct
0	477	14	491	97.15%
1	22	75	97	77.32%
Total	499	89	588	93.88%

Le taux de bon classement global est de 93.88% avec 97.15% pour les entreprises saines contre 77.32% pour les entreprises en défaut. Concernant l'erreur de type I, elle est de l'ordre de 2.585% contre 22.68% pour l'erreur de type II.

➤ **Le test de Hosmer – Lemeshow**

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Le test de Hosmer-Lemeshow est une mesure d'adéquation dans la mesure où il compare les probabilités prédites dans le modèle à celles observées dans l'échantillon. On cherche à avoir un modèle qui minimise la distance entre ces valeurs.

La statistique de ce test est basée sur le regroupement des individus en k classes de taille presque identiques (10 classes dans cette étude). Elle suit une loi Khi-deux à (k-2) ddl qui teste :

H0 : Les probabilités théoriques sont proches de celles observées (modèle calibré)

H1 : Les probabilités théoriques sont différentes de celles observées (modèle non calibré)

Le modèle est calibré si on ne rejette pas H₀, autrement dit lorsque la probabilité critique du test (p-value) est plus grande que le risque choisi (p-value > 0.05).

Les résultats de ce test sont les suivants :

Tableau 31 Test de Hosmer -Lemeshow

H-L Statistic	DDL	Probabilité
4.147	8	0.844

Notre modèle est validé puisque la p-value est supérieure au risque prédéfini (5%) comme on a évoqué précédemment ce qui implique une adéquation du modèle avec les données.

3.2.3. Significativité individuelle des ratios

Le test utilisé afin de juger la significativité individuelle des ratios est celui de Wald. Ce dernier est similaire au test de Student de la régression linéaire. La statistique de Wald se calcule ainsi : $\frac{\beta_i^2}{\sigma_i^2}$, elle est comparée à une loi de Khi deux à 1 degré de liberté. Il est par ailleurs possible de vérifier directement si la probabilité est inférieure à un seuil fixé (5% dans notre cas).

Avec : β_i : Le coefficient estimé ; σ_i : C'est l'écart type

Ce test a été effectué antérieurement afin de sélectionner les ratios significatifs à intégrer dans le modèle final. (Voir annexe 1)

3.2.4. La Courbe ROC et l'indicateur AUC

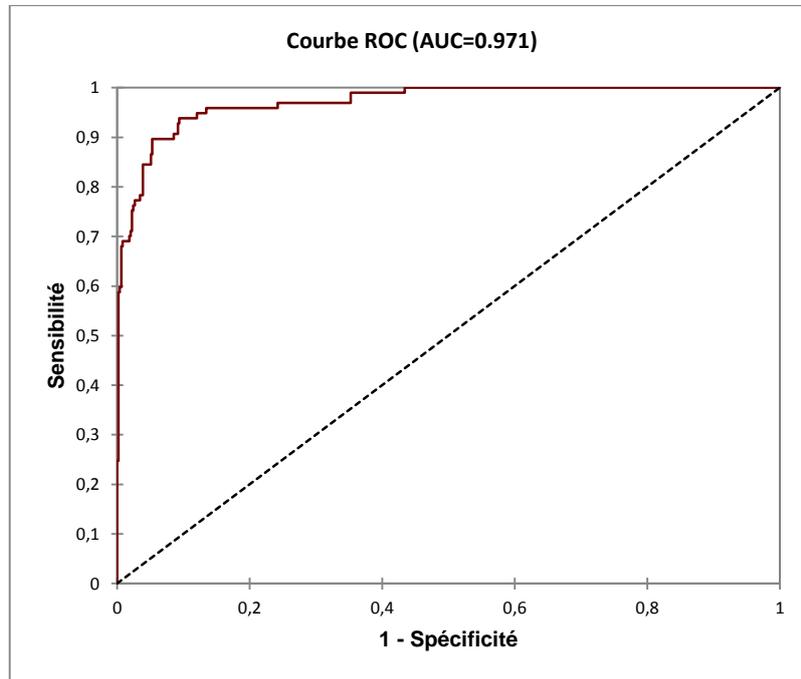


Figure 15 Courbe ROC régression logistique

L'AUC est de 0.971 et la courbe ROC est au dessus de la première bissectrice ce qui implique une discrimination parfaite.

➤ *Validation du modèle*

Appelée aussi phase de validation du modèle. Selon **Gueyié et al (2007)**, Cette technique consiste à laisser une partie de l'échantillon de base afin de la tester sur le modèle de base. En effet, nous avons laissé 154 entreprises dont 129 sont saines et 25 sont défailtantes.

Le test consiste à calculer les probabilités de défaut de chaque entreprise et appliquer par la suite la règle standard ($< \text{ou} <$ à 50%). Les résultats trouvés sont les suivants :

Tableau 32 La validation de la régression logistique

	0	1	Total	% correct
0	117	12	129	90.69 %
1	7	18	25	72%
Total	124	30	154	87.66%

3.3. Arbre de décision

Parmi les algorithmes d'arbres de décision, nous avons eu recours à l'arbre CHAID. Cet algorithme est considéré comme un arbre de classification, c'est-à-dire, il cherche à trouver les variables qui classent le mieux les observations dans les deux groupes.

