



Mémoire de fin d'Etudes

Thème :

**« Modélisation de Risque de Défaut :
Approche Comparative de Différentes Techniques de Scoring »**

Présenté et soutenu par :

▪ **Fida LOUHICHI**

Encadré par :

▪ **M. Radhouane GOUJA**

Etudiant(e) parrainé(e) par :

▪ **BNA « Banque Nationale Agricole »**

Dédicaces

Je dédie ce mémoire de fin d'études :

*A mes deux chers parents, Samira et Yahia, qui ont tout fait pour moi
pour que je puisse arriver à ce niveau.*

*A mes chères amies, Abir, Rawia et Zouhour, qui m'ont fourni la
motivation et l'encouragement*

A Sofiene qui m'a aidé et m'a fourni des conseils précieux

*A mes chères amies et collègues Bouthaina et Rayhane pour leur
disponibilité et encouragement...*

*Dans l'impossibilité de citer toutes les personnes, mes sincères
remerciements vont à tous ceux et celles, qui de près ou de loin, ont
permis par leurs conseils et leurs compétences la réalisation de ce
mémoire.*

Fida LOUHICHI

Remerciements

Tout d'abord, je remercie Dieu de m'avoir donné les forces, la volonté et le courage afin d'accomplir ce travail.

Je tiens à exprimer ma reconnaissance à mon encadrant académique M. Radhouane GOUJA que je remercie de m'avoir encadré, orienté, aidé et conseillé.

J'ai eu de la chance aussi d'avoir la confiance de M. Mohamed Taher TEJ, qui m'a offert la base de données de la présente étude. Ainsi, je l'adresse mes vifs remerciements.

Je remercie ma chère collègue Ahlem Hammami, pour son encouragement et sa disponibilité.

En guise de reconnaissance, je tiens à témoigner mes sincères remerciements à toutes les personnes : professeurs, intervenants et personnels de l'IFID qui ont contribué de près ou de loin à assister pour réussir les deux années d'études.

Je tiens aussi à remercier les membres du jury d'avoir accepté d'évaluer ce mémoire.

Sommaire

Introduction Générale	1
Partie Théorique	3
Chapitre I : Risques Bancaires et Réglementation en Vigueur	4
Introduction	4
Section I : Risques Bancaires : Concepts et Typologies	5
Section I : Généralités sur le Risque de Crédit	9
Section III : Réglementation Prudentielle Internationale : Normes Baloises	13
Section IV : Réglementation Prudentielle Nationale	26
Conclusion	29
Chapitre II : Méthodes d'Evaluation du Risque de Crédit	30
Introduction	30
Section I : Approche Traditionnelle Basée sur l'Analyse Financière	31
Section II : Notation Externe : « Rating »	34
Section III : Approche Statistique de « Crédit Scoring » :	39
Section IV : Approche Basée sur l'Intelligence Artificielle	55
Conclusion :	62
Partie Empirique	63
Chapitre I : Méthodologie Empirique d'Elaboration d'un Modèle de Scoring : Cas de la BNA	64
Introduction :	64
Section I : BNA en Chiffres : Indicateurs Clés de Performance	64
Section II : Méthodologie Empirique : Description des Données et Présentation des Variables	68
Section III : Etude des Pouvoirs Discriminants : Sélection Préalable des Variables	75
Conclusion :	84
Chapitre II : Modélisation du Risque de Crédit	85

Introduction :	85
Section I : Costruction d'un modèle de « Crédit Scoring »	85
Section II : Etude Comparative de Performance des Modèles	98
Conclusion :	111
Conclusion Générale	112

Liste des Abréviations

SNI	: Système de Notation Interne
ADF	: Analyse Discriminante Factorielle
AUC	: Area Under the Curve
BCBS	: Basel Committee on Banking Supervision
BCT	: Banque Centrale de Tunisie
BDD	: Base De Données
BFR	: B esoin en F onds de R oulement
BNA	: Banque Nationale Agricole
BRI	: Banque des Règlements Internationaux
CA	: Chiffre d'Affaires
CAF	: Capacité d'Autofinancement
CAHT	: Chiffre d'Affaires Hors taxes
CBCB	: Comité de Bâle sur le Contrôle Bancaire
CR	: Classe de Risque.
CT	: Court terme
CV	: Coefficient de Variation
D	: Défaillante
EAD	: Exposure At Default
EBE	: Excédent Brut d'Exploitation
EL	: Expected Loss
FDR	: Fonds de Roulement
FN	: Faux Négatifs
FP	: Fonds Propres
FP	: Faux Positifs
G-10	: Groupe des Dix
H0	: Hypothèse nulle
Ha	: Hypothèse alternative
IRB	: Internal Rating Based
IRB_F	: Internal Rating Based Foundation
IRB-A	: Internal Rating Based Advanced
LCR	: Liquidity Coverage Ratio
LGD	: Loss Given Default
LT	: Long terme
M	: Maturité
MD	: Million de Dinars
MLP	: MultiLayer Perceptron
N	: Année
ND	: non-Défaillante
NI	: Notation Interne

NSFR	: Net Stable Funding Ratio
OR	: Odds Ratio
P	: Précision
PD	: Probabilité de Défaut
PNB	: Produit Net Bancaire
R	: Rappel
RL	: Régression Logistique
RN	: Résultat Net
RNA	: Réseau de neurones Artificiel
ROA	: Return on Assets
ROC	: Receiver Operating Characteristic
ROE	: Return on Equity
RWA	: Risk Weighted Assets
SA	: Approche Standard
SPSS V26	: Statistical Package for the Social Sciences Versio-26
SSB	: Sum of Squares Between
SST	: Sum of Squares Total
TN	: Trésorerie Nette
TPE	: Très Petites Entreprises
TR	: Taux de Recouvrement
UL	: Unexpected Loss
VaR	: Value at Risk
VIF	: Variance Inflation Factor
VN	: Vrais Négatifs VN
VP	: Vrais Positifs

Liste des Tableaux

Tableau n°1 : Grille de notation selon les principales agences de notation	36
Tableau n°2: Résultats empiriques d'Ohlson (1980)	52
Tableau n°3 : Indicateurs de rentabilité de la BNA en [2016-2021]	65
Tableau n°4 : Répartition de la BDD initiale :	70
Tableau n°5 : Liste des variables indépendantes :	72
Tableau n°6 : Résultats du test de Student	76
Tableau n°7 : Résultats du test de Mann-Whitney	77
Tableau n°8 : Résultats du test de Khi-2	78
Tableau n°9 : Ratios financiers corrélés	79
Tableau n°10 : Eta^2 : Mesure de l'effet de taille	81
Tableau n°11 : Mesure de multi-colinéarité	82
Tableau n°12 : Liste des variables retenues	83
Tableau n°13 : Fonction 1 aux centroïdes des groupes	86
Tableau n°14 : Statistiques de validité globale de la fonction de score	87
Tableau n°15 : Matrice de confusion (ADF échantillon d'apprentissage)	88
Tableau n°16 : Variables significatives de la RL	89
Tableau n°17 : Odds Ratio par Variables	91
Tableau n°18 : Test de Hosmer et Lemeshow	92
Tableau n°19 : Matrice de confusion (RL échantillon d'apprentissage)	93
Tableau n°20 : Architecture du RNA retenue	95
Tableau n°21 : Pourcentage de prévisions incorrectes par échantillon	95
Tableau n°22 : Matrice de confusion (RNA échantillon d'apprentissage)	97
Tableau n°23 : Résultats de classement des modèles	99
Tableau n°24 : Sensibilité-spécificité par modèle	100
Tableau n°25 : Précision-Rappel par modèle	102
Tableau n°26 : Résultats de classement : validation « out of sample »	104
Tableau n°27 : Résultats de classement : validation « out of time »	105
Tableau n°28 : F1-Score des modèles par échantillon	106
Tableau n°29 : Interprétation des variables de la régression logistique	108

Liste Des Figures

<i>Figure n°1 : Résultats de discrimination des relations par deux modèles</i>	42
<i>Figure n°2 : Règle de décision du modèle d'Altman (1968)</i>	46
<i>Figure n°3 : Arbre de décision</i>	57
<i>Figure n°4 : Activité d'un neurone dans un RNA</i>	59

Liste Des Graphiques

<i>Graphique n°1 : Evolution du PNB de la BNA [2016-2021]</i>	65
<i>Graphique n°2 : Evolution du taux des provisions de la BNA en [2017-2021]</i>	66
<i>Graphique n°3 : Portefeuille des créances de la BNA par segment de clientèle en 2021</i>	67
<i>Graphique n°4 : Provision du segment « Corporate » par rapport au total des provisions de la BNA</i>	67
<i>Graphique n°5 : Importance en % par prédicteur</i>	96
<i>Graphique n°6 : Gains cumulés en %</i>	97
<i>Graphique n°7 : Courbe ROC de trois modèles</i>	101
<i>Graphique n°8 : AUC par modèle de scoring</i>	101
<i>Graphique n°9 : Courbe PR de trois modèles</i>	103

Liste des Annexes

<i>Annexe 1 : Statistiques descriptives</i>	<i>122</i>
<i>Annexe 2 : Présélection des variables discriminantes.....</i>	<i>123</i>
<i>Annexe 3 : Etude de dépendance des variables à retenir</i>	<i>127</i>
<i>Annexe 4 : Analyse discriminante factorielle.....</i>	<i>128</i>
<i>Annexe n°5 : Régression Logistique.....</i>	<i>129</i>
<i>Annexe n°6 : Perceptron multicouche</i>	<i>131</i>
<i>Annexe n°7 : Analyse de performance des modèles</i>	<i>135</i>
<i>Annexes n° 8 : Validation « out of sample » des modèles</i>	<i>136</i>
<i>Annexes n° 9 : Validation « out of time » des modèles.....</i>	<i>137</i>

Introduction Générale

Nul ne peut nier l'importance du rôle des banques dans le développement des pays et le financement des économies étant donné qu'elles sont considérées comme les canaux quasi-unique du drainage de l'épargne vers l'investissement.

Compte tenu de ce rôle d'intermédiation financière ainsi que de la complexité et de la diversité des services qu'ils offrent, les établissements de crédit subissent toujours une panoplie de risques pouvant menacer non seulement la performance des banques, mais aussi la stabilité et la résilience du système bancaire dans son ensemble.

En effet, la surveillance des risques relatifs à l'activité bancaire devient de plus en plus difficile, notamment avec les répercussions de la crise des « subprimes », qui mettent la gestion et l'évaluation du risque de crédit en général et le risque de défaut en particulier au centre des préoccupations des autorités de contrôle et de tout établissement de crédit.

Il s'agit du thème central des accords de Bâle II, dont l'objectif du premier pilier est l'amélioration du calcul des risques et leur couverture par des fonds propres plus solides à travers la mise en place de modèles permettant d'évaluer le risque que présente chaque relation, et ce en se basant sur des critères à la fois qualitatifs et quantitatifs.

Afin de garantir une bonne maîtrise du risque de crédit, les banques tunisiennes sont notamment invitées à développer leur propre système d'évaluation des risques par la mise en place de la circulaire n°2016-06 relative au système de notation interne des contreparties (SNI).

Dans ce contexte, les banques sont tenues d'adopter une des méthodes d'évaluation du risque de crédit afin de quantifier le défaut. En effet, il existe de multiples techniques mises à la disposition des établissements de crédit, parmi lesquelles nous pouvons citer :

- Les **techniques dites qualitatives**, elles sont fondées sur les jugements des experts de la banque sur la base des informations relatives à la nature de l'activité, à la stratégie suivie et au risque financier de l'emprunteur, dont l'objectif est de chercher à tirer parti de la relation privilégiée que l'institution financière peut entretenir avec ses clients.
- Les **techniques dites quantitatives** ou encore « le crédit scoring ». Sont des méthodes basées sur des approches statistiques utilisant des éléments financiers et non-financiers, des variables à la fois qualitatives et quantitatives.

Le choix de la méthode à adopter est tributaire de la stratégie de gestion de risque, de la structure et de la clientèle de la banque. Il dépend également des hypothèses et des conditions d'utilisation de la méthode retenue.

Au niveau de cet exercice de recherche, notre objectif est d'élaborer un modèle de « credit scoring » permettant d'évaluer le risque de défaut individuel des entreprises du segment « Corporate » au sein de la « Banque Nationale Agricole BNA ».

Pour arriver à cette finalité, nous allons faire recours, en premier lieu, à trois méthodes distinctes de scoring, telles que :

- L'analyse discriminante,
- La régression logistique,
- Le réseau de neurones artificiel.

En se basant sur une analyse comparative entre les résultats obtenus par les trois méthodes, nous allons identifier, en deuxième lieu, celle la plus pertinente.

Par ailleurs, l'objectif ultime de ce mémoire est d'apporter une réponse constructive à la question suivante :

« Quel est le modèle d'évaluation de risque de défaut le plus approprié aux relations du segment Corporate de la BNA ? »

Notre démarche méthodique se décline en deux parties (Théorique et Empirique) :

- Au niveau de la **première partie**, nous allons présenter, dans [le premier chapitre](#), les différents types de risque liés à l'activité bancaire en général, et le risque de crédit en particulier tout en précisant ses différentes composantes et facteurs ainsi que la réglementation internationale et nationale en matière de risque de crédit. Dans [le deuxième chapitre](#), nous allons expliquer les différentes méthodes mises à la disposition des établissements de crédit pour l'évaluation de risque de défaut de leurs relations.
- Au niveau de la **seconde partie**, nous allons exploiter trois différentes méthodes de crédit scoring pour l'évaluation de risque de défaut des relations appartenant au segment « Corporate », dans [le premier chapitre](#). Par la suite, à travers une étude comparative entre les résultats estimés par les trois modèles, nous allons sélectionner celui le plus performant pour le cas de la BNA, dans [le deuxième chapitre](#) de la partie empirique.

Partie Théorique

Introduction

L'internationalisation et l'interdépendance continues des marchés financiers, accompagnées par la révolution technologique ont privilégié, ces dernières décennies, les opportunités de profits pour les établissements financiers mais aussi la montée des différentes catégories de risques.

La nature de l'activité bancaire, essentiellement par son rôle d'intermédiation financière et ses services connexes, ainsi que sa diversité exposent les établissements bancaires à de nombreux risques. Cette exposition peut engendrer des pertes comme elle peut assurer des gains énormes. En effet, l'histoire a été marquée par de multiples crises ayant neutralisé la croissance et la stabilité du système bancaire.

De plus, dans un environnement de plus en plus concurrentiel, plusieurs facteurs peuvent inciter la banque à prendre des positions risquées permettant de mettre en péril l'établissement, voire de menacer la stabilité du système dans son ensemble par des effets de contagion.

De ce fait, les autorités de contrôle ont renforcé les mesures préventives pour faire face aux différents risques bancaires en général, et au risque de crédit en particulier.

Dans le contexte national, la banque centrale tunisienne a mis en place une réglementation prudentielle en matière de gestion de risque de crédit décrivant les normes, que tout établissement de crédit doit les suivre.

Au niveau de ce chapitre, nous allons, dans la première section, présenter les différents risques liés à l'activité bancaire.

Dans la deuxième section, nous allons aborder la notion du risque de crédit, ses composantes et ses facteurs.

Finalement, au niveau des deux dernières sections, nous allons présenter la réglementation bancaire internationale et nationale en matière de risque de crédit.

Section I : Risques Bancaires : Concepts et Typologies

Le risque bancaire représente la probabilité de réalisation d'un évènement pouvant engendrer une perte financière potentielle, il est principalement incorporé dans les transactions financières, bien qu'il puisse se produire en raison d'autres évènements opérationnels Brunel (2009)¹. De façon générale, nous pouvons adopter la définition déduite par Brunel (2009) : « un risque est un évènement qui peut affecter la chronique des flux à recevoir par un établissement bancaire ».

Dans le domaine bancaire, le risque est mesuré en termes de changement probable de la valeur d'un actif ou du prix d'un titre ou d'un produit de base par rapport à sa valeur ou son prix actuel suite à la survenance de certains évènements tels que :

- ✓ Les ralentissements économiques,
- ✓ Les changements défavorables de la politique fiscale et commerciale,
- ✓ Les mouvements défavorables des taux d'intérêt ou des taux de change,
- ✓ La chute des prix des actions.

I. Typologies des risques bancaires :

Il existe plusieurs critères de classification des différentes catégories des risques bancaires, au niveau du présent travail de recherche, nous avons suivi celui qui a été utilisé par Ghosh (2012)² dans son ouvrage « Managing Risks in Commercial and Retail Banking » 1^{ère} édition. Les risques auxquels une banque peut être confrontée sont classés en deux familles :

- Risques financiers.
- Risques non-financiers.

1. Risques financiers :

Cette famille de risques résulte de l'exercice quotidien de l'activité bancaire. Ils sont dus à la survenance des évènements prévus ou imprévus, dans l'économie ou au niveau des marchés financiers par exemple, qui peuvent engendrer, **directement**, une dégradation de la valeur des actifs et, par conséquent, une détérioration de la valeur intrinsèque de la banque, Ghosh (2012).

Outre le risque de crédit qui fera l'objet d'une présentation détaillée dans la deuxième section, cette catégorie de risques inclut notamment :

¹ BRUNEL, V. (2009). Gestion des risques et risque de crédit. *Article de recherche, Ecole Nationale des Ponts et Chaussées.*
²Ghosh, A. (2012). *Managing risks in commercial and retail banking.* John Wiley & Sons.

1.1. Risques de marché :

D'après le comité de Bâle de 2006 portant sur le contrôle bancaire, les risques de marché représentent les risques de réalisation des pertes financières sur des positions de bilan et de hors bilan suite aux fluctuations défavorables des prix des instruments financiers, titres ou contrats acquis avec l'intention de cession, composant le portefeuille de négociation de l'établissement financier (trading book). Ces risques peuvent porter sur :

- Cours des actions,
- Taux d'intérêt,
- Taux de change,
- Cours des matières premières.

1.2. Risque de taux :

C'est le risque de baisse du résultat d'une banque encouru en cas de mouvement défavorable des taux d'intérêts du fait de l'ensemble des opérations du bilan et de l'hors-bilan, à l'exception des opérations soumises aux risques de marché, Koch and MacDonald (2014)³.

En effet, les fluctuations des taux d'intérêt impactent les bénéfices de l'établissement financier en modifiant le revenu d'intérêts net ainsi que les autres revenus corrélés avec les taux d'intérêt et les dépenses d'exploitation.

De plus, ils ont également une répercussion sur la valeur des dettes des créances, et des instruments du hors-bilan, puisque la valeur actualisée des flux de trésorerie attendus est sensible aux taux d'intérêt.

→ Le risque de taux d'intérêt a un double effet, un de court terme sur la marge d'intérêt et un autre de long terme sur la valeur économique de la banque à travers le biais de l'actualisation par les modifications sur la valeur actualisée des flux futurs et donc sur la valeur actualisée des positions du bilan.

1.3. Risque de liquidité :

Selon Regrui et Al Meriouh (2015)⁴, le risque de liquidité se matérialise lorsque la banque se trouve dans l'incapacité à financer ses actifs et à honorer ses engagements

³ Koch and MacDonald (2014). Bank management. Nelson Education.