Ce type d'arbre présente la spécificité d'être un arbre non binaire, autrement dit, un arbre dans lequel nous pouvons avoir plus de deux branches connectées à un même nœud.

Le modèle décrit sous forme d'arbre peut être transcrit par une succession de règles disjointes et facilement interprétables. La structure arborescente intègre les attributs les plus marquants selon leur degré d'influence avec la variable à discriminer (variable binaire dans notre cas : classe d'entreprises). Les variables les plus pertinentes sont donc celles qui sont proches de la racine.

En utilisant l'arbre CHAID, nous avons abouti à l'arbre ci-dessous pour l'échantillon d'apprentissage :

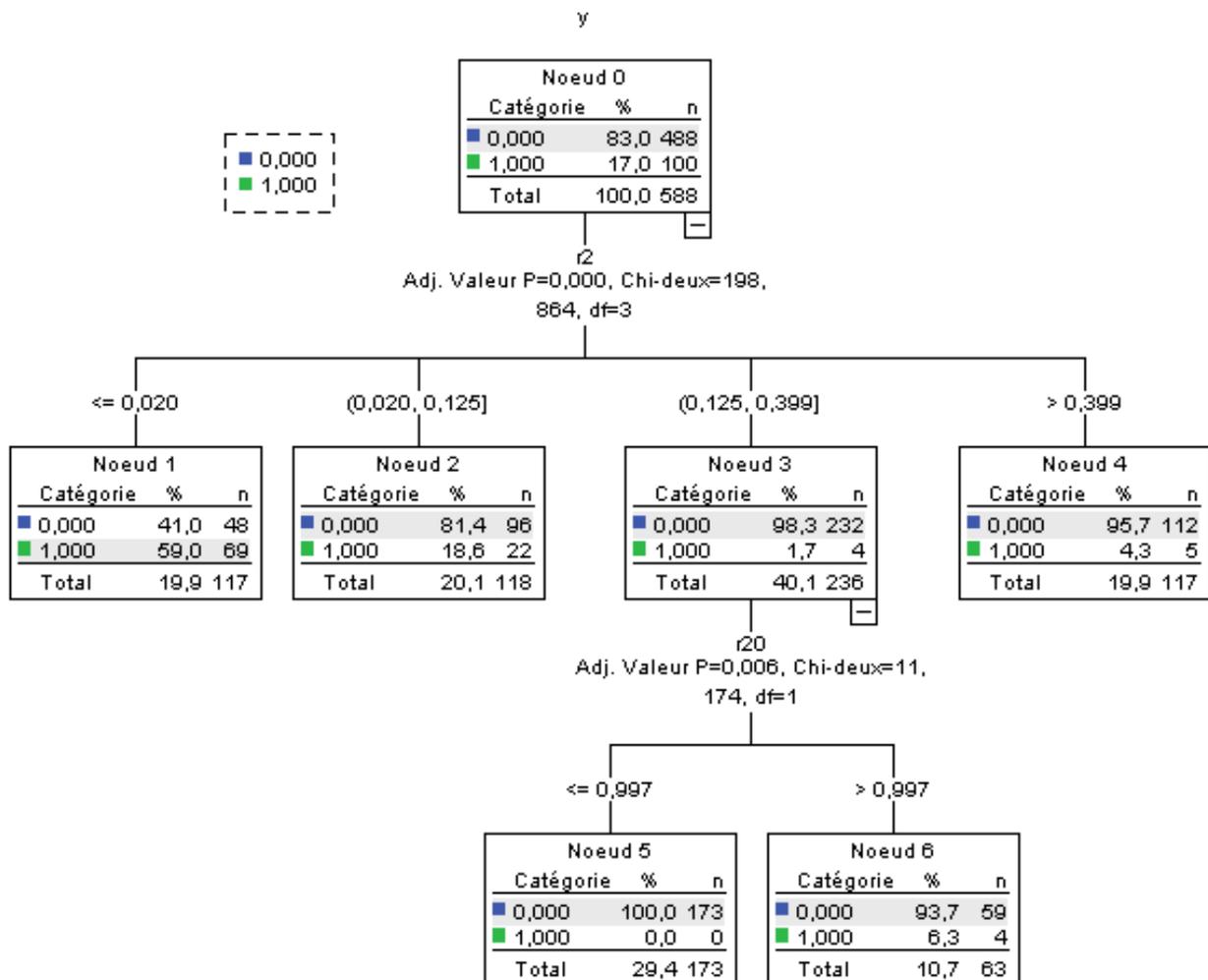


Figure 16 L'arbre de décision en utilisant le modèle CHAID

3.3.1. Analyse des résultats

En utilisant l'arbre CHAID, le ratio R2 (autonomie financière) s'avère le ratio qui discrimine le plus entre les deux catégories d'entreprises (la racine de l'arbre). En effet, les entreprises qui possèdent une autonomie financière inférieure ou égale à 0.020 sont considérées défaillantes dans 58.6% des cas alors que celles ayant ce ratio entre (0.020 ;0.125] sont saines à dans 81.35% des cas.

Les firmes ayant une autonomie financière dans l'intervalle (0.125 ;0.399] et un nombre de jours débiteurs inférieur ou égal à 0.997 sont saines à 100%. Toutefois, si ce dernier est supérieur à 0.997, elles sont saines dans 67.21% des cas, tandis que celles possédant un R2 entre 0.134 et 0.172 sont saines dans 93.65% des cas.

Au final, les entreprises ayant un $R2 > 0$ sont saines dans 95.72% des cas.

À la lumière de ces résultats, nous avons pu mettre en évidence la supériorité de l'autonomie financière par rapport aux restes des ratios en termes de discrimination entre les deux catégories d'entreprises. En effet, lorsqu'elle augmente, la probabilité de défaillance diminue et par la suite l'entreprise a plus de chance d'appartenir à la classe 0.

La variable comportemental R20 (nombre de jours débiteurs) semble être aussi significative pour la discrimination des entreprises ayant une autonomie financière dans l'intervalle (0.125 ; 0.399]

Pour conclure, l'arbre CHAID ci-dessus permet d'affirmer l'appartenance de l'entreprise à la classe 0 dans un seul cas :

- Si l'entreprise possède une autonomie financière appartenant à l'intervalle (0.0125 ; 0.399] et un nombre de jours débiteurs inférieur ou égal à 0.997.

3.3.2. Matrice de classement de l'arbre de décision

Nous avons précisé que l'échantillon d'apprentissage comporte 79.2% de l'ensemble de données contre 20.7% pour l'échantillon test. Le logiciel ajuste ces pourcentages dans le but d'obtenir les meilleurs résultats. La mesure de la performance de l'arbre CHAID repose sur la détermination du taux d'erreur de classification. Ceci est illustré par le tableau ci-dessous :

Tableau 33 : Le taux de bon classement de l'arbre CHAID

Échantillon	0	1	Total	% correct
Apprentissage				
0	440	48	488	90.2%
1	31	69	100	69%

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Global				86.6%
Test				
0				93.2%
1	123	9	132	75%
Global	5	15	20	90.8%

Le taux de bon classement global abouti par l'arbre CHAID pour l'ensemble des données de l'échantillon d'apprentissage est de l'ordre de 86.6% dont 90.2% est le taux de bon classement des entreprises saines contre 69% pour les entreprises défaillantes. L'erreur de type I est 9.83% contre 31% pour l'erreur de type II.

Concernant l'échantillon test, l'arbre permet d'assimiler 90.8% de l'ensemble des données de l'échantillon test.

Dans le tableau 30 produit par IBM SPSS Decision Tree, l'estimation du risque pour l'échantillon test vaut 0.092 ce qui signifie qu'environ une entreprise sur dix est mal classée par l'arbre. Cette estimation a diminué en passant de l'échantillon d'apprentissage à l'échantillon test puisque le taux de bon classement a augmenté de son part.

Tableau 34 : Coût d'un arbre de décision

Échantillon	Estimation	Erreur standard
Apprentissage	0.134	0.014
Test	0.092	0.023

3.4. Réseaux de neurones artificiels

Selon **Catherine Refait(2004)**, le principe des réseaux de neurones consiste en l'élaboration d'un algorithme dit d'apprentissage qui imite le traitement de l'information par le système neurologique humain. Trois sortes de neurones existent : les neurones d'entrée, les neurones de sortie et les neurones cachés. Les neurones d'entrée ont pour *inputs* les K ratios comptables et comportementales présélectionnés, les neurones de sortie ont pour *output* la variable dichotomique saine/défaillante. Concernant les neurones cachés, ils sont des neurones qui traitent l'information entre les neurones d'entrée et les neurones de sortie. L'objectif est d'aboutir, après une phase dite d'apprentissage, à la combinaison des poids d'*input* de chaque neurone qui conduit à la meilleure description de la réalité, c'est-à-dire, au meilleur classement des entreprises dans les deux groupes de firmes.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Pour une classification d'une base de données telle que la nôtre, le RNA nécessite l'emploi d'un apprentissage supervisé. Dans le but de développer notre modèle, nous devons procéder à un sous échantillonnage en construisant trois sous échantillons comme l'indique la Figure 18.

- *Le premier sous-échantillon* est celui d'apprentissage. Il doit être important en termes de nombre par rapport aux deux autres dans le sens où il fournit au RNA une base solide pour apprendre, et, sur laquelle il produit des résultats. Il occupe près de 65% de l'ensemble de notre base de données.
- *Le deuxième sous-échantillon* est celui de validation. Il représente près de 25% de l'ensemble des observations. Cet échantillon permet de tester l'aptitude du réseau à lire des données hors l'échantillon d'apprentissage.
- *Le troisième sous-échantillon* est celui de test dans la littérature, nommé sur SPSS comme l'échantillon traité. Il représente près de 10% de l'ensemble de nos données. Cet échantillon évalue la capacité de réseau à donner des résultats généralisés sans avoir eu la sortie correspondante.

Il est à noter que le logiciel lui-même ajuste ces sous-échantillons, à postériori, dans le but de réaliser les meilleurs résultats (voir annexe 4).