⁴Regrui et Al Meriouh (2015) Gestion du risque de liquidité : Problématique du comportement de la clientèle bancaire marocaine.

contractés lors de leur exigibilité avec ses actifs disponibles ou mobilisables à court terme. Dans ce cas, la banque est mise en difficulté par deux situations possibles :

- ✓ Le décalage temporel entre les flux entrants et les flux sortants.
- ✓ L'asymétrie de liquidité entre le passif et l'actif de la banque.

En effet une liquidité mal maîtrisée peut engendrer :

- ✓ Une perte d'opportunité par incapacité de financement du développement de l'activité.
- ✓ Une crise de liquidité par incapacité à honorer les engagements contractés.

1.4. Risque de solvabilité :

Nous pouvons parler du risque de solvabilité lorsque la banque ne dispose pas des fonds propres suffisants afin de couvrir ses pertes éventuelles. Cette incapacité résulte généralement de l'inadéquation entre les fonds propres de la banque et les risques pris par cette dernière. Il existe une grande confusion entre le problème de liquidité et problème de solvabilité.

- ➔ En effet, la solvabilité d'un établissement représente sa capacité à honorer ses dettes à tout horizon (court, moyen et long terme). Par conséquent, le montant des actifs dépasse celui des dettes.

1.5. Risques opérationnels :

La définition souvent retenue du risque opérationnel est celle du comité de Bâle (2005), ils sont définis comme étant « *le risque de pertes provenant de processus internes inadéquats ou défectueux, de personnes et systèmes ou d'évènements externes* ».

De plus, Le Comité de Bâle a retenu une classification listant les différents évènements de risques en sept catégories :

- Fraude interne.
 - Pratiques en matière d'emploi et de sécurité sur le lieu de travail.
 - Dommages aux actifs corporels.
 - Exécution, livraison et gestion des processus.
 - Fraude externe.
 - Clients, produits et pratiques commerciales.
 - Dysfonctionnement de l'activité ou/et des systèmes.
- ➔ Selon cette définition, nous pouvons déduire que les risques opérationnels sont inhérents à tous les activités, les produits et les services et concerne tous les employés

indépendamment de leur fonction à l'interne ou à l'externe de l'établissement. Ils touchent, donc, tous les domaines et regroupent ainsi le risque de conformité et le risque juridique. Par contre, au niveau de cette définition, le comité de Bâle a exclu les risques stratégiques et d'atteinte à la réputation qui seront développés dans la partie qui suit.

2. Risques non-financiers :

Les risques non-financiers ce sont des risques intrinsèques au métier de la banque, liés à son fonctionnement au quotidien et ses processus de gestion.

Ils affectent **indirectement** la situation financière de la banque Ghosh (2012). Cette famille de risques regroupe souvent :

2.1. **Risque de réputation :** ou encore risque d'image,

Il représente à la fois une perte de crédibilité, de confiance ou de notoriété d'une banque auprès du public suite à une mauvaise gestion liée à titre d'exemple à une défaillance en matière d'engagement et de conformité réglementaire avec le régulateur et de transparence, qui mettent en question la déontologie, l'éthique et le professionnalisme de l'établissement bancaire Gaultier-Gaillard et Pratlong (2011)⁵.

Selon Ghosh (2012), le risque de réputation peut résulter de tout type de situation liée à une mauvaise gestion des affaires de la banque ou au non-respect des codes de conduite en matière de gouvernance d'entreprise.

➔ Pour cette raison, ce risque est considéré comme l'un des plus difficiles à gérer car il est totalement immatériel et très sensible en fonction de l'actualité.

2.2. **Risque stratégique**

Il s'agit d'un risque lié aux décisions stratégiques prises par les dirigeants des établissements bancaires, dont l'objectif principal est de respecter les attentes des actionnaires. Il se manifeste à travers l'inefficacité des stratégies commerciales choisies.

⁵Gaultier-Gaillard, S., & Pratlong, F. (2011). Le risque de réputation: le cas du secteur bancaire. *Management Avenir*, (8), 272-288.

Section I : Généralités sur le Risque de Crédit

Dans ce travail nous allons évaluer, comme il est mentionné au niveau de l'introduction générale, le risque de crédit. Pour cette raison nous avons consacré la présente section pour définir les différents concepts liés à cette thématique.

D'abord, nous allons présenter les différentes composantes de risque de crédit. Par la suite, nous allons définir les facteurs liés à ce dernier.

I. Risque de crédit : Définition

En tant qu'intermédiaire financier entre les différents agents économiques et par l'exercice quotidien de son activité d'octroi de crédit, les banques sont fortement exposées à diverses catégories de risques, principalement le risque de crédit, Brunel (2009).

Le risque de crédit représente un événement pouvant impacter négativement les flux à verser par un emprunteur dans le cadre d'un contrat de crédit.

Dans le domaine bancaire, le comité de Bâle a défini le risque de crédit comme étant le risque de non remboursement associé à un prêt accordé par une banque.

→ Si nous généralisons, ce risque représente une baisse potentielle de la valeur d'une créance suite aux changements à la fois dans la **capacité de remboursement** ainsi que dans la **volonté des emprunteurs** à respecter leurs obligations contractuelles.

C'est un risque initié, généralement, par une opération de prêt. De ce fait, l'établissement de crédit est appelé à le supporter à partir du moment où il se met en situation d'attendre une entrée de fonds de la part d'une contrepartie.

En effet, il existe de nombreux événements déclenchant du risque de crédit tels que :

- La faillite,
- Le défaut de paiement,
- La restructuration,
- La dégradation de notation de l'emprunteur.

II. Composantes de risque de crédit :

Selon les travaux de recherches antérieurs, le risque de crédit regroupe plusieurs composantes. En se basant sur les travaux de Coussergues et al. (2017)⁶, Brunel (2009),

⁶De Coussergues, S., Bourdeaux, G., Péran, T. (2017). Gestion de la banque. Dunod. <https://doi.org/10.3917/dunod.decou.2017.01>

Ghosh (2012) et Roncalli (2009)⁷, nous avons constaté que le risque de crédit se subdivise en quatre grandes catégories :

1. Risque de défaut :

Ou encore le risque de non remboursement, il signifie l'incapacité de l'emprunteur à restituer les sommes prêtées à une échéance donnée. Le défaut bâlois a lieu lorsqu'au moins une des conditions suivantes est satisfaite :

- Condition I : la banque juge improbable que le débiteur rembourse en totalité son crédit sans prise des mesures appropriées telle que la réalisation d'une garantie.
- Condition II : le retard de paiement du débiteur dépasse les 90 jours.

2. Risque de dégradation de qualité de crédit :

C'est le risque que la qualité de la contrepartie se détériore « downgrade » suite à la dégradation de sa situation financière et donc l'augmentation de la probabilité de défaut même si le défaut n'est pas encore survenu.

Le risque de dégradation engendre une perte de la valeur de la dette durant sa durée de vie. Ces pertes sont donc réalisées en cas de vente anticipée de la dette sans qu'il y ait eu défaut.

➔ Les pertes sont supportées par le créancier seulement lors de la vente anticipée dans le cas où le débiteur n'est pas encore considéré défaillant.

En effet, si l'emprunteur bénéficie d'une notation de la part d'une agence de notation, il convient de noter que le risque de défaut et le risque de transition de rating sont fortement corrélés dans la mesure où une détérioration de la qualité de la contrepartie peut être un précurseur d'un défaut et où un défaut entraîne nécessairement une détérioration de la qualité du crédit, De Servigny et Zelenko (2010)⁸.

3. Risque de recouvrement :

Le taux de recouvrement, tel que présenté d'une manière détaillée au niveau de la partie suivante en se référant au travail de Bohn et Stein (2009)⁹, permet de déterminer le pourcentage de la créance qui sera finalement récupéré.

⁷Roncalli, T. (2009). *La gestion des risques financiers*. Paris: Economica.

⁸deServigny, A. & Zelenko, I. (2010). Chapitre 3. Défaut et migration de crédit. Dans : , A. de Servigny & I. Zelenko (Dir), *Le risque de crédit: Face à la crise* (pp. 79-127). Paris: Dunod.

⁹Bohn, J. R., & Stein, R. M. (2009). *Active credit portfolio management in practice*. John Wiley& Sons.

Le risque en question correspond à l'incertitude liée aux taux de recouvrement postérieur à un défaut constaté dans la mesure où il est déterminé à travers l'étude de plusieurs variables, notamment :

- Valeur réelle des garanties reçues,
- Durée des procédures judiciaires selon le pays,
- Rang de la banque dans la liste des créanciers.

→ Il s'agit du risque relatif au taux et à la durée de récupération des créances permettant de définir le pourcentage de l'exposition aux risques de la banque, De Coussergues et al. (2017).

4. Risque de dégradation du Spread :

Le spread de crédit représente la prime de risque qui lui est associée. Sa valeur est déterminée en fonction du volume de risque encouru. En effet, si le risque du débiteur augmente cette prime augmente aussi, Gabillon (2007)¹⁰.

→ Donc une baisse du risque de défaut entraîne une réduction du spread.

III. Risque de crédit : Facteurs :

Pour mieux cerner la notion du risque de crédit, il faut d'abord comprendre quelques paramètres de base :

1. Probabilité de Défaut (PD) :

C'est la probabilité que le débiteur n'arrive pas à rembourser la totalité ou une partie de ses engagements contractuels. Il s'agit d'un paramètre utilisé pour le calcul du capital économique et du capital réglementaire pour les institutions bancaires sous les normes prudentielles de Bâle II (voire la section suivante).

Par ailleurs, la PD fait référence à un horizon temporel particulier généralement un an.

2. Exposure At Default (EAD) :

Il s'agit d'un paramètre Bâlois qui représente le montant maximal du prêteur exposé au risque de défaut de l'emprunteur. Autrement dit, c'est l'exposition encourue par le créancier en cas de défaut de son débiteur.

¹⁰ Gabillon, J. (2007). Spreads de crédit et taux d'intérêt. Finance, 28, 121-160. <https://doi.org/10.3917/fin.282.0121>

L'exposition en cas de défaut regroupe à la fois :

- L'exposition du bilan : les sommes effectivement débloquées par l'établissement financier (le prêteur) au client (débiteur).
- L'exposition du hors bilan : c'est l'engagement de l'établissement financier auprès du client pouvant donner lieu à des futurs versements de fonds tels que les crédits non encore décaissés ou bien partiellement décaissés, autorisations de découvert ou quotes d'escompte non utilisées, engagements par signature et autres.

3. Loss Given Default (LGD) :

C'est l'évaluation de la perte effectivement encourue par un créancier en cas de défaut de la part d'une contrepartie. Il représente l'un des indicateurs de risque de crédit mentionné au niveau de la réglementation Bâle II.

Il s'agit d'un paramètre exprimé en pourcentage déterminé à partir du taux de recouvrement (TR) fixé par client, tel que :

$$LGD = 1 - TR$$

→ Si à titre d'exemple $LGD = 50\%$ nous pouvons dire que la perte du créancier représente 50% du montant emprunté.

4. Maturité (M) :

La maturité représente la durée moyenne durant laquelle les crédits seront remboursés.

5. Expected Loss (EL) :

La perte attendue désigne le montant à perdre par un créancier sur une exposition en raison de défaut potentiel d'une contrepartie.

Nous pouvons y avoir de nombreuses approches distinctes pour prévoir et estimer ce montant.

Dans le cadre de modélisation du risque de crédit, la perte attendue représente la somme des fréquences estimées des pertes (PD_i) multipliées par les gravités estimées des pertes ($LGD_i * EAD_i$), pour toute exposition i , dont la formule est la suivante :

$$EL = \sum PD_i * LGD_i * EAD_i$$

6. Unexpected Loss (UL) :

Nous pouvons définir la perte inattendue comme étant la fraction des pertes non couverte par les provisions, elle représente ainsi une mesure de risque formelle introduite en 2004 dans le cadre des réformes réglementaires de Bâle II.

Ces pertes inattendues correspondent aux pertes incertaines ayant des probabilités de survenance relativement faibles mais qui peuvent cependant se produire.

Statistiquement pour un intervalle de confiance donné de la fonction de distribution des pertes, l'UL correspond à la différence entre la perte maximale subie (mesurée généralement par the Value at Risk, VaR) et l'EL, telle que :

$$UL = VaR - EL$$

Cette mesure est principalement utilisée au niveau de la détermination des fonds propres afin de générer les ressources nécessaires à la couvrir l'UL.

- ➔ Les pertes futures qu'un établissement financier peut faire face sont composées des pertes attendues (EL) et d'autres inattendues (UL).
- ➔ Selon les directives de Bâle II concernant les exigences minimales en matière de fonds propres, il faut prendre l'UL uniquement comme base pour obtenir le capital économique à risque puisque l'EL est considérée comme étant « *the cost of doing business* », Morandi (2009).

— Section III : Réglementation Prudentielle Internationale : Normes Baloises —

La succession des crises financières a montré la nécessité de renforcer la surveillance et le contrôle des établissements bancaires. Elle a classé ainsi la sécurité des systèmes financiers parmi les préoccupations majeures des banques centrales dans le monde, Arbulu et Lamarque (2011)¹¹.

C'est dans le but de réduire les effets néfastes de la prise de risque et de renforcer la stabilité et la sécurité du système financier en sa totalité et de protéger les déposants en particulier que fut l'avènement de la réglementation prudentielle, Kefi et Maraghni (2011)¹².

La présente section offre un aperçu sur l'évolution de la réglementation internationale en matière de risque de crédit.

¹¹Arbulu, P., & Lamarque, É. (2011). Chapitre introductif.

¹²Kefi, M. & Maraghni, H. (2011). Respect des normes prudentielles et solvabilité des banques commerciales : étude empirique des banques tunisiennes. *Management & Avenir*, 48, 382-403. <https://doi.org/10.3917/mav.048.0382>

I. Création du comité de Bâle :

La faillite de la banque allemande « Herstatt » en 1974 a déclenché une réaction en chaîne et provoqué une grave crise sur le marché des changes, ce qui a entraîné la défaillance de plusieurs établissements financiers, Galati (2002)¹³.

Par conséquent, Peter Cooke, le directeur de la Banque d'Angleterre, a proposé la réunion des gouverneurs des banques centrales du G10¹⁴ et des contrôleurs bancaires d'où la création d'un comité appelé « le Comité de Bâle sur le Contrôle Bancaire », CBCB (ou encore Basel Committee on Banking Supervision, BCBS en anglais) au sein de la Banque des Règlements Internationaux (BRI) à Bâle, Suisse.

1. Composition du comité de Bâle :

Le CBCB se compose de représentants des autorités de surveillance ainsi que des banques centrales de 28 juridictions¹⁵ qui symbolise l'instance centrale en matière de coordination de la réglementation et de contrôle bancaire.

Ce comité représente le principal organisme à l'échelle internationale de normalisation de la réglementation prudentielle du secteur bancaire. Il se réunit quatre fois par an dont l'objectif ultime est de renforcer la surveillance bancaire et de favoriser ainsi la stabilité financière mondiale.

2. Missions du comité de Bâle

Il est à noter que le comité n'a aucune autorité juridique internationale ou nationale. Son rôle consiste à établir des normes et des lignes directrices générales et à formuler des recommandations que ses membres s'engagent à respecter afin de :

- ✓ Renforcer la régulation des banques,
- ✓ Diffuser les meilleures pratiques bancaires,
- ✓ Assurer la solidité financière du secteur.

Ainsi, pour arriver à ces finalités, le comité de Bâle a proposé un accord prudentiel international en matière des exigences minimales des fonds propres à détenir par les banques.

Cet accord a connu des améliorations permanentes comme suit :

¹³Galati, G. (2002). CLS Bank et le risque de règlement dans les opérations de change. *Rapport trimestriel BRI*.

¹⁴Les membres du G10 sont : Allemagne, Belgique, Canada, États-Unis, France, Italie, Japon, Pays-Bas, Royaume-Uni, Suède et Suisse.

¹⁵Argentine, Australie, Belgique, Brésil, Canada, Chine, Union européenne, France, Allemagne, RAS de Hong Kong, Inde, Indonésie, Italie, Japon, Corée du Sud, Luxembourg, Mexique, Pays-Bas, Russie, Arabie Saoudite, Singapour, Afrique du Sud, Espagne, Suède, Suisse, Turquie, Royaume-Uni et États-Unis.

II. Bâle I : Ratio COOKE :

En juillet 1988, le Comité de Bâle a élaboré le 1^{er} dispositif intitulé « *convergence internationale de la mesure et des normes de fonds propres* ». Ce dispositif a annoncé le **ratio international de solvabilité** sous le nom de « **ratio Cooke** ».

Il est défini comme étant le rapport entre les fonds propres (FP) réglementaires et les actifs pondérés (Risk Weighted Assets, RWA).

$$\text{Ratio Cooke} = \frac{\text{FP réglementaires}}{\text{Actifs pondérés}} \geq 8\%$$

$$\text{FP Réglementaires} \geq 8\% \text{ RWA}$$

→ A travers ce ratio, les autorités ont fixé les exigences minimales en matière des FP, qui doivent être respectées par les banques en fonction des risques pris, à 8 %.

La totalité des éléments d'actifs tel que les crédits consentis et les éléments du hors bilan d'établissements. Ces actifs sont pondérés selon la nature du crédit accordé et de la contrepartie ainsi que la zone géographique de cette dernière.

1. Réformes de 1996 : Amendement risque de marché :

Le scandale de la Barings¹⁶ a montré les insuffisances du ratio Cooke, ce qui a conduit le comité à renforcer cette norme en matière du **risque de marché** et à donner la possibilité aux banques d'utiliser des modèles de notation interne afin de calculer le niveau des FP nécessaire pour l'exercice quotidien de ses activités.

$$\text{Ratio Cooke} = \frac{\text{FP réglementaires}}{[\text{Risque de crédit}, \text{Risque de marché}]} \geq 8\%$$

2. Apports de Bâle I :

Ce ratio a offert plusieurs avantages au système bancaire international. Parmi ces derniers, nous pouvons citer :

- ✓ Adaptation d'une stratégie bancaire en matière de prise de risque plus prudente et plus rationnelle par la mise en œuvre une approche pragmatique et scientifique de gestion de risques car « *le ratio Cooke est devenu un élément de coût* » selon Kefi et Maraghni (2011).

¹⁶La faillite de la banque britannique suite à la prise des importantes positions sur l'indice Nikkei de la bourse de Tokyo par l'un de ses traders sans aucune mesure de contrôle, le 23 février 1995.

- ✓ Standardisation des normes prudentielles et la convergence internationale de la mesure et des normes de fonds propres par l'application du ratio Cooke au niveau de plusieurs pays à travers le monde, Daoui et Mouatassim (2021)¹⁷. Il permet ainsi de comparer équitablement les systèmes bancaires ayant des structures différentes.
- ✓ Renforcement de la solidité et de la stabilité du système bancaire international par l'amélioration du niveau des FP des établissements bancaires.
- ✓ Incorporation des risques de hors bilan dans la mesure de l'exigence des FP.
- ✓ Encouragement des établissements bancaires à déposer des actifs liquides de faible risque à travers les pondérations.