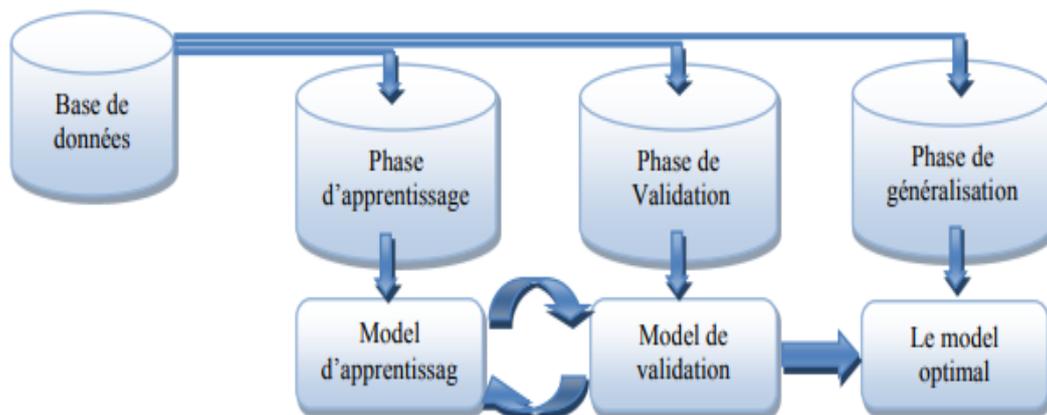


Figure 17 La généralisation d'un modèle

Source : Chapados.N(2010), « Data Mining Algorithms for Actuarial Ratemaking »,ApSTAT Technologies p22

Graphiquement, le réseau peut être schématisé ainsi par la Figure 19.

Le nombre des neurones dans la couche d'entrée est conforme au nombre des variables de notre échantillon qui est de quinze et le nombre des neurones dans la couche de sortie est de

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

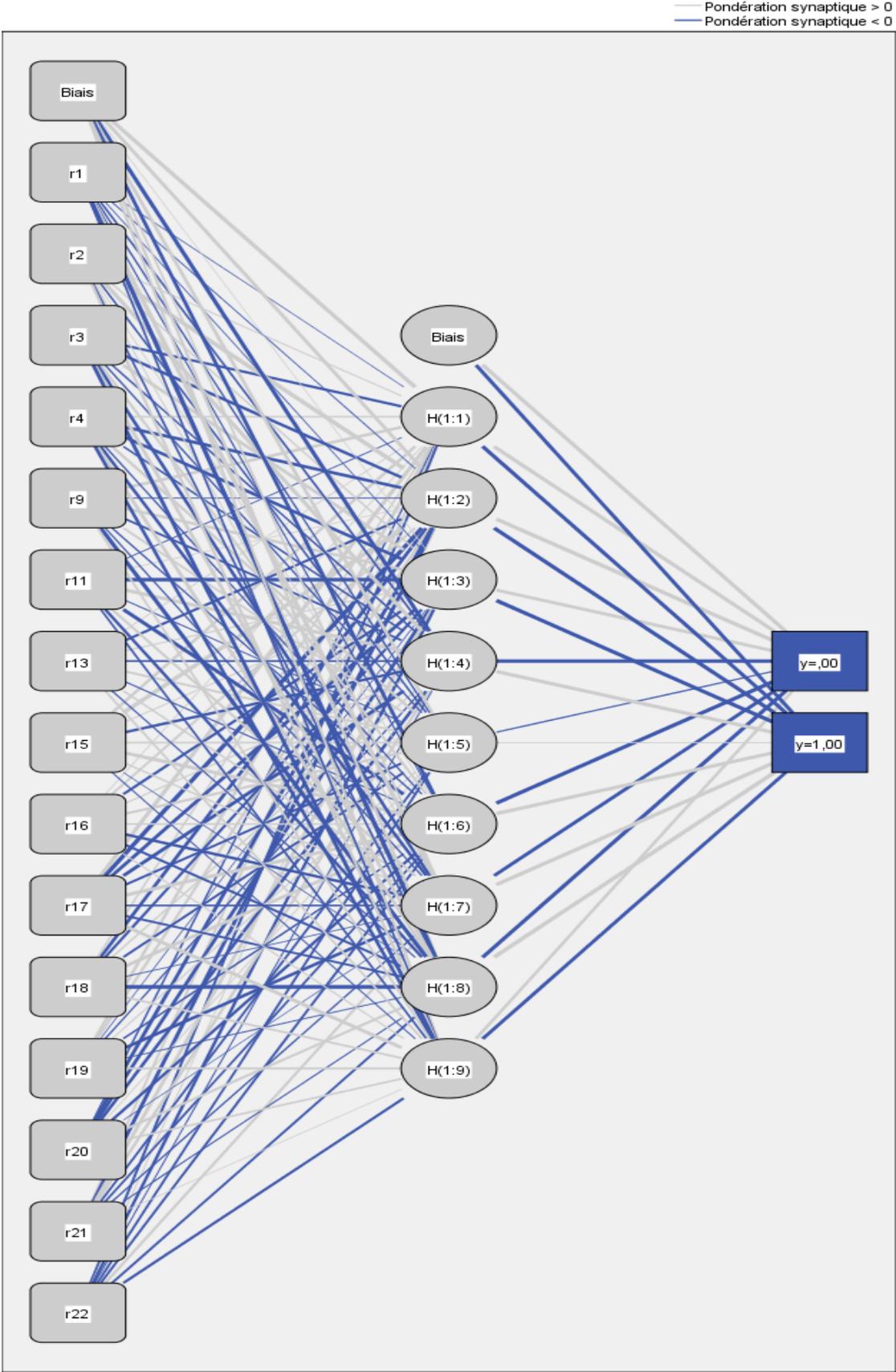
deux (0 et 1). Concernant la couche cachée, le modèle s'est porté sur neuf neurones. En parallèle, les fonctions d'activations utilisées dans la couche cachée et de sortie sont respectivement, la fonction tangente hyperbolique et la fonction soft max. De plus, l'erreur d'ajustement des poids synaptiques est calculée sur la base d'entropie croisée, une fonction généralement utilisée dans les modèles de classification. Cette dernière mesure la quantité d'incertitude dans un ensemble de données (dispersion). Elle est donnée par la formule suivante :