3. Limites de Bâle I :

Malgré les efforts fournis par les autorités internationales pour garantir l'efficacité du ratio de solvabilité, ce dernier a fait l'objet de multiples critiques. Nous pouvons citer à titre d'exemple :

- Cette mesure de solvabilité est standard pour tous les établissements de crédit indépendamment de leur profil de risque et de leur type d'activité.
- Les pondérations utilisées lors du calcul du risque de crédit sont peu sensible au risque associé à chaque contrepartie.
- Le ratio Cooke est jugé limité est statique parce qu'il ne prend pas en considération les durées des encours et les degrés de corrélation des créances constituant le portefeuille de la banque, Moussouni (2014)¹⁸.
- En dépit de la faillite de Barings qui a été provoquée par un trader comme nous avons mentionné, le comité de Bâle n'a pas introduit le risque opérationnel lors des réformes de 1996.

Pendant la période d'application de cette norme, le secteur bancaire international a enregistré 10 défaillances bancaires aux pays développés dont les pertes sont estimées à 4% de leur produit national brut et 59 défaillances bancaires aux pays en voie de développement avec des pertes de 250 milliards de dollars, Moussouni (2014). Selon Sardi (2004)¹⁹, tous les établissements de crédit défaillants ont respecté les exigences minimales des FP déterminées à travers le ratio Cooke.

¹⁷Daoui, D., & Mouatassim, M. (2021). La réglementation prudentielle de Bâle au Maroc: état des lieux, et apports pour le secteur bancaire. *International Journal of Accounting, Finance, Auditing, Management and Economics*, 2(5), 707-719.

¹⁸ HABIBA MOUSSOUNI : « LES ACCORDS DE BÂLE ET REGLES PRUDENTIELLES DES BANQUES: DEFIS ET CONTRAINTES POUR LE SYSTEME BANCAIRE ALGERIEN » (Thèse de doctorat : 2013-2014).

¹⁹Sardi (2004) « Bâle II », édition afgés, paris.

→ Pour ces raisons, l'accord de Bâle de 1988 a été jugé insuffisant d'où l'instauration d'autres normes prudentielles pour combler ces limites est obligatoire.

III. Bâle II :

En 1998, le comité de Bâle a décidé d'améliorer l'accord de 1988 à travers une nouvelle réglementation appelée « Bâle II » afin de renforcer la stabilité du système bancaire international. Cette réforme ne se limite pas aux ajustements des méthodes de calculs des exigences minimales de FP, mais elle englobe ainsi deux autres piliers à savoir la mise en place d'un système de surveillance prudentielle et le renforcement de la discipline de marché.

1. Pilier I : Exigences minimales de FP :

C'est le pilier central de l'accord Bâle II qui représente le passage du ratio Cooke à un autre coefficient de solvabilité mieux adapté aux risques relatifs à l'activité bancaire. Cette mesure est appelée « **ratio McDonough** » établi en 2003 et mis en application à la fin de l'année 2006. Elle regroupe trois types de risque, notamment :

- Risque de crédit,
- Risque de marché,
- Risque opérationnel.

$$\text{Ratio McDonough} = \frac{\text{Fonds Propres Nets}}{[\text{Risque de crédit} + \text{Risque de marché} + \text{Risque opérationnel}]} \geq 8\%$$

Chaque type de risque fait l'objet d'une approche distincte de calcul des exigences spécifiques en FP. Comme notre sujet tourne autour du risque de crédit, nous allons développer au niveau de cette partie seulement les approches relatives à ce dernier.

Au niveau de la nouvelle réglementation, le comité de Bâle a donné le droit aux banques de choisir entre deux grandes approches de détermination des exigences de FP par rapport au risque de crédit.

1.1. Approche Standard (SA) :

Il s'agit d'une amélioration de la sensibilité au risque tout en gardant une approche standard de risque de crédit jugée simple avec des pondérations plus détaillées par rapport à l'ancienne réglementation baloise.

Les pondérations de l'approche standard sont en fonction du rating accordé à chaque catégorie de contrepartie à travers les organismes de notation reconnus par les autorités de contrôle (à titre d'exemple Moody's ou Standard & Poor's) et donc de la situation

financière de l'emprunteur et non plus de la catégorie de créance comme par le ratio Cooke qui distingue les Etats et les entreprises par exemple

1.2. Approche basée sur la notation interne (IRB) :

Le deuxième choix proposé par le CBCB, qui représente une alternative à l'approche standard pour la mesure du risque de crédit, est d'employer les méthodes internes IRB (Internal Rating Based).

- La notation interne (NI) repose sur l'appréciation du risque de crédit par les banques elles-mêmes. C'est une méthode évolutive soumise à la validation et à l'approbation au préalable des autorités de contrôle.

En effet, les réformes de Bâle II imposent des contraintes aux estimations déterminées par la banque suite à l'utilisation de ses modèles internes pour calculer les FP réglementaires nécessaires à la couverture du risque de crédit.

- Mécanismes de l'approche IRB : comme il est souligné par Belás et Cipovová (2011)²⁰, le premier objectif attribué par les établissements de crédit à leur modèle de notation interne est d'affecter chaque emprunteur dans une catégorie de risque adéquate tout en reflétant son niveau de risque réel.

De ce fait, lors de l'utilisation de cette approche, les établissements bancaires sont appelés à classer eux-mêmes leur exposition en fonction de leurs catégories d'actif, telles que : pays, banques, entreprises, financement des projets, portefeuilles d'actions, particuliers et autres.

En effet, nous signifions par un système de notation, selon le comité de Bâle : « *l'ensemble des processus, méthodes, contrôle ainsi que les systèmes informatiques et de collecte des données qui permettent d'évaluer le risque de crédit, d'attribuer des notations internes et de quantifier les estimations de défaut et des pertes.* », document consultatif du CBCB, juin 2006.

- Quantification du risque de crédit par l'approche IRB :

L'approche IRB vise à combler les limites de l'approche standard en reliant directement le niveau d'exigences de FP réglementaires à la PD de l'actif.

²⁰Belás, J. & Cipovová, E. (2011), « Internal model of commercial bank as an instrument for measuring credit risk of the borrower in relation to financial performance (credit scoring and bankruptcy models », Journal of Competitiveness, vol. 4. p. 104-120.

Ces modèles sont fondés principalement sur la fonction de calcul des pondérations en utilisant les bases de données internes de la banque en matière d'évaluation du risque de crédit des actifs.

Pour le calcul des FP réglementaires, la principale étape consiste à déterminer la PD de panier correspondante à chaque catégorie de crédit. Cette « pooled PD » représente la probabilité de défaut moyenne des emprunteurs appartenant à la même catégorie de risque.

- **IRB-F et IRB-A :**

Le comité de Bâle a proposé ainsi aux banques deux familles de modèles IRB pour l'estimation interne des paramètres de risque :

- ➔ Approche **IRB-Foundation** : par l'application de cette approche, les banques déterminent seulement la PD pour chaque contrepartie et les autres paramètres tels que LGD, EAD et M vont être calculés et imposés par le superviseur.
- ➔ Approche **IRB-Advanced** : les modèles avancés donnent le droit aux banques de calculer elles-mêmes, tous les paramètres sous réserve du respect des exigences minimales de qualité des estimations.
- ➔ Apports des modèles IRB: ces modèles de mesure internes devraient permettre une identification plus pertinente du risque de crédit que les approches standards. Ils sont jugés plus efficaces et plus précis par rapport à l'approche des pondérations de Bâle 1.

De ce fait, l'application des méthodes IRB est considérée comme une opportunité pour la plupart des établissements assujettis à la réglementation bancaire car elles représentent une opportunité d'optimiser les processus internes de mesure et de gestion de risque.

- ➔ Difficultés de l'approche IRB :

Pour répondre aux exigences baloises en matière de FP réglementaires, les banques ont été incitées à exploiter leur propre modèle interne de notation, afin d'être au plus près de la réalité par l'estimation d'une manière plus fidèle leurs risques encourus. En vue d'atteindre cet objectif, elles doivent disposer et traiter un volume assez important des données que représente une étape très difficile à gérer.

De plus, l'établissement des modèles d'estimation ainsi que le traitement des données doivent satisfaire certaines conditions exigées par le CBCB, autrement, elles ne peuvent pas appliquer l'approche standard.

2. Pilier II : Processus de surveillance prudentielle :

Le comité a proposé, par les normes Bâle II, des principes fondamentaux relatifs à la surveillance prudentielle ainsi que des recommandations concernant la gestion des risques inhérents à l'exercice quotidien de l'activité bancaire.

Ce processus de surveillance prudentielle tel que présenté par le pilier II de l'accord bâlois a un double objectif :

- ✓ Garantir que les banques disposent des FP suffisants et au-delà des exigences réglementaires de base, afin de couvrir les risques auxquels elles sont exposées.
- ✓ Inciter les établissements bancaires à élaborer les meilleures techniques de surveillance et à les utiliser en matière de gestion des risques.

De plus, les banques appliquant l'approche IRB-A sont appelées de procéder à des simulations de crise extrême afin d'évaluer leur résilience et leur propre capacité de résistance face à un choc extérieur. Elles doivent réaliser :

- ✓ Back testing: afin de prouver la validité des méthodes statistiques utilisées sur des périodes assez longues (5 à 7 ans).
- ✓ Stress testing²¹ : pour prouver la solidité de ses FP en cas de crise économique.

En se basant sur les résultats de ces tests, les autorités prudentielles peuvent imposer à ces banques de réduire leurs prises de risque ou de renforcer les exigences des FP réglementaires par rapport à celle déterminées par le calcul sous le premier pilier avec des FP supplémentaires.

3. Pilier III : la discipline de marché :

L'objectif principal du pilier III « la discipline de marché », selon le rapport du comité, est de compléter les deux premiers : les exigences minimales de fonds propres et le processus de surveillance prudentielle.

²¹Les *stress tests* c'est un moyen de gestion du risque, qui nous informe sur les conséquences défavorables inattendues de certains risques. Il représente ainsi un indicateur sur les montants de FP nécessaires à l'absorption des pertes en cas de crise.

Les autorités de contrôle et de surveillance prudentielle ont estimé que la transparence et la communication financière représentent des outils nécessaires pour restaurer la confiance des marchés et promouvoir la stabilité du secteur bancaire international.

De ce fait, le comité a essayé d'instaurer la discipline de marché à travers un ensemble d'exigences de communication financière pour garantir aux intervenants une bonne appréciation des éléments d'information, essentiellement sur :

- ✓ Les FP,
- ✓ Les expositions au risque,
- ✓ Les procédures d'évaluation des risques,
- ✓ L'adéquation des FP de l'établissement au profil de risque.

De plus, le CBCB a considéré que cette panoplie d'informations est jugée assez importante, en cas d'application des méthodes IRB, particulièrement les IRB-A, qui offrent une souplesse dans l'évaluation des exigences de fonds propres. En effet le comité a exigé ainsi que la publication doit être régulière afin d'assurer une meilleure discipline de marché.

4. Insuffisances de Bâle II :

Au niveau de Bâle II, le comité n'a pas pris en considération le risque de liquidité lors de calculs des exigences minimales en FP réglementaires.

4.1. Limites de l'approche standard :

En dépit de l'amélioration de la sensibilité au risque à l'aide des pondérations plus détaillées par rapport à l'ancienne réglementation baloise, cette classification des paniers de risque réglementaires de l'approche standard a montré des insuffisances.

En effet, le regroupement par exemple, au niveau de la 2^{ème} classe de risque, des contreparties notées A⁺ avec celles notées A⁻ malgré qu'elles disposent des PD largement différentes, par conséquent, elles subissent aux spreads du taux d'intérêt différents, représente une opportunité d'arbitrage qui entraîne une augmentation des positions de risque des banques sans avoir augmenté leur niveau des FP réglementaire.

4.2. Limites de l'approche IRB :

- Bruno & Girod (2008)²² ont montré, à travers l'évaluation de l'application du nouvel accord bâlois pour certains pays émergents, que même avec la corrélation positive

²²Bruno, O. & Girod, A. (2008). Bâle II, réallocation des portefeuilles de crédits et incitation à la prise de risque: Une application au cas des pays émergents d'Asie du Sud-Est. *Revue économique*, 59, 1193-1213. <https://doi.org/10.3917/reco.596.1193>.

entre le niveau des FP réglementaire et la PD associé à chaque catégorie d'actif, la charge en FP issue d'une prise de risque supplémentaire est décroissante. D'où l'aggravation des positions risquées d'un établissement bancaire fortement risqué est proportionnellement moins coûteuse en FP réglementaire par rapport à un autre peu risqué.

- ➔ Étant donné que cette charge supplémentaire est couverte par le spread de taux d'intérêt sur les prêts à risque élevé, certaines banques ont intérêt à augmenter le risque de crédit de leur portefeuille.
- De plus, selon la réglementation de Bâle II, la PD par catégorie d'actif est estimée indépendamment de la conjoncture et des facteurs économiques car elle doit être stable dans le temps pour le calcul des exigences en FP réglementaire. Or le risque de crédit est très sensible à la dynamique macroéconomique Rochet (2008)²³.
- ➔ De ce fait, les banques peuvent avoir des FP insuffisants pour couvrir le risque de crédit réellement encourus même si elles ont bien respecté les réformes de calculs des exigences réglementaires, Bruno & Girod (2008).
- En outre, par les modèles avancés et les techniques d'évaluation de risque sophistiquées, les banques ont la possibilité de sous-estimer leurs risques d'une manière scientifique.

Ces insuffisances ont été mises en évidence par la crise financière des subprimes, qui a incité le Comité de Bâle à envisager de mettre en place une nouvelle réglementation prudentielle d'où le passage aux normes de Bâle III.

IV. Bâle III :

« Si la crise économique et financière, qui a commencé en 2007, a été aussi grave, c'est notamment parce que, dans de nombreux pays, le secteur bancaire avait développé un effet de levier excessif au bilan et au hors-bilan, tout en laissant se dégrader progressivement le niveau et la qualité de ses fonds propres ». CBCB, Bâle III, (Décembre 2010)²⁴.

En effet, cette crise financière, dite des « Subprimes », a souligné de nombreuses failles dans le système de réglementation prudentielle, ce qui a poussé les autorités de contrôle à mettre en place une surveillance adéquate plus sévère sur le système bancaire à travers

²³Rochet, J. C. (2008). Le futur de la réglementation bancaire.

²⁴ Accords Bâle III : « dispositif réglementaire mondial visant à renforcer la résilience des établissements et systèmes bancaires », décembre 2010.

l'instauration des nouvelles exigences pour les banques à l'échelle internationale, Delaite (2012)²⁵.

D'où l'établissement du nouveau cadre prudentiel avec la mise en place des accords de « Bâle III », en décembre 2010, afin

- ✓ D'améliorer la quantité, la qualité et la convergence des exigences des FP réglementaires,
- ✓ De renforcer la liquidité des établissements de crédit,
- ✓ D'optimiser la gestion des différentes catégories de risques bancaires,
- ✓ De renforcer la supervision bancaire, la gouvernance, la transparence ainsi que la communication financière.

Bâle III a été mis en place pour combler les insuffisances liées aux normes de Bâle II dévoilées par la crise de 2007 et réduire donc ses répercussions sur le système financier, à travers :

1. Renforcement du dispositif mondial des FP :

Comme une première réflexion, le comité de Bâle a opté pour le renforcement du niveau des FP requis pour garantir la solvabilité et la solidité des banques face aux pertes qu'elles pourraient enregistrées.

Depuis Bâle III, le Tier one, qui représente le noyau dur des FP bancaires, a devenu 4,5% contre 2% sous Bâle II.

En outre, les exigences minimales du ratio de solvabilité bancaire de 8% de Bâle II ont été majorées par un coussin de conservation de 2,5% pour atteindre la valeur de 10,5%.

2. Effet de levier:

Afin de compléter le niveau réglementaire des FP nécessaire pour la couverture des risques bancaires, le comité de Bâle a établi une nouvelle exigence par l'instauration d'un ratio de levier.

$$\text{Ratio de levier} = \frac{\text{Total Actifs}}{FP} \leq \mathbf{3\% \text{ du Tier One}}$$

²⁵Delaite, M. (2012). Les fonds propres bancaires au cœur de la crise financière. *Mondes en développement*, 160, 117-132. <https://doi.org/10.3917/med.160.0117>

Ce ratio a un double objectif :

- ✓ Premièrement, il permet de limiter l'effet de levier dans le secteur bancaire et de prendre en considération les expositions du hors-bilan de la banque afin d'atténuer le risque de déstabilisation de l'économie en général, et du système financier en particulier.
- ✓ Deuxièmement, ce ratio offre des garanties supplémentaires contre le risque du modèle interne de la banque (erreur de mesure et stabilité des résultats) complétant la mesure basée sur le risque (à travers le ratio de solvabilité) par une mesure simple et indépendante du risque.

3. Constitution du volant contra-cyclique :

La crise des subprimes a mis en évidence la cyclicité des exigences en matière des FP réglementaires.

Dans l'objectif de réduire la procyclicité, le CBCB a instauré tout un dispositif pour encourager la conservation des FP et la constitution des volants adéquats, au-delà du minimum, à mobiliser en périodes de tension et à reconstituer en période de reprise économique.

$$\text{Volant Contra – cyclique} \in [0, 2, 5\%] \text{ des RWA}$$

4. Instauration des normes mondiales de liquidité:

Un nombre important des établissements de crédit ont fait face à des graves lacunes durant la période de crise de 2007 malgré qu'ils disposent un niveau des FP largement supérieur aux exigences réglementaires de Bâle II, Visnovsky (2027)²⁶.

Cette crise a montré que les risques inhérents à l'activité bancaire ne se limitent pas en capital et elle a souligné sous le problème de liquidité.

De ce fait, le comité a mis en place des exigences minimales de liquidité bancaire à court et à long terme :

- **Ratio de liquidité à court terme** : Liquidity Coverage Ratio (LCR) : les banques sont tenues de disposer au minimum un LCR de 100% pour faire face à une course aux guichets durant une période de 30j.

$$\text{LCR} = \frac{\text{Actifs Liquides}}{\text{Flux sortant à 1 mois}} \geq 100\%$$

²⁶VISNOVSKY, F. (2017) Secrétaire général adjoint : « Bâle 1, 2, 3 ... de quoi s'agit-il ? Banque de France.

- **Ratio de liquidité à long terme** : Net Stable Funding Ratio (NSFR) : les établissements de crédit doivent disposer des ressources stables pour financer leurs activités afin de garantir l'exercice de l'activité bancaire pour un horizon de 1an.

$$NSFR = \frac{\text{Ressources Stables Disponibles}}{\text{Fiacemets stables requis}} \geq 100\%$$

V. Bâle IV :

Finalisation des réformes de Bâle III : les réformes de 2017²⁷ :

La mise en place de ces réformes a pour objectif de compléter la première version des accords de Bâle III en vue de renforcer les pratiques des banques à l'échelle internationale, afin d'améliorer la stabilité financière.

Ces réformes sont couramment appelées « Bâle IV » par l'industrie financière, étant donné que les modifications annoncées sont de grande importance.

Les réformes de 2017 sont basées principalement sur les éléments suivants :

1. Révision du calcul des RWA :

Cet axe vise essentiellement l'amélioration de :

- ✓ La robustesse des modèles internes afin d'assurer un niveau minimum de significativité des outputs.
- ✓ La sensibilité d'approche standard par la modification des pondérations des actifs.

2. Fixation d'« output floor » :

Par les modèles IRB, les banques ont arrivé à sous-estimer leurs risques, donc, de réduire leurs exigences en matière des FP réglementaires.

De ce fait, le CBCB a instauré un plancher en capital ou bien un capital « floor », c'est-à-dire un niveau minimum de FP à ne pas dépasser par les établissements de crédit.

*Exigences minimales en FP ≥ **Output Floor** = 72,5% FP approche standard*

L'objectif de fixation de cette limite est de réduire les écarts entre les exigences en FP des établissements bancaires, tout en gardant la sensibilité au risque, pour assurer une gestion plus prudente des risques bancaires.

²⁷Note d'information sur la finalisation des réformes « Bâle III » Banque de France.

En effet, les établissements appliquant l'approche IRB ne peuvent tirer d'avantage que 27,5%.

Section IV : Réglementation Prudentielle Nationale

La réglementation prudentielle nationale en matière de gestion de risque de crédit est élaborée par la banque centrale de Tunisie (BCT) à travers la circulaire n° 91-24 du 17 décembre 1991 qui a été modifiée en 29 juin 2012 par la circulaire n° 2012-09.

Cette circulaire décrit les normes, que tout établissement de crédit doit les respecter, relatives à :

- ✓ La division et couverture des risques.
- ✓ La classification d'actifs bancaires.
- ✓ La constitution des provisions.

Il y a également la circulaire n° 16-06 relative au système de notation des contreparties, qui a été publiée le 11 octobre 2016 ainsi que la circulaire n° 2020-01 du 29 janvier 2020 aux banques et institutions financières portant sur les mesures préliminaires à la mise en place des normes internationales d'information financière (IFRS).

I. Division et couverture des risques :

1. Division des risques :

Afin d'atténuer les pertes potentielles en cas de défaut et dans le but de répartir l'exposition des établissements de crédit au risque de contrepartie, la banque centrale a mis en œuvre la circulaire n° 91-24.

Le 1^{er} article de cette dernière exige que le montant total des risques encourus ne doit pas excéder :

- ✓ **3 fois** les FP nets bancaire, pour les bénéficiaires dont le risque encouru est supérieur ou égal à **5%** des FP nets.
- ✓ **1,5 fois** les FP nets bancaire, pour les bénéficiaires dont le risque encouru est supérieur à **15%** des fonds propres nets.

Au niveau du 2^{ème} article de la même circulaire, la BCT limite la valeur maximale des risques encourus sur un même bénéficiaire à **25%** des FP nets de la banque.

De plus, dans le cadre de l'article 23 de la loi n°2001-65 relative aux établissements de crédit, le montant total des risques encourus sur les personnes ayant lien avec la banque, ne doit pas dépasser **25%** des FP nets. Il est à noter que ces personnes doivent être identifiées par le système de mesure de risque de la banque.

2. Couverture des risques :

Concernant la couverture des différentes catégories de risque, selon les quotités des risques fixées au niveau de l'article 6 de la circulaire n°91-24 le ratio de solvabilité bancaire doit être obligatoirement supérieur ou égale à **8%**. En d'autres termes, les FP net de chaque établissement de crédit ne doit pas être inférieur à **8%** du son total actif, du bilan ainsi que du hors-bilan, pondérés en fonction des risques.

3. Classification des actifs :

A partir de l'année 1991, tout établissement de crédit est appelé à procéder à la classification de ses actifs d'une manière régulière afin de calculer le montant des provisions nécessaires à la couverture des risques encourus.

Il est à noter que, cette procédure est applicable sur tout type d'actif quel que soit du bilan ou du hors-bilan, libellé en dinar ou en devise à l'exception des actifs détenus directement sur l'État ou encore sur la BCT, ils ne font pas l'objet de la classification car il s'agit d'un risque souverain.

Selon la circulaire n° 91-24 du 17 décembre 1991, il existe deux grandes catégories d'actifs, telles que : actif courant et actif classé. Nous pouvons les présenter comme suit :

3.1. Actifs courants :

- **Classe 0 : actifs sains :** Les actifs qui sont détenus sur des entreprises dont la situation financière est solide ou qui ne présente pas de sujets d'inquiétudes majeurs.

3.2. Actifs classés :

- **Classe 1 : actifs nécessitant un suivi particulier :** Ce sont les actifs qui sont détenus sur des entreprises qui honorent leur engagement financier à leur échéance, mais ayant une situation financière qui se dégrade ou qui opèrent dans un secteur d'activité en difficulté.

- **Classe 2 : actifs incertains** : Ils sont détenus par des entreprises dont le retard de paiements des intérêts ou du principal est compris entre **[90j-180j]** ainsi que tous les actifs détenus par des entreprises qui présentent des éléments préoccupants.
- **Classe 3 : actifs préoccupants** : Ils sont affiliés à cette classe des entreprises dont le retard de paiements des intérêts ou du principal appartient à l'intervalle de **[180j-360j]** et tous les actifs détenus sur des entreprises qui rencontrent de sérieuses difficultés.
- **Classe 4 : actifs compromis** : Ce sont les actifs détenus par des entreprises dont le retard de paiements des intérêts ou du principal sont supérieurs à **360 jours**, les actifs restés en suspens au-delà de 360 jours ainsi que les autres actifs qui doivent être passés par pertes après que la banque ait veillé à utiliser toutes les procédures de droit tendant à la réalisation de ces actifs.

4. Constitution des provisions et suivi des engagements :

La constitution des provisions est en fonction de la classification des actifs. Par ailleurs, les banques doivent constituer des provisions d'au moins :

- ✓ **20 %** pour les actifs de la classe 2,
- ✓ **50 %** pour les actifs de la classe 3,
- ✓ **100 %** pour les actifs de la classe 4.

L'assiette de calcul des provisions par actif, ayant un montant au moins égale à 50.000 DT (ou encore 0,5 % des FP nets), est la valeur de la créance nette de la valeur estimée des éventuelles garanties.

De plus, les garanties reçues de l'état, des compagnies d'assurance et des banques ainsi que des hypothèques dûment enregistrées sur biens meubles ou immeubles et qui ont fait l'objet par un organisme indépendant d'évaluations régulières sont admises en déduction de la valeur de l'actif à provisionner.

En effet, la circulaire **n°2013-21** a apporté des améliorations par rapport à celle **n°91-24** en termes de couverture du risque net pour les actifs appartenant à la classe 4. Elle a proposé ainsi une nouvelle décote de la garantie réelle reçue par la banque en fonction de l'ancienneté des actifs dans la classe 4 :

- ✓ Décote de **40%** pour les actifs d'ancienneté $\in [3, 5]$ ans.
- ✓ Décote de **70%** pour les actifs d'ancienneté $\in [6, 7]$ ans.

- ✓ Décote de **100%** pour les actifs d'ancienneté ≥ 8 ans.

5. Etablissement d'un système de notation interne :

Pour une bonne maîtrise du risque de crédit, la banque centrale a mis en place la circulaire (n°2016-06) décrivant le SNI à appliquer par tous les établissements de crédits.

Cette circulaire est adoptée vu l'importance des créances classées depuis plusieurs années et vu leurs poids sur les provisions des différents établissements de crédits.

Ce cadre réglementaire va pousser les institutions à implémenter un système de notation des contreparties afin d'instaurer une gestion plus prudente de risque de crédit en amont et en aval et ce en présentant une feuille de route de la gestion d'octroi et du recouvrement ainsi que de mettre en relief les différents autres paramètres dont la probabilité de défaut représente le paramètre le plus important.

Conclusion

Pour conclure, il est indispensable pour les banques de gérer efficacement les divers risques auxquels elles sont exposées. Parmi ces risques, le risque de crédit constitue un enjeu majeur, étant donné qu'il peut avoir des conséquences importantes sur la rentabilité de l'établissement.

Pour cette raison, les autorités de contrôle nationales et internationales ont mis en place des mesures pour inciter les banques à gérer efficacement ce risque et à évaluer correctement la qualité de leur portefeuille de crédits.

En outre, les banques doivent mettre en œuvre des systèmes d'évaluation du risque de contrepartie et gérer leur portefeuille de crédit afin de minimiser le risque de défaillance.

En effet, nous réservons le deuxième chapitre lié à la partie théorique à la présentation de multiples méthodes d'évaluation du risque de crédit.

Introduction

L'évaluation du risque de crédit d'une contrepartie vise à estimer, en utilisant des variables quantitatives et/ou qualitatives, la probabilité qu'elle rencontre des difficultés financières et qu'elle soit incapable d'honorer ses engagements contractuels. En d'autres termes, c'est la probabilité que la dégradation de sa situation financière engendre un incident de crédit, Kharoubi et Thomas, (2016).

En se référant aux travaux d'Altman et Saunders (1998)²⁸, la mesure du risque de crédit a énormément évolué au cours des années 80 et 90. Cette progression a été influencée par de divers facteurs. Entre autres, nous pouvons citer :

- Une augmentation considérable du nombre de faillites dans le monde.
- Une désintermédiation financière chez les grandes entreprises et le recours direct au marché des capitaux.
- Un secteur bancaire plus compétitif.
- Une dégradation de la valeur réelle des garanties reçues auprès de la clientèle.

Devant le nouveau contexte économique induit par ces évolutions, les praticiens ont essayé de développer des modèles de mesure du risque de crédit plus sophistiqués, Gadhoun et al. (2010)²⁹.

Par ailleurs, pour les modèles individuels, il existe deux catégories : d'une part, les modèles qualitatifs traditionnels basés sur le jugement des experts en se référant à des informations liées à la nature de l'activité, à la stratégie suivie et au risque financier de l'emprunteur et d'autre part, les modèles quantitatifs qui reposent sur des approches statistiques fondées sur des éléments financiers et non-financiers.

Nous consacrons le présent chapitre à la description des différentes méthodes d'évaluation du risque de crédit adoptées par les établissements de crédit tout en mettant en relief les avantages et les limites de chacune entre elles.

²⁸ Altman, EI, & Saunders, A. (1997). Mesure du risque de crédit : évolutions au cours des 20 dernières années. *Journal of banking & finance*, 21 (11-12), 1721-1742.

²⁹Gadhoun, Y., Gueyie, J. & Karim Siala, M. (2007). La décision de crédit : Procédure et comparaison de la performance de quatre modèles de prévision d'insolvabilité. *La Revue des Sciences de Gestion*, 224-225, 177-183. <https://doi.org/10.3917/rsg.224.0177>

Section I : Approche Traditionnelle Basée sur l'Analyse Financière

Il existe plusieurs méthodes traditionnelles d'évaluation du risque de crédit, qui sont généralement employées par les banques et les autres institutions financières en vue de vérifier la solvabilité et la qualité du crédit des emprunteurs.

En effet, l'analyse financière représente une technique parmi les techniques traditionnelles permettant d'évaluer la performance financière d'une entreprise et de prévoir son évolution future.

Nous nous focalisons, au niveau de la présente partie, sur l'analyse financière traditionnelle en tant qu'une méthode largement utilisée par les établissements de crédit jusqu'à présent. Nous allons présenter le principe de base, les avantages ainsi que les limites de cette dernière.

II. Principe de l'approche traditionnelle :

Le principe fondamental de ces techniques est de traiter et d'analyser un ensemble de données afin de déduire une évaluation du risque du demandeur de crédit fondée sur une appréciation subjective. Ces méthodes, dites qualitatives, sont réservées aux professionnels et aux entreprises.

En effet, l'analyse financière repose sur les jugements des experts en se basant sur un nombre de paramètres décrivant la situation financière de la relation tels que les informations relatives au secteur d'activité, à la stratégie adoptée et au risque financier de l'emprunteur, Levy et al. (2021)³⁰.

Cette approche est largement descriptive, elle permet d'avoir une idée sur la qualité globale du crédit d'un emprunteur par le calcul et l'interprétation des différents ratios financiers reflétant la performance de la relation.

En effet, les méthodes traditionnelles consistent à établir un diagnostic financier pour qualifier et apprécier la solidité financière des demandeurs de crédit.

En se référant au livre de Vernimmen : Finance d'entreprise dans sa 3^{ème} édition, l'approche traditionnelle vise à « *étudier le passé pour diagnostiquer le présent et prévoir l'avenir* »,

³⁰Levy, A., Baha, R. & Bouzar, S. (2021). Évaluation du risque de défaut de solvabilité des PME par une double approche: Une application du modèle Logit et du modèle *CreditRisk⁺*. *Gestion* 2000, 38, 51-70. <https://doi.org/10.3917/g2000.381.0051>

Vernimmen (1998)³¹, dont l'objectif principal est d'évaluer la solvabilité future de l'entreprise.

L'analyse financière est assez complexe puisqu'elle repose sur de nombreux indicateurs. Par conséquent, l'établissement de crédit va instaurer et calculer différents ratios pour pouvoir juger la performance ainsi que la solidité financière de l'entreprise en question à travers son compte de résultat et son bilan.

Le diagnostic financier permettant de détecter les entreprises en difficulté doit être en mesure de saisir la probabilité d'un éventuel échec à travers certains signes avertisseurs, car l'existence de difficultés se traduit généralement par certains indicateurs spécifiques.

De ce fait, le traitement des différents documents communiqués par la contrepartie est basé sur des soldes et des ratios analysés dans le temps et généralement sur au moins trois exercices comptables.

III. Avantages de l'approche traditionnelle :

La démarche du diagnostic financier représente la méthode la plus ancienne et la plus utilisée en évaluation du risque de crédit. Selon les recherches de Dietsch et Petey (2003)³², la méthode d'analyse financière présente quelques avantages, nous pouvons citer à titre d'exemple la prise en compte des paramètres non quantitatifs ou encore mal formalisés au niveau du traitement du dossier du client.

De plus, le diagnostic financier intègre le savoir requis par les années d'expérience par la confrontation de toutes les règles formalisées par un groupe d'experts permettant d'avoir une vision complète ainsi qu'une idée approfondie sur la santé financière des emprunteurs et donc d'identifier ceux qui seront défaillants, Makany et GabsouboYienezoune(2013)³³.

IV. Limites de l'approche traditionnelle:

Toutefois, l'approche traditionnelle basée sur le diagnostic financier présente des limites qui aboutissent à des conséquences néfastes.

³¹Vernimmen P. (1998), Finance d'entreprise, 3ème édition par Quiry P. & Ceddaha F., Dalloz.

³²Dietsch, M., & Petey, J. (2003). *Mesure et gestion du risque de crédit dans les institutions financières* (No. 2013/14375). ULB--Université Libre de Bruxelles.

³³Makany, J. & GabsouboYienezoune, C. (2013). L'évaluation du risque de crédit des entreprises: cas de la banque congolaise de l'habitat. *Revue Congolaise de Gestion*, 17, 87-130. <https://doi.org/10.3917/rcg.017.0087>

Cette méthode est jugée lourde et coûteuse en termes de ressources humaines, du temporelles et de matérielles qu'elle requiert. Elle est considérée, parfois, comme une appréciation subjective du risque, qui peut mener à un volume important de créances impayées.

De plus, au niveau des relations banque-client, il existe un certain degré d'asymétrie d'information. Les demandeurs de crédit ont tendance à financer des projets assez risqués par des emprunts bancaires. De ce fait, le client peut, dans certains cas, masquer des informations significatives qui ne sont pas favorables à la décision d'octroi de crédit. Cela engendre des conséquences importantes que la banque doit maîtriser rapidement afin de pouvoir contrôler le risque de contrepartie, Coussergues et al. (2017).

En règle générale, l'approche qualitative est réservée à l'évaluation des grandes entreprises, qui opèrent souvent sur le marché financier et dont leur exposition au risque est encore plus importante.

Finalement, il est jugé difficile de définir une démarche scientifique pour tester et valider les outputs de la méthode traditionnelle puisque cette dernière repose principalement sur le background des experts et sur leur expérience dans le domaine, Sadi (2010)³⁴.

Pour conclure, l'analyse financière est une étape fondamentale pour l'évaluation du risque de chaque contrepartie. Néanmoins, elle est jugée insuffisante pour la prise de décision puisqu'elle est basée sur des états comptables historiques. De plus, elle ne permet pas de prévoir la situation financière future de l'emprunteur.

Donc pour combler ces limites, les établissements de crédit doivent employer d'autres méthodes.

³⁴ SADI, K. (2010). Un état des lieux des différentes techniques d'analyse statistiques pour la construction des modèles de prévision de défaillance. *Revue d'économie et de statistique appliquée*, Volume 7, Numéro 2, Pages 83-95.

Section II : Notation Externe : « Rating »

Passons maintenant à une autre méthode mise à la disposition des banques en vue d'apprécier la qualité de son emprunteur. La notation externe, le rating ou encore le « creditworthiness » tel que indiqué par Brunel (2009), est une méthode d'évaluation du risque de la contrepartie. Elle est principalement utilisée dans le cadre des créances négociables sur le marché financier, De La Bruslerie (2014)³⁵.

Nous allons présenter, au niveau de cette partie la méthodologie générale adoptée par les agences de rating en vue d'évaluer la solvabilité de la contrepartie. Par la suite, nous allons citer quelques avantages offerts par la notation externe et souligner sous les insuffisances de cette dernière.

I. Principe de la méthode :

Il s'agit d'une appréciation de la solvabilité ainsi que de la volonté d'une contrepartie à honorer ses obligations à court et à long terme. Il s'agit d'une note attribuée par les organismes externes d'évaluation du crédit, dont les plus connus sont l'agence Fitch Ratings, Moody's et Standard and Poor's.

Chaque institution a une méthodologie et une échelle de notation différentes. Pour mener leurs analyses et afin de déduire une note reflétant la qualité crédit d'un émetteur.

En effet, pour évaluer le risque financier de la contrepartie, selon Degos et al. (2012)³⁶, les agences de notation traitent, en premier lieu, les éléments quantitatifs et purement financiers, à savoir :

- ✓ La rentabilité financière globale,
- ✓ Le niveau d'endettement,
- ✓ Le niveau de liquidité,
- ✓ Le rendement des capitaux employés,
- ✓ Les cash-flows futurs.

Afin de déduire une appréciation globale de risque, elles intègrent les éléments non quantitatifs au niveau de son processus d'évaluation, notamment :

- ✓ La gouvernance,

³⁵ De La Bruslerie, H. (2014). Chapitre 12. L'analyse du risque crédit par le prêteur. Dans : , H. de La Bruslerie, Analyse financière: Information financière, évaluation, diagnostic (pp. 453-506). Paris:Dunod.

³⁶Degos, J., Ben Hmiden, O. & Henchiri, J. (2012). Les agences de notation financières: Naissance et évolution d'un oligopole controversé. *Revue française de gestion*, 227, 45-65. <https://doi.org/10.3166/RFG.227.45-65>

- ✓ La responsabilité sociale de l'entreprise,
- ✓ La stratégie future de l'emprunteur

II. Output du rating :

Autre que l'évaluation de la solvabilité des emprunteurs par l'attribution des notes, les organismes de notations offrent ainsi une perspective « Outlook » indiquant l'orientation la plus probable de la note dans les 6 à 24 mois suivants. En effet, nous pouvons distinguer trois perspectives de notation :

- Perspectives positives c'est-à-dire la possibilité de voir la note relevée.
- Perspectives stables représentent un probable statu quo.
- Perspectives négatives signifient que la probabilité de voir une dégradation future de la note attribuée est importante.