$$E(w) = -\sum \sum t_k^{(n)} \ln(y_k^{(n)})$$

Avec : N : la taille des données d'apprentissage ; C : Nombre de classe ; t^n : le $n^{\text{ièmes}}$ vecteur cible ; y^n : le $n^{\text{ièmes}}$ vecteurs de sortie, calculé par le réseau.

Lorsque $E(w) = 0$, l'ensemble des données est parfaitement classé.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat



Couche masquée pour la fonction d'activation : Tangente hyperbolique
Couche de sortie pour la fonction d'activation : Softmax

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

Figure 18 Le réseau de neurone artificiel propre à notre étude

3.4.1. Résultats de la classification

Arrivant à ce stade, nous allons percevoir le taux de classement réalisé par le réseau de neurones artificiel en donnant opinion sur son niveau d'efficacité à savoir s'il est performant ou pas.

Le tableau ci-dessous mentionne le temps exact pour la réalisation de cette analyse, ainsi que les taux d'erreurs :

Tableau 35 : Récapitulatif du modèle

Apprentissage	
Erreur d'entropie croisée	32.081
Pourcentage de prévisions incorrectes	3.3%
Durée d'apprentissage	0 :00 :00 :53
Test	
Erreur d'entropie croisée	21.627
Pourcentage de prévisions incorrectes	3.2%
Élément traité	
Pourcentage de prévisions incorrectes	1.7%

Afin de vérifier la solidité de notre modèle, nous observons que l'erreur est réduite de 32.081 dans la phase d'apprentissage à 21.627 dans la phase de test. Cette baisse est aussi un indicateur qui nous a amené à choisir ce modèle, car le taux d'erreur de test est un facteur déterminant dans l'architecture du modèle optimal. En outre, la durée de traitement n'a pas dépassé une seconde.

La courbe ROC, de même, corrobore notre choix puisqu'elle joue un rôle primordial pour la sélection du meilleur modèle. Cette dernière se situe en haut près du coin gauche, de plus, la zone sous la courbe (AUC) est proche de 1 ce qui atteste que notre modèle est performant.