En outre, ces agences construisent également des matrices de transition représentant l'évolution moyenne des entreprises dans l'échelle de notation, en d'autres termes, elles estiment la probabilité de migration d'une note à une autre ainsi que la probabilité de défaut par :

- ✓ Année,
- ✓ Classes de notations,
- ✓ Secteur d'activité,
- ✓ Régions géographiques.

Tableau n°1 : Grille de notation selon les principales agences de notation

Agences			Signification des notes
Standard & Poor's	Moody's	Fitch Rating	
Catégorie « Investissement »			
AAA	Aaa	AAA	Valeurs de premier ordre
AA+	Aa1	AA+	Qualité haute
AA	Aa2	AA	
AA-	Aa3	AA-	
A+	A1	A+	Qualité moyenne
A	A2	A	
A-	A3	A-	
BBB+	Baa1	BBB+	Qualité moyenne inférieure
BBB	Baa2	BBB	
BBB-	Baa3	BBB-	
Catégorie « spéculative »			
BB+	Ba1	BB+	Éléments spéculatifs
BB	Ba2	BB	
BB-	Ba3	BB-	
B+	B1	B+	Hautement spéculatif
B	B2	B	
B-	B3	B-	
CCC+	Caa1	CCC+	Risques élevés
CCC	Caa2	CCC	
CCC-	Caa3	CCC-	
CC	Ca	CC	Ultra spéculatif
C	C	C	Faibles perspectives
SD et D	/	RD et D	En défaut

Source : élaboré par l'auteur à partir de : www.fitchratings.com, www.moody.com et www.standardandpoors.com.

Comme le montre le tableau ci-dessus, il existe deux grades catégories de notation :

- **Catégorie « Investissement »** [AAA-BBB-] ou [Aaa-Baa3] : regroupe les notes attribuées aux entreprises jugées solides financièrement et ayant un risque de défaut faible.

- **Catégorie « spéculative »** [BB+-D]ou [Ba1-C] : représente les contreparties à risque qui sont très sensibles aux chocs économiques. Pour cette raison, elle offre une rémunération plus importante que la première catégorie.

III. Avantages de la notation externe :

Cette technique d'évaluation du risque de crédit offre plusieurs avantages aux différents intervenants dans le système bancaire international.

Selon les travaux de Wakeman (1998)³⁷, les agences de notation facilitent l'évaluation de la solvabilité d'une contrepartie à travers la publication de ses notations. Elles comblent ainsi un éventuel manque d'informations et dispensent les agents économiques de mener eux-mêmes des évaluations coûteuses. De ce fait, elles offrent un outil pertinent d'aide à la décision pour les investisseurs ainsi que les établissements de crédit à la recherche des informations fiables.

IV. Limites de la notation externe :

Malgré les efforts fournis par les autorités internationales pour garantir l'efficacité de cette technique d'évaluation, la méthodologie des agences de notation a montré de multiples critiques, nous pouvons citer à titre d'exemple :

1. Concentration oligopolistique :

Le problème principal est la concentration de ces agences de notation, un oligopole qui domine le marché, Collard (2012)³⁸. Alors que nous dénombrons plus de 130 agences de notation dans le monde entier, seulement trois entre elles détiennent ensemble à-peu-près 95% du marché, ce sont les anglo-saxonnes Standard & Poor's, Moody's et Fitch Ratings.

2. « Juge et partie » :

Le problème principal est que les organismes de notation sont à la fois « juge et partie ». Elles contribuent à la création des produits structurés en proposant aux établissements bancaires des modèles d'évaluation du risque de défaut.

➔ Les agences de notation ont donc à la fois un rôle de conseil et de sanction.

³⁷Wakeman M. (1998). "The real function of bond rating agencies", The revolution in corporate finance, Stern & Chew (Eds), Malden MA Blackwell Business.

³⁸ Collard, F. (2012). Les agences de notation. Courrier hebdomadaire du CRISP, 2156-2157, 5-60. <https://doi.org/10.3917/cris.2156.0005>

3. Champ d'application restreint :

Puisque les agences de notation exigent une commission assez importante en contrepartie de l'analyse d'évaluation du risque. Cette opération est considérée très coûteuse pour les petites et moyennes entreprises (PME) et encore plus pour les très petites entreprises (TPE).

→ Cela signifie que cette catégorie de clientèle ne sera pas notée.

4. Conflit d'intérêt :

Darbellay (2011)³⁹ a indiqué que le principe « *émetteur-payeur* » dans lequel l'émetteur, et non l'investisseur, engage l'agence de notation pour lui attribuer une notation financière en contrepartie d'une commission, remet en question l'indépendance de l'agence de notation par rapport à celui qui paye la notation. Un conflit potentiel peut donc survenir dans le cas où cette émetteur utilise sa position afin d'exercer des pressions sur l'agence employée, Darbellay et Partnoy (2012)⁴⁰.

En outre, dans certains cas, les analystes de l'agence se trouvent impliqués d'une part dans la relation commerciale, pour négocier les honoraires, et d'autre part dans le processus d'évaluation de risque, Degos et al. (2012).

→ Des pratiques douteuses qui puissent compromettre leur indépendance et leur objectivité en raison de ce conflit d'intérêts.

5. Opacité de méthodologie et manque de transparence :

« Un secret bien gardé » Collard (2012).

Chaque agence de notation suit sa propre méthode de notation. Après la crise des Subprimes, les organismes régulateurs ont instauré de multiples règles relatives aux publications régulières d'informations et ils ont imposé aux agences de notation d'apporter quelques précisions sur les ajustements qu'elles effectuent lors de processus d'élaboration de la note. Cependant, seuls les aspects méthodologiques les plus généraux sont rendus publics.

→ La plus part des procédures publiées sont jugées, floues et complexes.

³⁹Darbellay, A. (2011). *Régulation des notations : l'oligopole des agences de notation de crédit d'un point de vue réglementaire*. Schulthess.

⁴⁰Darbellay, A. & Partnoy, F. (2012). Agences de notation et conflits d'intérêts. *Revue d'économie financière*, 105, 309-318. <https://doi.org/10.3917/ecofi.105.0309>

6. Choix du « Timing » et manque d'anticipation :

Les agences de notation ont du mal à communiquer, au bon moment, une note ou une modification d'une note déjà attribuée. Ce bon timing est particulièrement sollicité surtout en période de crise, Degos et al. (2012).

Leur réaction est souvent tardive, comme l'exemple de la crise de 2007-2008 où ces agences n'ont pas été réactives à la détérioration des marchés financiers et elles ont conservé des bonnes notations sur des produits financiers jugés défaillants.

Parfois, au contraire, elles dégradent trop rapidement leur note, notamment dans la crise de la zone euro en 2010. Celles-ci ont été vivement critiquées pour leur manque de réactivité au bon moment et leur rôle amplificateur.

En conclusion, la notation externe représente l'un des outils d'appréciation du risque de solvabilité de l'emprunteur ainsi que sa capacité à honorer ses engagements contractuels que la banque peut l'employer pour la prise de décision. Cependant, Les agences de notation n'ont pas été en mesure d'anticiper plusieurs crises financières et n'ont pas pu prévoir la faillite de certaines grandes entreprises, Seffar et al. (2019)⁴¹. Ces événements remettent toujours en question la fiabilité de ces évaluations.

Donc, les établissements bancaires doivent fonder leur processus de prise de décision d'octroi sur d'autres méthodes d'évaluation de crédit plus pertinentes.

———— Section III : Approche Statistique de « Crédit Scoring » : ————

Comme nous avons présenté dans le 1^{er} chapitre au niveau de la partie relative à la réglementation baloise, dans le cadre de l'approche IRB (Fondation ou Advanced), les banques ont le droit d'utiliser leurs estimations internes de la solvabilité de leurs emprunteurs afin de mesurer le risque de crédit de leur portefeuille sous réserve de respecter des normes rigoureuses en matière de méthodologie exploitée, Fekir (2009)⁴². D'où, le recours à l'approche statistique de « crédit scoring » en tant qu'outil d'aide à la décision.

Au niveau de cette section, nous allons présenter, dans une première partie, le principe de « crédit scoring », la méthodologie et les conditions d'application de cette approche. Par la suite, dans une deuxième partie, nous allons développer les différentes

⁴¹Seffar, M., Olivero, B., &Luu, P. (2019). Dynamique de la Notation et Performance Financière de l'Entreprise.

⁴²Fekir, H. (2009). Présentation Du Nouvel Accord De Bale Sur Les Fonds Propres. arXivpreprint arXiv:0905.2546.

techniques statistiques de « crédit scoring » toute en précisant les avantages et les limites de chacune entre elles.

I. Crédit scoring : principe, méthodologie et qualité de prédiction :

1. Principe de « Crédit scoring » :

Ce moyen d'appréciation du risque peut être défini selon Mester (1997)⁴³ comme étant « *une méthode statistique utilisée pour prédire la probabilité qu'un demandeur de prêt ou un emprunteur existant soit en défaut de paiement ou en retard de paiement* » tout en tenant compte de tous les critères quantitatifs ainsi que qualitatifs pouvant renseigner sur la contrepartie.

A l'aide des techniques statistiques et des données historiques, les méthodes de scoring cherchent à identifier la contribution de chacune des caractéristiques du demandeur de crédit au défaut de paiement en attribuant une pondération à chaque information représentative de la solvabilité de l'emprunteur, De Coussergues et al. (2017).

Elles tentent ainsi de synthétiser le risque de la contrepartie par un « score » qui se traduit par une probabilité mathématique de défaut permettant de prendre une décision immédiate d'accorder ou de refuser la demande de crédit.

2. Méthodologie de « crédit scoring » :

La conception d'un modèle de notation suit une démarche relativement standard. En effet l'objectif commun entre les différentes méthodes de scoring est de faciliter le diagnostic financier en permettant de sélectionner les caractéristiques des relations défailtantes afin de construire un modèle de discrimination, Gadhoun et al. (2007).

Cette approche consiste à observer ex post sur la base des données historiques (généralement financières et comptables) des entreprises déjà défailtantes et à déterminer la combinaison optimale des facteurs financiers, tel que les ratios financiers, et d'autres non financiers qui les distingue le mieux des entreprises non défailtantes. Le but est d'identifier les variables les plus discriminantes individuellement, puis de construire un modèle statistique d'anticipation de défaillance, Kharoubi et Thomas (2016).

Une fois cette évaluation chiffrée est accomplie, il suffit d'intégrer le score obtenu dans une grille d'évaluation préalablement définie ou encore le comparer à un score critique, Gadhoun et al. (2007).

⁴³Mester, LJ (1997). Quel est l'intérêt de la cote de crédit. Business review , 3 (septembre/octobre), 3-16.

La lecture d'un score suppose donc la détermination d'un niveau de risque qui permet l'interprétation pertinente du score obtenu. Si ce dernier est supérieur, le risque de la relation évaluée est jugé acceptable. Dans le cas contraire, lorsque le risque est considéré important, le client fait l'objet d'une analyse plus approfondie, ou la demande de crédit est tout simplement refusée.

De plus, les établissements de crédit peuvent classer ses clientèles en fonction de leur proportion du risque en deux catégories : des clients sains ou autres vulnérables, Bouhlal (2021)⁴⁴.

3. « Crédit scoring » exigence en matière de qualité de prédiction :

3.1. Degré de précision du modèle :

Étant donné que la prise de risque est en relation positive avec la rentabilité, les établissements de crédit ont toujours intérêt à développer des modèles d'évaluation du risque de crédit plus précis et plus performants dont le but est d'optimiser le rendement des prêts accordés, Boujelbène et Khemakhem (2013)⁴⁵.

La construction de n'importe quel modèle de scoring exige le respect d'un certain niveau de précision. En effet, le modèle adopté doit être en mesure d'évaluer avec précision les caractéristiques spécifiques à chaque emprunteur, de différencier ces risques et de les classer avec suffisamment de cohérence. Ils doivent également permettre une utilisation interne pertinente des résultats obtenus, Fekir (2009).

Cependant, Elhamma (2009)⁴⁶ a précisé dans son travail d'évaluation du risque de crédit par le scoring au sein de la banque populaire de Rabat-Kénitra au Maroc que aucun modèle n'est parfait et qu'il n'a aucune fonction de score à pouvoir séparateur absolu.

En effet, il y a toujours une zone de coïncidence entre les deux sous-groupes (des relations saines et des relations vulnérables) qui génèrent deux types d'erreur :

- **Erreur de type I** : le modèle utilisé classe une relation défaillante parmi les relations saines → nous parlons donc des faux positif.

⁴⁴ BOUHLAL, N. (2021). Internal Rating: A Tool for Managing the Bank Credit Risk of SMEs. La Notation Interne: Un Outil De Gestion Du Risque De Crédit Bancaire Des PME. Revue d'Economie et de Statistique Appliquée, 18(2).

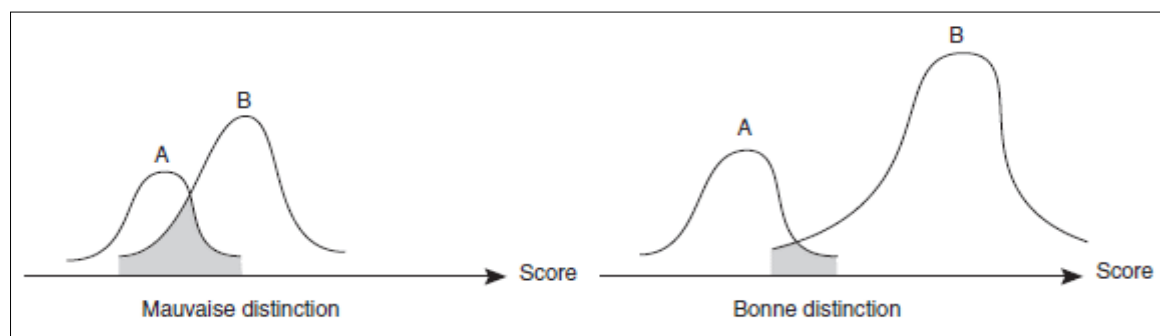
⁴⁵Boujelbène, Y., &Khemakhem, S. (2013). Prévission du risque de crédit: Une étude comparative entre l'Analyse Discriminante et l'Approche Neuronale. arXivpreprint arXiv:1311.4266.

⁴⁶Elhamma, A. (2009). La gestion du risque crédit par la méthode du scoring: cas de la Banque Populaire de Rabat-Kénitra. *Revue marocaine de recherche en management et marketing*, 291.

- **Erreur de type II** : les relations saines seront considérées comme défaillantes par le modèle → nous parlons donc des faux négatif,

La figure ci-dessous illustre une mauvaise et une bonne capacité d'un modèle de scoring à discriminer entre une sous-population A de relations vulnérables et une sous-population B de relations saines.

Figure n°1 : Résultats de discrimination des relations par deux modèles



Source : « Analyse Financière », De La Brulerie (2014)⁴⁷.

En effet, une zone de recouvrement (zone grise) réduite signifie que le risque de confusion entre une relation saine et une autre défaillante est minimisé. Nous pouvons conclure, donc, que le modèle utilisé dispose un pouvoir de discrimination important et un risque d'erreur faible.

3.2. Conditions d'utilisation :

De plus, De Coussergues et al. (2017) ont exigé deux autres conditions nécessaires pour que la technique du crédit scoring employée soit performante.

- ✓ **1^{ère} condition** : il faut que les clients de la banque présentent une certaine homogénéité de comportement afin de standardiser des critères décisionnels valables pour toute la population.
- ✓ **2^{ème} condition** : il faut que les crédits présentent notamment une certaine identité d'objet, de durée et de montant de sorte que les risques encourus soient comparables.

En outre, pour élaborer un bon modèle de scoring, les développeurs ont besoin d'une base de données historique suffisante reflétant la performance des crédits accordés dans des périodes économiques favorables et défavorables.

⁴⁷ De La Brulerie H. (2014), Analyse financière : Information financière, évaluation, diagnostic, Paris, Dunod.

Il est à noter que la mise en œuvre de l'approche IRB est soumise à l'approbation préalable des autorités de contrôle en se référant aux critères définis par le Comité de Bâle.

Dans cette partie, nous présentons les principales catégories des méthodes de mesure du risque par le « crédit scoring », en suivant la classification adoptée par Kharoubi et Thomas, (2016), à savoir : les modèles paramétriques basés sur l'analyse discriminante, les modèles semi-paramétriques ou encore les techniques économétriques sur données qualitatives, tel que le logit et le probit, en présentant les forces et les faiblesses de chaque catégorie.

II. Modèles paramétriques :

Selon Refait-Alexandre (2004)⁴⁸, les méthodes statistiques de classification dites paramétriques visent essentiellement à instaurer une relation fonctionnelle dont la forme est donnée à priori, entre la variable à expliquer et la ou les variables explicatives en supposant que la loi de distribution est connue.

Dès les années 1930, les travaux académiques ont montré que, en se basant sur des données financières et comptables, certains ratios étaient significativement différents entre les bonnes et les mauvaises entreprises.

En effet, les travaux de Beaver (1966)⁴⁹ et d'Altman (1968)⁵⁰ représentent la référence et le véritable point de départ des multiples études empiriques.

Il existe deux grandes familles de méthodes paramétriques de classification statistique soit la méthodologie unidimensionnelle de Beaver (1966) ou l'analyse discriminante multivariée d'Altman (1968).

1. Modèle unidimensionnel de Beaver (1966) :

1.1. Méthodologie de Beaver (1966) :

Parmi les premiers essais de prédiction de défaillances des entreprises à partir de données financières et comptables nous trouvons la méthode unidimensionnelle de Beaver (1966). Il s'agit d'élaborer une classification dichotomique unidimensionnelle qui vise à séparer deux échantillons d'entreprises en sociétés défaillantes et autres non

⁴⁸Refait-Alexandre, C. (2004). La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux. *Économie & prévision*, <https://doi.org/10.3917/ecop.162.0129>

⁴⁹Beaver, WH (1966). Les ratios financiers comme prédicteurs d'échec. *Revue de recherchecomptable* , 71-111.

⁵⁰Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4), 589-609.

défaillantes à travers le seul ratio financier le plus discriminant entre les deux sous-groupes, qui est « Cash-flow / total dette ». Cet indicateur unique est souvent choisi à partir des outils classiques de l'analyse financiers des risques.

1.2. Limites du modèle de Beaver (1966) :

Cependant, malgré sa simplicité, cette ancienne méthode a montré rapidement ces limites. En effet, l'identification d'un seul ratio, dont le pouvoir prédictif est le plus important, représente une étape très délicate parce qu'il représente le seul critère de décision.

De plus cette approche n'a pas pris en compte la corrélation et la linéarité vis-à-vis du risque de l'entreprise des ratios examinés ainsi que l'interaction entre ces derniers, Refait-Alexandre (2004).

➔ De ce fait, le manque de robustesse relatif à l'unicité du ratio retenu justifie sans doute que cette méthodologie n'est plus utilisée.

Pour remédier aux limites d'analyse univariée, les chercheurs recourent à l'analyse discriminante multidimensionnelle, permettant une description plus riche de la situation de l'entreprise, sont maintenant utilisées de manière systématique.

2. Analyse discriminante Multivariée d'Altman (1968)

Par sa publication de 1968, Altman est le premier chercheur utilisant une panoplie des ratios financiers simultanément à travers un modèle de classification fondé sur l'analyse des données, à savoir l'analyse discriminante multidimensionnelle linéaire, afin de prédire la situation financière des entreprises.

2.1. Variables discriminantes :

Pendant la période de [1946-1965], sur une base de 66 entreprises dont 33 saines et 33 défaillantes, Altman a calculé 22 ratios à partir des états financiers. Par la méthode de l'analyse discriminante.

A partir de la liste initiale de 22 variables, seulement cinq variables ont été sélectionnées comme étant celles les plus prédictifs et les plus discriminants entre les deux sous-échantillons d'entreprises (saines et défaillantes) et qui, ensemble, permettent le meilleur de prédire la faillite des entreprises, à savoir :

- ✓ Ratio de liquidité,

- ✓ Ratio de rentabilité,
- ✓ Ratio de levier,
- ✓ Ratio de solvabilité,
- ✓ Ratio relatif à l'activité de la relation.

2.2. Fonction discriminante « Z-score » :

Par la suite, Altman (1968) a développé, sur la base de la combinaison optimale de ces 5 indicateurs, une fonction nommée « Z-score » permettant de prévoir la faillite de l'entreprise et de distinguer au mieux les deux classes de risques.

$$Z - score = 1.12R1 + 0.14R2 + 0.33R3 + 0.06R4 + 0.99R5$$

Dont :

$$\left\{ \begin{array}{l} \blacksquare R1 = \frac{\text{Fonds de roulement}}{\text{Actif Total}} \\ \blacksquare R2 = \frac{\text{Bénéfices non répartis}}{\text{Actif Total}} \\ \blacksquare R3 = \frac{\text{Bénéfices avant intérêts et impôts}}{\text{Actif Total}} \\ \blacksquare R4 = \frac{\text{Bénéfices avant intérêts et impôts}}{\text{Valeur comptable de la dette totale}} \\ \blacksquare R5 = \frac{\text{Ventes}}{\text{Actif Total}} \end{array} \right.$$

La démarche suivie par Altman (1968) afin de parvenir à cette combinaison finale des variables est la suivante :

- ✓ Observation de la signification statistique de diverses variables.
- ✓ Détermination des contributions relatives de chaque variable seule.
- ✓ Evaluation des inter-corrélations entre les variables pertinentes.
- ✓ Observation de la précision prédictive des diverses combinaisons.

Cette fonction est basée essentiellement sur le principe de maximisation de la variance entre les deux sous-groupes en minimisant celle du même groupe. Elle a ainsi la capacité de prédire la défaillance de la relation deux ans avant sa survenance avec 72% de précision et avec 80 à 90% un an avant l'événement.

2.3. Hypothèses du modèle :

Pour développer un tel modèle de discrimination, Altman (1968) a choisi la combinaison qui remplit, principalement, les conditions suivantes :

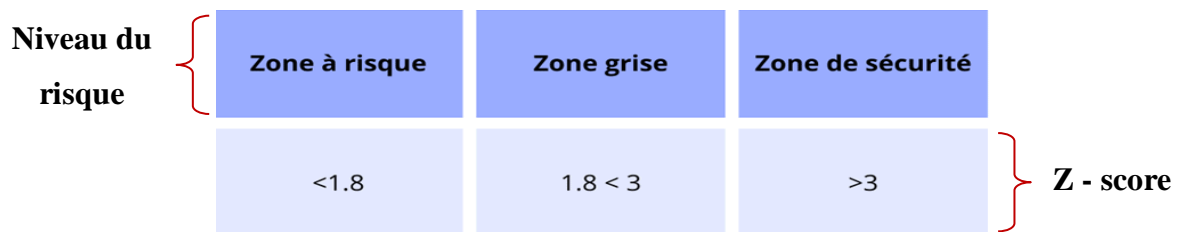
- ✓ Présentant les résultats les plus fiables et le meilleur pouvoir discriminant dont les variables explicatives retenues doivent ainsi posséder une bonne capacité discriminante.
- ✓ Permettant la prise en compte de plusieurs variables simultanément ainsi que de leur inter-corrélation.
- ✓ Fournissant des prévisions avec le meilleur degré de précision possible.
- ✓ Normalité des variables retenues au niveau de l'équation finale.

2.4. Règle de décision :

Les entreprises ayant un score au-dessus du seuil critique seront classées parmi les relations saines, et celles dont le score est au-dessous du seuil critique seront considérées comme étant potentiellement en faillite.

En effet, la figure suivante résume les différentes classes de risque selon le modèle d'Altman (1968)

Figure n°2 : Règle de décision du modèle d'Altman (1968)



Source : élaboré par l'auteur, Altman (1968).

Comme le montre la figure ci-dessus, le risque encourus par les établissements de crédit est en relation négative avec la valeur de Z. Un Z-score d'Altman tend vers 0 signifie que la relation en question risque de faire faillite, tandis qu'un score supérieur à 3 signifie que la société est classée dans la zone de sécurité de la banque où les entreprises ont une bonne situation financière et un risque de faillite est négligeable.

2.5. Modèle dérivé de l'analyse discriminante d'Altman (1968) :

Ce modèle a été la base de multiples travaux de recherche, il a connu ainsi des améliorations et des modifications en termes de variables utilisées et d'approche

appliquée. La version la plus connue est celle élaborée par Altman et al. (1977)⁵¹ intitulée le « ZETA analysis ».

Ils ont combiné à la fois les données financières disponibles dans les états financiers de la firme et la valeur marchande des capitaux propres.

Ce modèle garantit un degré de précision au-delà de 96 % pour le court terme (une période avant la faillite) et de 70 % pour le long terme (cinq périodes avant la faillite) par rapport à 36% par le modèle Z-score. Contrairement à l'ancien modèle de 1968, la mesure ZETA est jugé plus pertinente pour les grandes et les petites entreprises telle que les manufacturiers et de commerçants.

2.6. Avantages de l'analyse discriminante multivariée :

L'utilisation des modèles d'évaluation de risque de crédit basés sur l'analyse discriminante multivariée a fait l'objet des multiples travaux de recherche [Beaver (1966), Altman (1968), Altman et al. (1977), Collongues (1977)⁵², Bardos et Zhu (1997)⁵³]. Ce choix se justifie par le fait que, selon plusieurs auteurs, cette technique statistique présente plusieurs avantages.

➤ Prise en compte du profil complet :

Selon Horrigan (1968)⁵⁴, l'usage des ratios financiers calculés à partir des données comptables pour estimer la performance des entreprises afin de les classer en des sous-groupes résume une large quantité d'informations. En effet, la technique d'analyse discriminante multivariée a l'avantage de prendre en compte le profil complet de la relation en combinant les ratios relatifs à la situation financière ayant le pouvoir discriminant le plus important.

➤ Meilleure qualité des résultats :

Altman (1968) a précisé qu'en raison du caractère multidimensionnel de la fonction, le score offre une prédiction plus fine de la vulnérabilité des entreprises que l'interprétation de chaque ratio individuellement. De plus, le score obtenu représente

⁵¹Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of banking & finance*, 1(1), 29-54.

⁵²Collongues Y. (1977), "Ratios Financiers et Prévision des Faillites des Petites et Moyennes Entreprises", in Banque, vol. n°365 (septembre).

⁵³Bardos, M., & Zhu, W. H. (1997). Comparaison de l'analyse discriminante linéaire et des réseaux de neurones. Application à la détection de défaillance d'entreprises. *Revue de statistique appliquée*, 45(4), 65-92.

⁵⁴Horrigan, JO (1968). Une brève histoire de l'analyse des ratios financiers. *La revue comptable*, 43 (2), 284-294.

un outil fiable de prise de décision d'accorder ou non un crédit ainsi que un indicateur du niveau de risque et le fait de disposer d'un modèle d'évaluation du risque de crédit à haut pouvoir prédictif représente un avantage concurrentiel très important, Modrik (2016)⁵⁵.

En suivant le principe fondamental de l'analyse discriminante tel que la maximisation de la variance entre les N groupes d'entreprises ayant la même notation et la minimisation de l'intra-groupe, cette technique est souvent utilisée pour discriminer essentiellement être les grades entreprises dotées de la même notation car elle offre une mesure plus fine du risque de solvabilité de la contrepartie, Degos et al. (2010)⁵⁶.

2.7. Limites de l'analyse discriminante multivariée :

Plusieurs méthodes ont été développées pour évaluer le risque de crédit. La technique la plus utilisée par les établissements de crédit est celle du crédit scoring par l'analyse discriminante. Cependant, les modèles fondés sur l'analyse discriminante ont été largement critiquée par plusieurs auteurs [notamment par : Altman et al. (1994)⁵⁷, Bardos et Zhu (1997)⁵⁸, Kharoubi et Thomas (2016) et autres].

➤ Hypothèses restreintes :

Malgré le niveau de précision qu'elle garantie lors de discrimination entre les différents groupes d'entreprises, l'usage ainsi que la performance de cette technique sont conditionnés par la validation de plusieurs hypothèses statistiques strictes jugées restrictives qui ne peuvent pas être respectées pour des données financières, notamment, celle relative à la normalité, à l'indépendance des variables retenues ainsi que à l'égalité des matrices de variance-covariance entre les deux sous-échantillons, Bardos et Zhu (1997).

➤ Insensibilité de la méthode au degré de risque de la relation :

Altman et al. (1994) ont conclure, dans le cadre d'une étude comparative, que la robustesse des modèles discriminants est limitée dans sa capacité à différencier entre

⁵⁵Modrik, K. (2016). *L'évaluation du risque de crédit des PME françaises internationalisées* (Doctoral dissertation, Université d'Angers).

⁵⁶Degos, JG, Hmiden, OB et Henchiri, JE (2010). Agences de notation : ajustements comptables et analyse économétrique. *Revue internationale d'économie et de comptabilité*, 1 (1-2), 88-106.

⁵⁷Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of banking & finance*, 18(3), 505-529.

⁵⁸Bardos, M., & Zhu, W. H. (1997). Comparaison de l'analyse discriminante linéaire et des réseaux de neurones. Application à la détection de défaillance d'entreprises. *Revue de statistique appliquée*, 45(4), 65-92.

les entreprises ayant un niveau de risque différent mais qui appartiennent au sous-ensemble des entreprises défaillantes telles que relations non viables et relations vulnérables nécessitant un suivi particulier.

En outre, Altman et al. (1994) ont conclu qu'il est certainement beaucoup plus difficile pour ces modèles de distinguer entre deux échantillons d'entreprises défaillantes (non viables et vulnérables) qu'entre des entreprises clairement saines et non viables.

➔ Donc, avec l'augmentation de la taille de l'échantillon, par conséquent l'augmentation de la variabilité des situations possibles, le degré de précision du modèle diminuent.

Pour surmonter ces hypothèses restrictives et les limites présentées ci-dessus, différentes méthodes alternatives sont à disposition des établissements de crédit. En effet, les travaux de recherche ont conduit à des méthodes qui ne modélisent pas les lois conditionnelles à priori, mais les probabilités à posteriori. Ces méthodes sont basées sur des techniques économétriques, telles que le modèle Logit et le modèle Probit, Kharoubi et Thomas (2016).

III. Modèles semi-paramétriques

En pratique, l'hypothèse de l'égalité des matrices de variance-covariance entre les deux sous-groupes d'entreprise est rarement respectée empiriquement, donc le recours aux méthodes probabilistes représente une solution utile.

1. Méthodologie des modèles économétriques :

Les modèles économétriques des variables qualitatives supposent des lois de distribution des variables financières retenues différentes et lèvent, par conséquence, la limite relative à la technique de l'analyse discriminante, Kharoubi et Thomas (2016).

Ces modèles cherchent à déterminer la probabilité d'occurrence du l'événement de faillite ou de non faillite compte tenu du profil du risque d'une entreprise donnée.

Il existe deux modèles qui se distinguent par la loi statistique de distribution des résidus :

- **Modèle « Logit »** pour des erreurs suivant une loi logistique.
- **Modèle « Probit »** pour des erreurs suivant une loi normale.

Remarque : le modèle logit est considéré comme une approximation du modèle probit. Mais dans le cadre paramétrique, par rapport au modèle probit, la régression logistique a des

hypothèses moins restrictives que celles de l'analyse discriminant linéaire, Bardos (2008)⁵⁹. C'est pour cette raison qu'au niveau de la présente partie nous nous concentrons sur la régression logistique.

En fait, quel que soit le type du modèle choisi, logit ou probit, la variable endogène Y est toujours une variable qualitative dichotomique ou encore polyatomiques dont le comportement n'est pas nécessairement linéaire, Degos et al. (2010)⁶⁰.

2. Écriture de modèle :

Dans le cas particulier d'évaluation du risque de crédit d'un demandeur de prêt, Y représente soit une entreprise « i » défailante soit une autre saine comme suit :

$$Y_i = \begin{cases} 0 & \text{si } i \in ND \\ 1 & \text{si } i \in D \end{cases}$$

Avec :

- **D** représente le groupe des entreprises jugées défailtantes,
- **ND** représente le groupe des entreprises saines ou non-défailtantes.

L'objectif des modèles qualitatifs est d'expliquer la variable endogène Y en fonction d'un vecteur de variables exogènes composé de K ratios financiers estimés les plus discriminants entre les deux sous-groupes (D et ND) tout en minimisant la corrélation entre eux. Telle que la fonction à estimer est la suivante :

$$\text{Score}(Y_i) = \beta + \alpha X_i + \varepsilon_i$$

Avec :

- **X** : (X_1, X_2, \dots, X_k) : le vecteur des k ratios sélectionnés,
- **α** : le vecteur des coefficients représentant les poids attribués aux ratios retenus,
- **β** : la constante du modèle estimé,
- **ε_i** : les termes d'erreur sont indépendants et identiquement distribués.

Donc :

$$Y_i = \begin{cases} 0 & \text{si } \beta + \alpha X_i + \varepsilon_i > 0 \text{ ou encore } \varepsilon_i > -\beta - \alpha X_i \\ 1 & \text{si } \beta + \alpha X_i + \varepsilon_i \leq 0 \text{ ou encore } \varepsilon_i \leq -\beta - \alpha X_i \end{cases}$$

⁵⁹Bardos, M. (2008). Scoring sur données d'entreprises: instrument de diagnostic individuel et outil d'analyse de portefeuille d'une clientèle. Revue Modula, 38, 159-177.

⁶⁰Degos, J. G., Hmiden, O. B., & Henchiri, J. E. (2010). Credit rating agencies: accounting adjustments and econometrical analysis. International Journal of Economics and Accounting, 1(1-2), 88-106.

Les termes d'erreur ε_i représentent les résidus du modèle supposés indépendants, de moyenne égale à 0 et de variance égale à 1. Puisque nous traitons le cas d'un logit, ils sont supposés suivre une loi logistique dont la fonction de répartition F est présentée comme suit :

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

Les paramètres du modèle α et β sont obtenus en suivant la méthode du maximum de vraisemblance. Il s'agit d'un problème d'optimisation par lequel nous cherchons à trouver le maximum de la fonction de vraisemblance de sorte que les probabilités des réalisations observées soient aussi maximum.

La régression logistique permet de déterminer la probabilité à posteriori P_i pour l'entreprise étudiée « i » appartient au groupe des ND sous les hypothèses suivantes :

$$P_i = P(Y_i = 1/X_i) = P(\varepsilon_i \leq -\beta - \alpha X_i) = F(-\beta - \alpha X_i) = F(\beta + \alpha X_i) = \frac{1}{1 + e^{-(\beta + \alpha X_i)}}$$

$$1 - P_i = P(Y_i = 0/X_i) = P(\varepsilon_i > -\beta - \alpha X_i) = 1 - F(-\beta - \alpha X_i) = F(\beta + \alpha X_i) = \frac{1}{1 + e^{\beta + \alpha X_i}}$$

Nous avons donc : $\mathbf{Logit}(P_i) = \ln\left(\frac{P_i}{1-P_i}\right) = \beta + \alpha'X_i$

3. Etudes empiriques antérieures :

3.1. Etude d'Ohlson (1980)⁶¹ :

Un des premiers travaux de recherches utilisant l'approche économétrique qualitative pour prévoir la défaillance des entreprises à travers la régression logistique, nous trouvons l'étude d'Ohlson (1980).

Sur un échantillon de 105 entreprises défaillantes durant la période de [1970-1976] et un autre composé de 2058 entreprises saines, Ohlson (1980) a identifié les ratios financiers ayant le pouvoir discriminant le plus important, tels que :

⁶¹Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of accounting research, 109-131.

Tableau n°2: Résultats empiriques d'Ohlson (1980)

Ratios financiers retenus par Ohlson (1980)	
$R1 = \text{Log} \frac{\text{Actif total}}{\text{PNB ajusté}}$	$R2 = \frac{\text{Dette totale}}{\text{Actif total}}$
$R3 = \frac{\text{Fonds de roulement}}{\text{Actif total}}$	$R4 = \frac{\text{Dette à CT}}{\text{Actif circulant}}$
$R5 = \begin{cases} 0 \text{ si } \text{Dette totale} > \text{Actif total réel} \\ 1 \text{ si } \text{Dette totale} \leq \text{Actif total réel} \end{cases}$	$R6 = \frac{\text{Résultat net}}{\text{Actif total}}$
$R7 = \frac{\text{Fonds générés par l'exploitation}}{\text{Dette total}}$	$R8 = \begin{cases} 0 \text{ si } \text{Résultat net}_{t-1} \text{ et/ou } \text{Résultat net}_{t-2} \geq 0 \\ 1 \text{ si } \text{Résultat net}_{t-1} \text{ ou } \text{Résultat net}_{t-2} < 0 \end{cases}$
$R9 = \frac{(\text{Résultat net}_t - \text{Résultat net}_{t-1})}{(\text{Résultat net}_t + \text{Résultat net}_{t-1})}$	

Source : Ohlson (1980).

Par les 9 ratios mentionnés ci-dessus, Ohlson (1980) a déduit la fonction suivante :

$$O - score = -1,32 - 0,407 R1 + 6,03 R2 - 1,43 R3 + 0,07 R4 - 2,37 R5 - 1,83 R6 - 1,72 R7 + 0,285 R8 - 0,521 R9$$

Avec : $P_i = \frac{e^{O-score}}{1+e^{O-score}}$ représente la probabilité à posteriori de défaillance.

Afin d'affecter une entreprise A à l'un des deux groupes, D ou ND, sur la base de la probabilité de défaillance, Ohlson (1980) a déterminé un seuil de décision à 0,5 choisi en maximisant la qualité du classement.

- ✓ Si sa probabilité de faillite $P_i \geq 0,5$: alors l'entreprise A est jugée ND.
- ✓ Si sa probabilité de faillite $P_i < 0,5$: alors l'entreprise A est jugée D.

3.2. Etude de Matoussi et al. (1999)⁶² :

Une des premières tentatives d'estimation de défaillance des entreprises tunisiennes en utilisant l'approche économétrique qualitative telle que la régression logistique, nous pouvons citer celle de Matoussi et al. (1999).