Tableau 36 : Zone sous la courbe

	Zone
,00	,990
1,00	,990

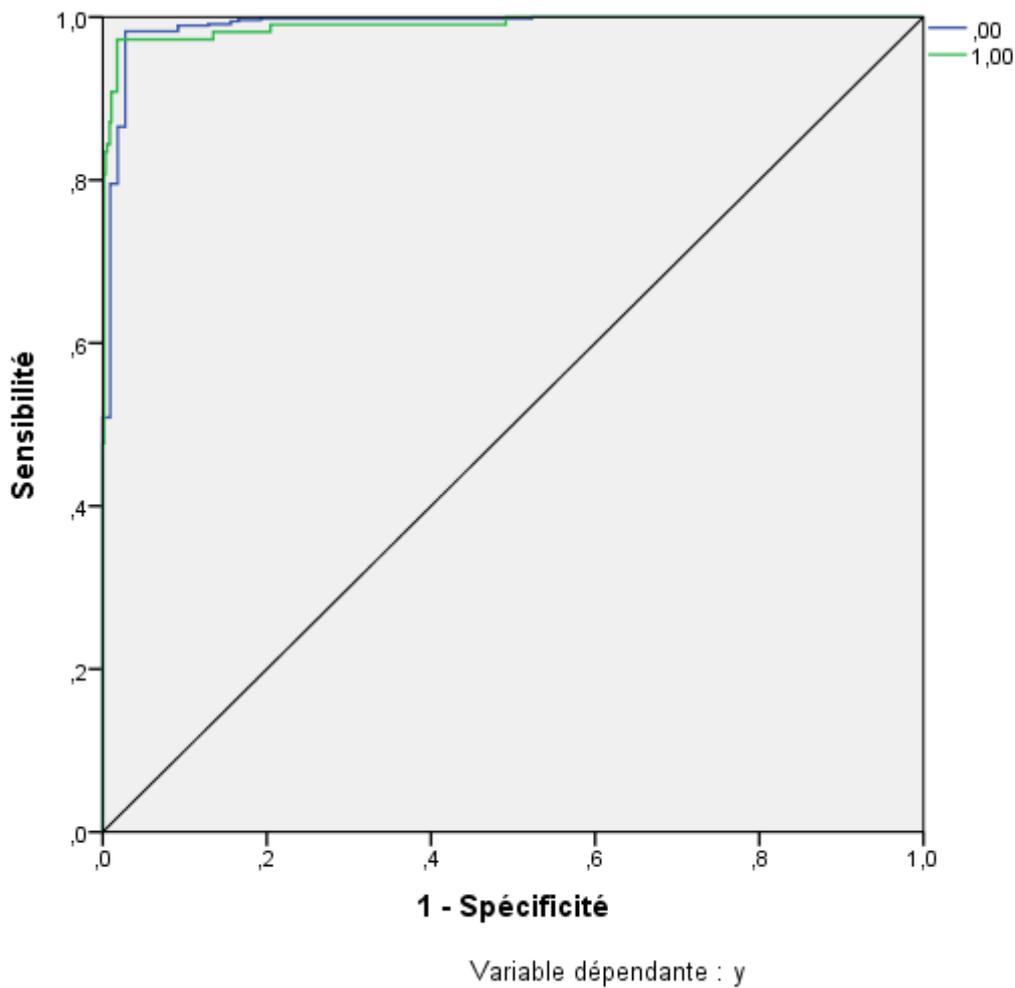


Figure 19 :La courbe ROC RNA

Les deux courbes sont au dessus de la bissectrice et l'AUC est proche de 1 (0.990) indiquent une discrimination parfaite pour les deux groupes d'entreprises.

Pour ce qui est des résultats de classification, le tableau 38 partage les proportions de classification correctes faites par les réseaux de neurones :

Tableau 37 : La classification des données

Échantillon	Observations	Prévisions		
		0	1	Pourcentage correct
Apprentissage	0	408	3	99.3%
	1	13	68	84%
	Global			96.7%
Test	0	159	2	98.8%
	1	4	24	85.7%

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

	Global			96.8%
Traité	0	48	0	100%
	1	1	10	90.9%
	Global			98.3%

Le réseau de neurones artificiel a assimilé correctement près de 96.7% des données de l'ensemble d'apprentissage dont il a détecté 99.3% des entreprises saines contre 84% des entreprises en défaut. En ce qui concerne l'échantillon test, il détecte correctement la situation de 96.8% de l'ensemble des données test. Quant à l'ensemble des données traitées à but de généralisation, il arrive à détecter la zone appropriée de 98.3% de la totalité des entreprises dans cet échantillon.

3.4.2. Analyse des résultats

Reprenons maintenant les principaux résultats dégagés par ce modèle :

Premièrement, le réseau de neurones arrive à classifier correctement 97.7% des entreprises dans l'échantillon d'apprentissage ce qui indique qu'il représente un outil très efficace pour l'accomplissement des tâches de classification des deux catégories d'entreprises : Il s'agit d'un résultat extrêmement intéressant par rapport aux résultats des recherches antérieurs qui n'ont pas abouti à un taux équivalent.

Deuxièmement, le temps imparti à l'analyse est surprenant, le réseau fournit des résultats dans un temps record, moins d'une seconde. Le facteur temps sera très bénéfique dans la mesure où il réalise des prévisions et des classifications raisonnables dans un très peu de temps voire un temps opportun.

Troisièmement, comme notre étude le montre, l'utilisation du réseau de neurones artificiels n'est pas aussi difficile qu'elle ne le paraît. Il faut juste suivre quelques commandes pour accomplir la tâche de classification avec succès.

3.5. La comparaison du pouvoir prédictif des modèles

Dans le but de comparer entre les différents modèles traités, nous avons procédé à un récapitulatif du pouvoir prédictif sur l'échantillon d'apprentissage fourni par ces derniers :

Tableau 38 : Comparaison entre les résultats des modèles utilisés

Le modèle	Saines	Défaillantes	Globale
Analyse discriminante	88.39%	77.39%	82.89%
Régression	97.15%	77.32%	93.88%

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

logistique			
Arbre de décision	90.2%	69%	86.6%
RNA	99.3%	84%	96.7%

La comparaison entre les modèles en termes de prévisibilité montre la performance de la technique neuronale par rapport aux modèles traités. En effet, le taux de bon classement, issu de l'application des réseaux de neurones artificiels, est meilleur que l'analyse discriminante, la régression logistique et l'arbre de décision. Ce résultat confirme ainsi les études empiriques déjà établies (Kerling et Podding, 1994, Oden et Sharada, 1990, Abdou et al, 2008). Toutefois, ce modèle se présente comme une boîte noire incapable d'expliquer les résultats qu'il fournit.

L'analyse discriminante se positionne la dernière en matière de classification, ceci est attendu vu l'exigence des hypothèses assez restrictives exigées par ce modèle.

Ainsi, au vue des considérations théoriques et des résultats empiriques, nous avons opté pour la construction de notre modèle de notation interne pour la régression logistique. Pour ce faire, nous avons commencé par étudier la stabilité de ce modèle sur une année après l'année de référence, nous avons eu recours à un échantillon d'entreprises dont les états financiers datent de 2016(une année après l'année de référence 2015). Cet échantillon comporte 654 entreprises dont 642 sont saines (0) et 12 sont défailtantes(1).

Les résultats obtenus sont présentés ci-dessous :

Tableau 39 La stabilité de la régression logistique

	0	1	Total	% correct
0	615	27	642	95.79 %
1	10	2	12	16.16%
Total	625	29	654	94.34%

Le taux global de bon classement est de l'ordre de 94%, nous pouvons conclure ainsi que la régression logistique est stable au cours du temps.