Sur un échantillon de 88 entreprises tunisiennes de différente taille et de différents secteurs d'activité dont 50 saines et 38 défailtantes durant la période de [1994-1996], Matoussi et al. (1999) ont essayé d'estimer la probabilité de défaillance en utilisant

⁶²Matoussi, H., Mouelhi, R., & Salah, S. (1999, May). La prédiction de faillite des entreprises tunisiennes par la régression logistique. In 20ÈME CONGRES DE L'AFIC (pp. CD-Rom).

initialement une batterie de 31 ratios financiers calculés à travers les bilans et les comptes de résultats des entreprises étudiées.

Le modèle finalement retenu, en se basant sur la significativité des résultats, le pouvoir explicatif et le taux de bon classement, est le suivant :

$$\text{Logit}(P_i) = \ln \left(\frac{P_i}{1-P_i} \right) = 2,259 - 55,702 R21 - 6,989 R25 + 6,674 R5 - 1,144 R30$$

Dont les ratios retenus sont les suivant :

- $R5 = \frac{\text{Dette LT et MT}}{\text{Capitaux permanents}}$
- $R21 = \frac{\text{Cash-flows}}{\text{Actifs total}}$
- $R25 = \frac{\text{Capitaux propres}}{\text{Passif total}}$
- $R30 = \frac{\text{Fonds de roulement}}{\text{Actif circulant}}$

Ce modèle a permis de reclasser les entreprises avec un degré de précision de 87% deux années avant et il arrive à 91% une année avant la défaillance avec un seuil de probabilité P_i de 0,5.

4. Avantages des modèles économétriques :

La régression logistique a été considérée comme l'une des techniques d'évaluation du risque de défaillance des contreparties les plus utilisées du fait qu'elle offre des multiples avantages par rapport aux autres méthodes présentées auparavant :

4.1. Domaine d'application plus large :

Cette méthode dispose d'un champ d'application plus vaste. Elle permet de traiter à la fois des variables quantitatives et d'autres qualitatives, Kharoubi et Thomas (2016).

En tant que procédure semi-paramétrique, Paget (2003)⁶³ a précisé que le modèle économétriques qualitatifs en général et le logit en particulier présentent l'avantage de ne pas exiger des strictes contraintes. En fait, elle ne nécessite pas d'hypothèses a priori sur la distribution des variables retenues dans le modèle.

De plus, la relation déduite entre la variable à expliquer et celle explicatives n'est pas nécessairement linéaire et les matrices de variance-covariance de deux groupes n'ont pas obligatoirement égales entre les deux sous-groupes d'échantillon.

⁶³ Éric Paget-Blanc (2003), Rating et probabilité de défaut des entreprises européennes : détermination par un modèle de régression logistique ordonné. Banques & marchés. N0 65.

4.2. Outil d'aide à la décision :

En se référant à la recherche de Refait-Alexandre (2004), le recours aux méthodes probabilistes permet de tenir compte de la probabilité à posteriori d'appartenance à un groupe.

Par ailleurs, cette technique indique le degré de risque d'un demandeur de crédit, et non seulement une simple affectation en termes de probabilité de faillite. Ce qui facilite la prise de décision d'accorder ou de refuser l'octroi de crédit à la relation étudiée.

4.3. Outil d'explication de défaut :

La régression logistique permet d'étudier la relation entre le défaut et plusieurs ratios reflétant au mieux la situation financière des entreprises étudiées en accordant à chacun entre eux un poids selon leur importance dans l'explication de défaut, Bouazzara et al. (2020)⁶⁴. A travers cette relation elle permet aux analystes d'examiner les effets conjoints de plusieurs interactions, Degos et al. (2010)⁶⁵.

En outre, Ohlson (1980) a choisi le modèle logit pour prédire la défaillance des entreprises car il a jugé l'usage des variables financières employées en tant que prédicteurs de la défaillance des entreprises de l'échantillon étudié comme plus intéressant et plus utile plutôt que de les utiliser à des fins de comparaison entre deux types d'entreprises. Il est beaucoup plus important pour un analyste de prêt d'estimer le degré de risque et les difficultés financières éventuelles que de chercher à estimer le résultat final d'un prêt, Refait-Alexandre (2004).

5. Limites des modèles économétriques :

La régression logistique est une méthode utile et jugée pertinente pour prédire de défaut. Mais cette approche qualitative d'évaluation du risque de crédit présente ainsi quelques inconvénients :

Bien qu'elle a levé la limite relative aux strictes contraintes exigées par l'analyse discriminante, Kharoubi et Thomas (2016) ont précisé que « *les résultats obtenus par ces méthodes sont bons mais reposent sur des hypothèses restrictives* ».

Afin de garantir des estimations significatives, il faut que les termes d'erreurs du modèle suivent soit la loi logistique soit la loi normale (selon le choix du modèle). Par ailleurs, la

⁶⁴BOUZZARA, A., Riad, B. A. H. A., & BEKTACHE, F. (2020). Évaluation du risque de défaillance de solvabilité des PME: une application du modèle de la régression logistique. *Dirassat Journal Economic Issue*, 11(2), 491-505.

⁶⁵Degos, J. G., Hmiden, O. B., & Henchiri, J. E. (2010). Credit rating agencies: accounting adjustments and econometrical analysis. *International Journal of Economics and Accounting*, 1(1-2), 88-106.

corrélation entre les variables sélectionnées doit être faible, tandis que, par définition, les ratios retenus sont corrélés (formule de calcul : numérateur et dénominateur).

De plus cette méthode n'est robuste qu'en absence du problème de multi-colinéarité. Finalement, la régression logistique n'est pas adaptée pour prédire des variables qualitatives avec plus de deux modalités.

Section IV : Approche Basée sur l'Intelligence Artificielle

En pratique, les données ne valident pas généralement les hypothèses imposées par les modèles cités précédemment. Par conséquent, nous ne pouvons pas utiliser les modèles paramétriques et semi-paramétriques, Altman et Saunders (1997).

Dans la littérature financière, d'autres modèles ont été récemment développés, tels que les méthodes basées sur l'intelligence artificielle. En effet, cette technique est en plein développement, elle est utilisée dans de nombreuses disciplines telles que : la médecine, le marketing, la finance et autres.

La solution consiste alors à employer des méthodes non-paramétriques. Celles-ci ont l'avantage de ne pas dépendre d'hypothèses paramétriques pour les distributions des variables. Elle représente ainsi une alternative intéressante aux techniques classiques de traitement des données statistiques, Lotfi et Mesk (2020)⁶⁶.

De ce fait, nous allons réserver cette dernière section du chapitre à la présentation de deux méthodes non paramétriques fondées sur l'utilisation des techniques l'apprentissage automatique, à savoir :

- La méthode de partitionnement récursif.
- Les réseaux neuronaux artificiels multicouches.

I. Partitionnement récursif :

Tandis que les méthodes économétriques produisent un score, puis éventuellement une classification de la population étudiée en appliquant une règle discriminante donnée, les modèles arborescents commencent par produire une classification des individus en des sous-groupes, puis déterminent un score pour chacun entre eux, Giudici (2003)⁶⁷.

⁶⁶Lotfi, S., &Mesk, H. (2020). Prédiction du risque de crédit: étude comparative des techniques de Scoring. *International Journal of Accounting, Finance, Auditing, Management and Economics*, 1(2), 511-527.

⁶⁷Giudici, P. (2003) *Applied Data Mining: Statistical Methods for Business and Industry*. Wiley, Hoboken.

Parmi les modèles arborescents, nous trouvons la méthode de partitionnement récursif. Cet algorithme de classification supervisée permet de construire un arbre dichotomique dit décisionnel.

Elle représente l'une des premières méthodes non-paramétriques utilisées dans la prévision de défaillance des contreparties.

L'objectif final de cette technique de segmentation est de regrouper les modalités qui ne se distinguent pas par rapport à Y par la construction d'un arbre maximal dont :

- ✓ Le critère d'arrêt de division de l'arbre est satisfait,
- ✓ Les individus appartenant à chaque classe sont identiques,
- ✓ La subdivision entre les individus du même groupe n'est plus possible.

1. Algorithme de construction :

La phase de construction de l'arbre de décision maximal consiste à segmenter les individus de l'échantillon en N classes prédéfinies selon la nature statistique de la variable explicative dans notre cas $N = 2$ soit une entreprise défaillante ou saine.

Donc, il faut sélectionner la variable explicative $X_j \in X$ ayant le pouvoir discriminant le plus important afin de séparer au mieux les individus de chaque classe au niveau de chaque nœud j .

Au niveau de chaque nœud j de l'arbre, l'échantillon étudié, composé des entreprises défaillantes et d'autres non-défaillantes, est décortiqué de manière séquentielle à l'aide de la variable explicative choisie permettant de déterminer si le parcours de l'arbre doit être poursuivi vers :

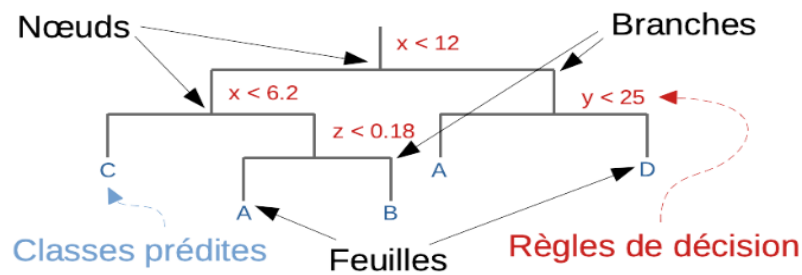
- La gauche du nœud : si sa valeur est inférieure au seuil fixé.
- La droite du nœud : sinon.

La fixation de la valeur du seuil à chaque nœud est déterminée par l'algorithme en suivant la même démarche de détermination de la variable explicative, c'est-à-dire en maximisant la distance entre les deux sous-échantillons (des défaillants et des non-défaillants), Rudd et al. (2017)⁶⁸.

⁶⁸Rudd, M. P. H., GStat, J. M., & Priestley, J. L. (2017). A comparison of decision tree with logistic regression model for prediction of worst non-financial payment status in commercial credit.

La figure suivante schématise les différentes composantes de l'arbre de décision :

Figure n°3 : Arbre de décision



Source : Science des données biologiques III, Grosjean et Engel (2020)⁶⁹.

Le traitement de données par cet algorithme est poursuivi tout au long de l'arbre de classification de façon récursive jusqu'à atteindre un point nommée « feuille » représentant l'extrémité d'une classe donnée (A, B, C ou D).

Au niveau de chaque nœud de l'arbre, chaque sous-échantillon est décomposé en deux sous-échantillons plus petits en appliquant la règle de décision relative.

2. Condition d'arrêt :

Cette opération est répétée en boucle tant que la subdivision en sous-populations est encore possible, c'est-à-dire jusqu'au l'arrivé aux groupes finaux composés de 100% d'entreprises défaillantes ou 100% d'entreprises saines au niveau des nœuds dits « purs ».

Afin d'optimiser le processus de segmentation par la réduction du nombre de nœuds au niveau de l'arbre et le nombre d'itération de l'algorithme de classification, nous pouvons accepter un seuil inférieur à 100% dite de tolérance au niveau duquel l'homogénéité des sous-échantillons d'entreprises est jugée suffisante.

Par conséquent, les nœuds finalement retenus sont qualifiés « impurs » dont le degré de l'impureté du nœud reflète le degré d'hétérogénéité du sous-échantillon qui le caractérise, Refait-Alexandre (2004).

3. Avantages du partitionnement récursif :

La qualité majeure de cette technique prédictive est la flexibilité, en tant que méthode non paramétrique, l'arbre décisionnel ne nécessite pas d'hypothèse sur la loi de distribution des variables utilisées. En effet, ce caractère signifie que la méthode est généralement applicable indépendamment de la nature des variables à expliquer et explicatives.

⁶⁹Grosjean, P. & Engels, G. (2022), Science des données biologiques III. <https://wp.sciviews.org/sdd-umons3>

Giudici (2003) a montré, à travers une étude comparative entre la régression logistique et la technique de l'arbre de décision, que cette dernière est plus performante et plus discriminante en utilisant deux critères de performance : l'indice de Gini et la courbe ROC (Receiver Operating Characteristic) qualifiant le degré de précision des modèles.

De plus, Refait-Alexandre (2004) a indiqué que l'utilisation de la méthode de l'arbre de segmentation nous permet de prendre en considération à la fois le coût des erreurs de classification et les probabilités à priori de défaut.

4. Limites du partitionnement récursif :

Cependant, le partitionnement récursif présente de multiples limites, la flexibilité offerte par cette méthode demande une masse importante de données, ce qui n'est pas une affaire évidente pour plusieurs banques. En effet, les résultats obtenus en utilisant ces techniques séquentielles dépendent fortement des données exploitées, Giudici (2003).

4.1. Complexité :

Cette technique nécessite l'exécution répétée de calculs et de comparaisons, ce qui peut entraîner une surcharge du processeur et ralentir le traitement des données. En effet, elle peut être difficile à mettre en œuvre et à comprendre par des utilisateurs inexpérimentés.

De plus, la complexité des algorithmes de nature séquentielle rend difficile la standardisation ainsi que la généralisation de la structure de l'arbre choisie.

4.2. Instabilité :

C'est une technique jugée instable temporellement et sensible aux petits changements pouvant altérer la structure de l'arbre, Refait-Alexandre (2004).

Passant maintenant au deuxième modèle d'inélégance artificielle instauré afin de pallier les insuffisances relatives aux techniques traditionnelles : réseaux de neurones artificiels.

II. Réseaux de neurones artificiels :

Les réseaux de neurones artificiels (RNA) ont été développés initialement dans le domaine de l'apprentissage automatique afin d'essayer d'imiter la neurophysiologie du cerveau humain en

combinant des unités de calculs informatiques simples appelées neurones dans un système hautement interconnecté Anandarajan et al. (2001)⁷⁰.

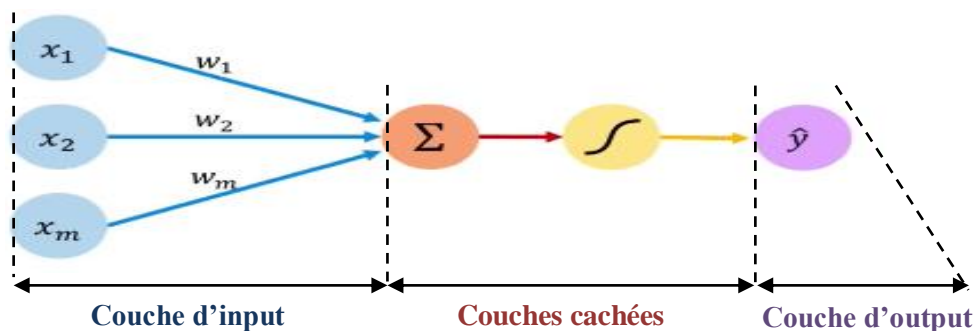
1. Architecture de modèle :

Les unités de calculs informatiques de réseau de neurones sont organisées en couches de telle sorte que chaque neurone d'une couche est exclusivement relié à d'autres de la couche précédente et suivante.

Chaque neurone, représente une unité de calcul autonome qui reçoit pour son activation des entrées sous la forme d'une série de signaux $[x_1, x_2, \dots, x_m]$. Cependant, ce dernier ne produit qu'un seul signal de sortie \hat{y} .

Chaque signal d'entrée est associé à un poids de connexion $[w_1, w_2, \dots, w_m]$, qui reflète l'importance relative du signal d'entrée dans la production de l'impulsion finale transmise par le neurone (la fonction f), Giudici (2003).

Figure n°4 : Activité d'un neurone dans un RNA



Source : élaborée par l'auteur à partir de Giudici (2003).

Cette approche est devenue une des méthodes importantes et puissantes de traitement non linéaire de données, où aucune hypothèse n'est posée à priori, Du Jardin (2007)⁷¹, employées dans des nombreux domaines, à savoir l'extraction des données descriptives et prédictives d'analyse du risque.

L'une des caractéristiques importantes de cette méthode est l'apprentissage. Ceci consiste à développer le réseau de neurones afin d'atteindre le résultat attendu. En d'autres termes, le fonctionnement du RNA repose sur la réception de données d'entrée que le neurone doit

⁷⁰Anandarajan, M., Lee, P., &Anandarajan, A. (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *Intelligent Systems in Accounting, Finance & Management*, 10(2), 69-81.

⁷¹Du Jardin, P. (2007). *Prévision de la défaillance et réseaux de neurones : l'apport des méthodes numériques de sélection de variables* (Doctoral dissertation, Université Nice Sophia Antipolis).

apprendre à reconnaître afin de les classer en sous-groupes homogènes tels que défaillants et non-défaillants, Lotfi et Mesk (2020).

2. Différents types d'architecture :

Il existe plusieurs modèles de RNA utilisés par les auteurs pour la classification des entreprises qui se différencient en fonction de quatre caractéristiques suivantes :

- Degré de différenciation de la couche d'entrée et de sortie.
- Nombre de couches.
- Direction du flux pour le calcul.
- Type de connexions.

Cependant, il y a toujours trois composantes fondamentales quelque soit le type utilisé, Giudici (2003), notamment :

- L'architecture des neurones.
- La topologie du réseau.
- L'algorithme d'apprentissage.

3. Etudes empiriques antérieures :

De multiples études comparatives ont employé la méthode de RNA au problème de classification du risque de crédit afin d'améliorer la qualité des résultats obtenus par d'autres méthodes classiques en général et par l'analyse discriminante en particulier [Anandarajan et al. (2001), Altman et al (1994), Bardos et Zhu (1997), Du Jardin (2007), Giudici (2003), Lotfi et Mesk (2020),...].

La majorité des travaux ont conclu que les résultats de RNA sont plus précis et robustes que les méthodes statistiques traditionnelles pour l'évaluation du risque de crédit.

En effet, Altman et Saunders (1997) ont considéré l'application de cette approche comme étant l'analyse discriminante non linéaire en abandonnant l'hypothèse de linéarité et d'indépendance des variables utilisées dans la fonction de prédiction de faillite.

Sur un échantillon de 1000 entreprises italiennes opérantes dans le secteur d'industrie durant la période de [1982-1992], Altman et al. (1994) ont montré que pour obtenir des résultats plus significatifs, il est recommandé de combiner plus qu'une méthode de prédiction. De ce fait, ils ont utilisé à la fois la technique de l'analyse discriminante et celle de RNA.

4. Avantages des réseaux de neurones :

Les réseaux de neurones sont des modèles de « Machine Learning »largement utilisés pour la prédiction de défaut de crédit car ils présentent plusieurs avantages par rapport aux méthodes classiques d'évaluation du risque de crédit.

4.1. Capacité de traitement des données complexes:

Les réseaux de neurones sont capables de traiter des données de

- Différents types : numériques, catégorielles, temporelles, et autres,
- Différentes dimensions : unidimensionnelles, multidimensionnelles.

Cela les rend particulièrement adaptés lorsque les données sont complexes et/ou de nature hétérogène. En outre, cette technique est capable de traiter même les bases de données présentant des valeurs manquantes, Giudici (2003).

4.2. Capacité de traitement de données non-linéaires:

Les réseaux de neurones sont des modèles fondés sur les techniques d'intelligence artificielle permettant de modéliser des relations complexes entre les variables d'entrée et de sortie, même lorsque ces relations ne suivent pas une forme linéaire simple.