3.6. Transformation des probabilités de défaut en rating

Une fois le modèle construit, chaque entreprise se voit attribué un score. Il convient par la suite de construire une échelle de notation adaptée en vue de regrouper les contreparties en classes homogènes de risque.

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

On peut recourir à plusieurs méthodes afin de construire cette échelle. Nous allons par la suite procéder par le découpage iso effectif (probabilité de défaut= nombre d'entreprises en défaut dans une classe divisé par le nombre total d'entreprises dans une classe).

Tableau 40 Construction de l'échelle de notation

Classe	0	1	2
Probabilité de défaut (par effectif)	0%	48.64%	100%
Probabilité moyenne de défaut	0.0242	0.4514	0.9057

L'échelle de notation fait apparaître 3 classes distinctes :

Les entreprises qui appartiennent à la classe 0 sont considérées comme « saines » à 2.42%, ils présentent un risque très faible ce qui reflète la bonne situation financière de l'entreprise concernée. Cette dernière peut par la suite solliciter un crédit.

Quant à la deuxième classe, le taux de défaut moyen est élevé par rapport à la classe qui le précède et il est proche de 50%. Il s'agit d'une situation de prudence ce qui nous émane à proposer une analyse financière plus approfondie d'un tel dossier afin de prendre la bonne décision.

Au final, les entreprises de la deuxième classe sont défaillantes dont leur probabilité de défaut moyenne s'élève à 90.57%. Cette catégorie présente un danger très élevé, de ce fait, nous devons refuser tout dossier de crédit vérifiant ce critère.

L'affectation à une classe peut s'avérer utile notamment en matière de perception du risque et de tarification. En effet, en estimant la probabilité de défaut, la banque peut calculer la perte attendue provenant de chaque classe d'entreprise et déterminer par la suite les provisions nécessaires à mettre en place afin de combler cette perte.

Il est à noter que le taux de défaut s'élève en passant d'une classe à une autre : Les entreprises qui possèdent le score le moins élevé sont les plus saines.

La croissance du taux de défaut est illustrée par le graphe ci-dessous :

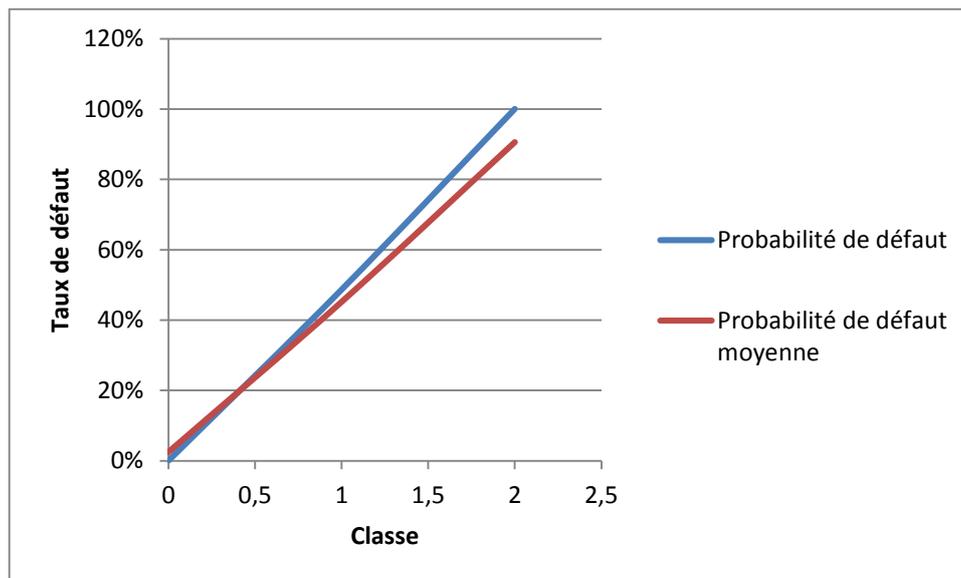


Figure 20 La croissance du taux de défaut

Conclusion

Dans ce quatrième chapitre, nous avons abordé la présentation de la structure d'accueil et l'application du modèle de crédit scoring sur les données obtenues de la banque. En effet, l'application de l'analyse discriminante, la régression logistique, les arbres de décision ainsi que les réseaux de neurones nous a permis d'analyser la défaillance des entreprises et de détecter les variables les plus significatives, c'est-à-dire celles qui expliquent au mieux la défaillance de ces dernières.

Dans un premier temps, nous avons commencé par utiliser une technique statistique classique : *L'analyse discriminante*, nous avons tiré une fonction score composée de sept ratios qui sont les plus discriminants dont cinq variables sont financières et deux comportementales à savoir : l'autonomie financière, l'endettement à moyen et long terme, la rotation de l'actif, le délai crédits clients consentis par les entreprises, la liquidité, le nombre de jours débiteurs et la durée de la relation entreprise-banque.

Dans un second temps, nous avons eu recours à *la régression logistique*, nous avons tiré aussi une fonction score comportant onze ratios qui expliquent le plus la défaillance des entreprises. Ces derniers incluent sept variables financières et toutes les variables comportementales à savoir : l'indicateur de taille, l'autonomie financière, l'endettement à moyen et long terme, la rotation de l'actif, le délai crédits clients consentis par les entreprises, les deux ratios de liquidité, les mouvements créditeurs par rapport au chiffre d'affaires, le nombre de jours

Chapitre 3 Étude empirique des méthodes du « Credit Scoring » : Application sur un échantillon de la banque de l'Habitat

débiteurs, la durée de la relation entreprise-banque et le solde des impayés par rapport au chiffre d'affaires.

En outre, nous avons appliqué la méthode de *l'arbre de décision*, nous avons ainsi trouvé que l'autonomie financière est le ratio qui découpe au mieux les entreprises selon un intervalle bien déterminé. D'où la pertinence de ce dernier à discriminer les deux groupes d'entreprises.

A l'addition de ces méthodes, nous avons utilisé une technique d'intelligence artificielle intitulée *les réseaux de neurones artificiels*. Cette dernière a abouti à une meilleure classification des entreprises dans un temps record (moins d'une seconde).

Dans une étape ultérieure, nous avons essayé de comparer la performance de tous les modèles utilisés. Premièrement, nous avons comparé leur performance sur la base du taux de bon classement de l'échantillon de base, dans ce cas, le réseau de neurones s'avère le plus performant, ceci a été approuvé par plusieurs chercheurs, en fait, ce résultat trouve ses explications principalement dans la caractéristique qu'offre un modèle non paramétrique : l'absence d'hypothèses quant à la distribution des données.

Deuxièmement, nous avons construit notre modèle de notation interne pour la régression logistique vu que cette dernière nous a offert une fonction de score stable dans le temps (un taux de bon classement est de l'ordre de 94.34%) dont l'utilisation est simple et très utile au sein d'une institution financière. En effet, en outre de l'attribution de score, ce modèle de notation interne nous a permis d'estimer la probabilité de défaut au sein de chaque classe qui sera utile ultérieurement pour le calcul des provisions.

Conclusion générale

Conclusion générale

La banque possède désormais un statut d'une entreprise commerciale dans la recherche constante du profit et simultanément, elle dispose du statut financeur de l'économie par excellence et ce, en comblant les besoins non seulement des ménages (besoins de consommation), mais aussi des entreprises (besoin d'investissement).

Partant de la relation étroite liant le risque de crédit et les entreprises, la banque doit s'assurer de leur capacité de remboursement afin de maximiser son profit et gérer ce risque de la manière la plus vigilante.

L'objectif principal de notre étude était donc de développer un modèle interne fiable et assez robuste afin de répondre aux exigences d'une tarification efficiente des crédits en fonction d'un niveau de risque bien déterminé.

Afin de construire un tel modèle, nous avons passé par différentes étapes : nous avons considéré tout d'abord une base de données comprenant 634 entreprises qui ont eu des crédits auprès de la BH.

Nous avons ainsi commencé par le choix des variables sur lesquelles va porter l'étude. Ces variables sont de type financier (dit aussi comptable) et comportemental. Ensuite, nous avons entamé la construction des différents modèles à savoir : l'analyse discriminante, la régression logistique, les arbres de décision et les réseaux de neurones artificiels.

Ces différents modèles nous ont dévoilé les ratios les plus significatifs, c'est-à-dire, ceux qui expliquent au mieux la défaillance des entreprises. A cet effet, nous avons conclu que les variables comportementales s'avèrent pertinentes dans la mesure où elles ont intervenu dans les deux fonctions score établies. En outre, nous avons remarqué la présence du ratio expliquant l'autonomie financière dans les différents modèles construits ce qui implique sa pertinence.

Après avoir construit les différents modèles, nous avons conclu que les réseaux de neurones artificiels ont abouti au meilleur taux de bon classement, soit 96.7%, suivis par la régression logistique (93.88%), les arbres de décision (86.6%) et au final, l'analyse discriminante avec un taux de 82.89%.

En s'appuyant sur ces résultats, nous avons utilisé la régression logistique pour construire notre modèle de notation interne. Ce résultat trouve ses explications principalement par la

Conclusion générale

simplicité et la clarté de cette dernière en matière d'interprétation. En outre, la régression logistique nous a permis d'aboutir à une fonction score à travers laquelle, nous avons pu segmenter les différentes entreprises dans des différentes classes de risques. Cette méthode est stable aussi dans le temps : elle a engendré un taux de bon classement de l'ordre de 94% une année après celle de l'étude.

A la lumière de cette conclusion, la régression logistique est considérée comme un modèle d'aide à la décision dans le sens où nous pouvons refuser ou accepter un nouveau dossier de crédit à partir d'un certain score. Et par conséquent, la banque peut gérer son risque en éliminant les dossiers qui présentent un niveau de risque élevé. Elle peut aussi estimer sa perte attendue en utilisant la probabilité de défaut tout en essayant de la combler par les provisions.

Finalement, il faut signaler qu'il n'existe pas un modèle de « Credit Scoring » idéal, parfait et exclusif. En effet, toutes les méthodes utilisées sont des méthodes d'aide à la décision afin d'avoir des résultats plus proches de la réalité et aider les analystes de crédit dans le traitement des dossiers les plus compliqués ou ceux qui se situent dans une zone de forte incertitude par le biais des scores.

Références bibliographiques