Cette capacité à traiter des données non linéaires est un atout majeur des réseaux neuronaux, car elle leur permet de saisir des modèles et des structures complexes dans les données avec plus de précision par rapport à d'autres méthodes d'évaluation, Altman et al. (1994).

5. Limites des réseaux de neurones :

Toutefois, il faut savoir que les réseaux neuronaux ont également certaines limites qui sont susceptibles de réduire leurs performances et leur précision. Nous pouvons citer à titre d'exemple :

5.1. Modèle difficile à comprendre :

Comme tout modèle basé sur l'intelligence artificielle, le réseau de neurones est qualifié comme étant une « boîte noire ». Car il s'agit d'un modèle qui ne permet pas de comprendre comment il estime ses résultats, ni comment il a été construit.

Cela peut être un problème en particulier lorsqu'il s'agit de modèles de crédit scoring, car il est important de pouvoir expliquer et justifier les décisions de crédit prises par la banque.

5.2. Modèle difficile à interpréter :

Parfois, l'architecture du modèle neuronal est considérée comme complexe et peut se révéler difficile à interpréter même pour des experts, Boujelbène et Khemakhem (2013).

Cette complexité peut rendre difficile l'explication des décisions prises par le modèle. Comme il peut poser des problèmes de confiance et de transparence dans certaines applications.

5.3. Modèle difficile à valider :

Pour ces différentes raisons, il est difficile de savoir si le réseau neuronal artificiel est juste et équitable dans son traitement des différentes demandes de crédit. En d'autres termes, si le modèle est utilisé pour prendre des décisions importantes, il peut être difficile de savoir s'il est fiable et fonctionne correctement. Cela rend la validation ainsi que l'approbation du modèle très difficile.

Conclusion : _____

Le choix de la méthode d'évaluation du risque de crédit représente une décision importante pour les établissements de crédit. En effet, il existe de multiples conditions à vérifier avant d'adopter une approche bien déterminée.

Selon la méthode choisie, l'établissement de crédit pourra évaluer différemment le risque de crédit d'un client et cela peut avoir des implications sur la décision de prêt.

L'objectif du présent chapitre était la présentation des différentes approches mises à la disposition des banques pour l'évaluation du risque de ses clients, afin d'avoir une vision plus claire et complète sur la situation financière de l'emprunteur en général, et sur sa solvabilité en particulier.

Chacune des méthodes présentées ci-dessus dispose des avantages et des limites. Pour choisir entre elles, il faut vérifier les conditions d'application de la méthode et s'assurer de la validité des hypothèses relatives à cette dernière.

Partie Empirique

Chapitre I : Elaboration d'un Modèle de Scoring : Cas de la BNA

Introduction :

Pour se conformer aux exigences réglementaires toujours plus strictes et absorber l'augmentation des dépenses en capital liées au risque de crédit, les établissements de crédit revoient leur approche de la gestion des risques.

En effet, la gestion du risque de crédit vise à réduire les pertes en s'assurant de l'adéquation des fonds propres de la banque avec ses provisions pour pertes de crédit attendues à un moment donné. De plus, par l'optimisation de la gestion des risques, les banques peuvent aussi améliorer considérablement leur performance globale et obtenir un avantage concurrentiel.

A cet effet, l'objectif de notre étude consiste à élaborer un modèle de « crédit scoring » permettant de prévoir la défaillance des entreprises et d'assurer, en occurrence, une meilleure visibilité sur le risque de crédit des relations.

Avant de procéder à la construction des modèles, une étude statistique exploratoire 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 lesquelles nous allons modéliser la défaillance des relations.

Le plan de ce chapitre est comme suit :

Nous allons décrire au niveau de la première section la situation actuelle la « Banque Nationale Agricole » (BNA) en se focalisant sur les indicateurs de risque de crédit.

Dans la deuxième section, nous allons présenter notre méthodologie empirique tout en précisant l'échantillon de l'étude et les variables à utiliser.

Finalement, nous allons identifier les variables les plus discriminantes entre les entreprises saines et défaillantes nécessaire à la construction du modèle, dans la troisième section.

Section I : BNA en Chiffres : Indicateurs Clés de Performance

Au niveau de la présente section, nous allons étudier les différents indicateurs d'exploitation, de la rentabilité, et des risques tout en présentant la situation actuelle de la banque. De ce fait, nous faisons recours à plusieurs variables clés, ainsi par exemple :

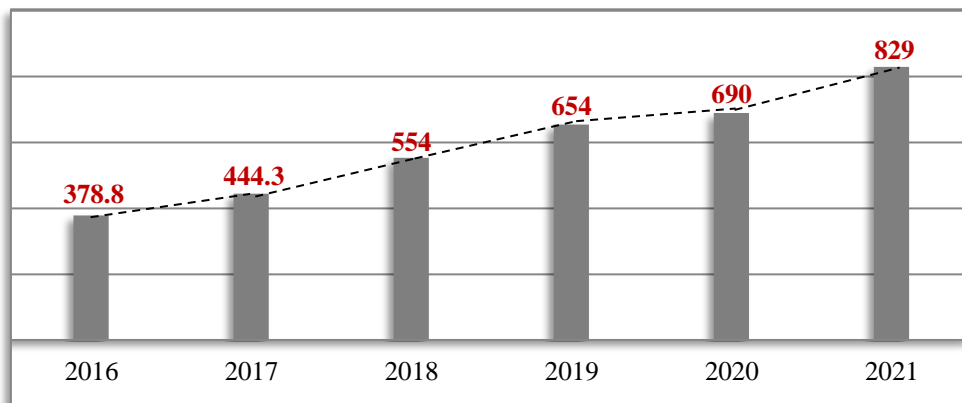
- Indicateur d'exploitation.
- Indicateurs de rentabilité.
- Indicateurs de risque de crédit.

I. Indicateurs d'exploitation : produit net bancaire (PNB) :

En observant le PNB tout au long de la période [2016-2021], nous pouvons remarquer que la banque a presque doublé ses chiffres. Au profit de l'année 2021, malgré l'augmentation notable du total des charges opératoires de 18,2%, la BNA a enregistré un PNB de 829 MD soit une progression de 20,1% par rapport à l'année 2020.

Ceci peut être expliqué par l'amélioration des produits d'exploitation bancaire qui ont atteint un taux de progression de 10,5%.

Graphique n°1 : Evolution du PNB de la BNA [2016-2021]



Source : élaboré par l'auteur à partir des rapports annuels de la BNA.

II. Indicateurs de rentabilité :

Malgré l'amélioration continue du PNB, la BNA annonce toujours une baisse du résultat net dès 2018 pour enregistrer une baisse d'à peu près 17% au terme de l'année 2020 expliquée généralement par l'augmentation du coût de risque.

Cette baisse considérable du résultat net de la banque a engendré une dégradation de la rentabilité des actifs et par conséquent de sa capacité à générer des revenus importants à partir de ses ressources. D'ailleurs, nous pouvons remarquer que la rentabilité des capitaux propres a suivi la même tendance durant notre période de référence.

Tableau n°3 : Indicateurs de rentabilité de la BNA en [2016-2021]

	2016	2017	2018	2019	2020	2021
RN	104	199	175	123	102	163

ROE	28,52%	34,6%	23,6%	11,2%	7,2%	10,8%
ROA	1,57%	2%	1,6%	1%	0,7%	1%

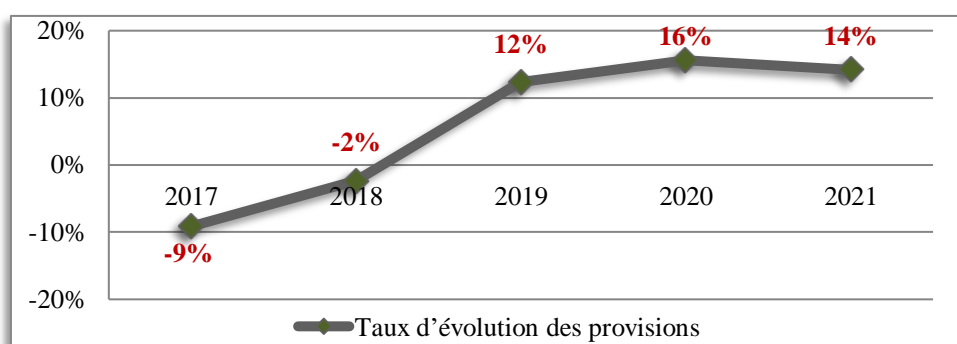
Source : élaboré par l'auteur à partir des rapports annuels de la BNA.

Cependant, au profit de l'année 2021, la BNA a amélioré ses chiffres en enregistrant une augmentation notable du résultat net de 60% ce qui a engendré une amélioration de la rentabilité des actifs ainsi que celle des capitaux propres de la banque.

III. Indicateurs de risque de crédit :

Au cours de la période [2017-2021], la BNA a enregistré une progression importante du taux d'évolution des provisions constituées de 60%.

Graphique n°2 : Evolution du taux des provisions de la BNA en [2017-2021]

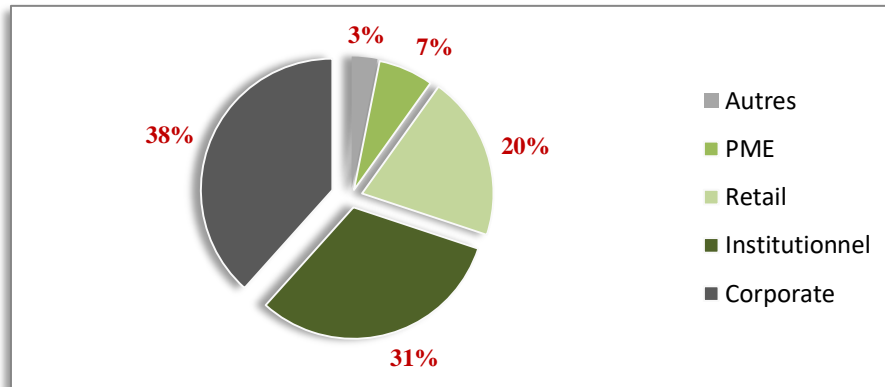


Source : élaboré par l'auteur à partir des rapports annuels de la BNA.

Ce gonflement du montant des provisions constituées peut être expliqué, entre autres, par l'augmentation des créances classées.

Par ailleurs, pour avoir une idée plus détaillée sur la décomposition du portefeuille des créances de la BNA, nous présentons la part de chaque catégorie dans le total des engagements de la banque en 2021 au niveau du graphique en secteur ci-dessous :

Graphique n°3 : Portefeuille des créances de la BNA par segment de clientèle en 2021



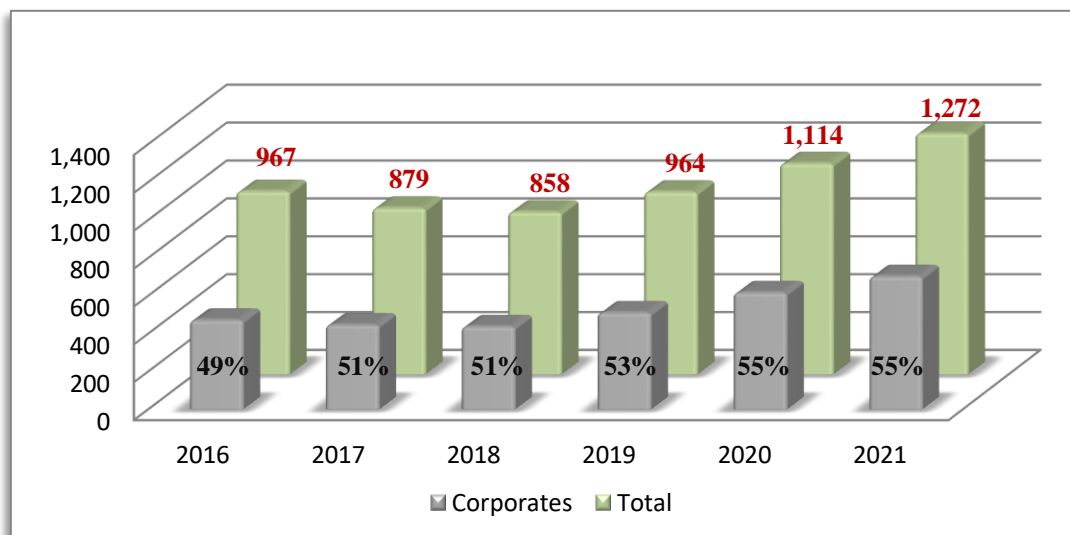
Source : élaboré par l'auteur à partir de la BDD⁷² de la BNA.

Avec : la catégorie « Autres » regroupe le segment des « banques et établissements financiers » et des « très petites entreprises » TPE car ils représentent une minorité.

Nous remarquons que le portefeuille des créances de la BNA est concentré sur le segment « Corporate » avec une part de 38% et sur celui des « institutionnels » avec une part de 31%.

En effet, le segment « Corporate » accapare la part la plus importante des provisions constituées par la banque pour la période [2016-2021].

Graphique n°4 : Provision du segment « Corporate » par rapport au total des provisions de la BNA



Source : élaboré par l'auteur à partir des données de la BNA.

⁷² La BDD a été obtenue auprès de la direction « surveillance de risque de crédit » de la BNA.

En observant le graphique ci-dessus, nous remarquons que les provisions constituées pour couvrir les créances de ce segment sont en évolution continue durant la période d'étude.

Donc, il est indispensable de mettre en place une stratégie de gestion de risque de crédit du portefeuille de la BNA en général, et du segment « Corporate » en particulier.

Section II : Méthodologie Empirique : Description des Données et Présentation des

Variables

Dans le but de réduire l'exposition de la banque au risque de crédit et d'améliorer le taux des créances classées, nous décidons de développer un modèle de scoring pour le segment « Corporate », qui permet d'évaluer la qualité des relations avant toute décision d'octroi de crédit ainsi que pendant la durée de l'engagement.

Au niveau de cette section, nous allons présenter en premier lieu notre méthodologie empirique en spécifiant notre échantillon d'étude ainsi que l'ensemble des variables à utiliser. Nous allons identifier en deuxième lieu, et à travers divers tests statistiques, les variables à retenir dans l'étape modélisation.

I. Méthodologie empirique :

1. Démarche empirique :

L'objectif de notre travail de recherche est d'élaborer une fonction de crédit scoring permettant de détecter la défaillance des relations avec précision. Pour arriver à cette finalité nous suivons la démarche empirique suivante :

La première étape consiste à collecter les données nécessaires pour l'élaboration de notre modèle. Avant de pouvoir utiliser ces données, nous procédons au pré nettoyage afin d'éliminer les valeurs manquantes et les valeurs extrêmes de notre base car elles peuvent nuire à la modélisation.

Nous passons par la suite à la sélection des variables significatives les moins corrélées et les plus discriminantes entre les entreprises saines et les entreprises défaillantes de notre échantillon en utilisant le test statistique de Student, de Mann-Whitney ainsi que de Khi-2.

Après l'identification des variables, nous procédons à la construction du modèle de crédit scoring. Au niveau de cette étape, nous modélisons à la fois la technique de l'analyse discriminante, de la régression logistique et des réseaux de neurones artificiels.

Une fois les résultats sont estimés, nous exploitons plusieurs critères pour vérifier la significativité des variables et pour juger la qualité globale des résultats obtenus.

En outre, nous nous assurons de la stabilité globale des modèles par le biais à la fois d'une validation « out of time » au niveau de laquelle nous utilisons un échantillon de testing d'une période différente et d'une validation « out of sample » au niveau de laquelle nous utilisons un échantillon de testing de groupe d'entreprises différent.

Comme nous avons présenté au niveau de la partie théorique, chacune des méthodes utilisées a ses propres avantages et limites. De ce fait, en se basant sur une analyse comparative entre les trois techniques, nous choisissons celle jugée la plus performante.

2. Constitution de l'échantillon d'étude :

Pour mener cette étude, nous constituons un échantillon de 355 entreprises. Il est à noter que l'échantillon a été sélectionné de manière aléatoire parmi les clients de la BNA appartenant au segment « Corporate » durant la période de [2019-2021].

Après la phase de nettoyage de la BDD et du près traitement des observations, nous ne gardons que 296 observations, dont :

- 156 entreprises pour la période de [2019-2020].
- 140 entreprises pour la période de [2020-2021].

Pour éviter le fameux problème de surapprentissage des techniques à utiliser, entre autres la technique d'intelligence artificielle, nous décortiquons notre échantillon d'étude en 2 sous-échantillons.

Selon Giudici (2003) la meilleure répartition pour élaborer un modèle de crédit scoring est de grader 75% pour l'apprentissage du modèle et 25% pour la validation de ce dernier d'une manière stratifiée, c'est-à-dire en gardant la même composition de notre échantillon initial (71,8% des relations sont saines et 28,2% des relations sont défailtantes).

En suivant la méthodologie empirique de Giudici (2003), nous répartissons la BDD finalement retenue en trois sous-échantillons comme suit :

Tableau n°4 : Répartition de la BDD initiale :

Echantillon [2019-2020] 156 entreprises		Echantillon [2020-2021] 140 entreprises
75% Apprentissage (117 entreprises)	Testing	
	25% Testing « out of sample » (39 entreprises)	100% Testing « out of time » (140 entreprises)

Source : élaboré par l'auteur.

3. Présentation des variables :

En se référant à l'ensemble des études empiriques antérieures et en fonction de la disponibilité des données, nous présentons au niveau de cette partie les différentes variables à utiliser par la suite pour la construction de notre modèle, telles que :

- Variable dépendante ou encore à expliquer : le passage en défaut.
- Variables indépendantes ou encore explicatives.

3.1. Variable dépendante : le passage en défaut :

En suivant la méthodologie de construction d'un modèle de crédit scoring recommandée par le comité de Bâle, nous définissons notre variable dépendante « Passage en défaut » comme étant l'événement de défaillance au 31/12/N+1 d'une entreprise jugée saine au 31/12/N.

Pour la discrimination entre les saines et les défaillantes, nous prenons ainsi la définition du comité présentée au niveau de notre partie théorique. En effet, toute entreprise est considérée comme défaillante si elle a enregistré un retard de paiement d'au moins 90j.

Par analogie, toute entreprise est considérée comme défaillante si elle est classée « classe 2 » ou plus au niveau du centrale des risques de la BCT et elle est jugée saine dans le cas échéant. En d'autres termes :

$$CR_n (\text{classe de risque}) = \begin{cases} 0 & \text{si classe BCT} = (0; 1) \\ 1 & \text{si classe BCT} = (2; 3; 4) \end{cases}$$

Avec : **n** représente l'indice de l'année.

Pour l'année N, nous ne prenons, dans notre base d'étude, que les entreprises saines, c'est-à-dire que celles ayant une CR=0.

De ce fait, notre variable « Passage en défaut » catégorielle binaire est présentée comme suit :

$$Y(\textit{Passage en défaut}) = \begin{cases} 0 & \textit{si } CR_{N+1} = 0 \\ 1 & \textit{si } CR_{N+1} = 1 \end{cases}$$

3.2. Variables indépendantes :

Dans le but de déterminer une fonction de score permettant de prédire avec précision le passage en défaut des relations de la BNA, nous essayons de traiter le maximum de variables qui peuvent refléter la situation financière de la relation. En effet, les variables indépendantes sont réparties entre :

- Ratios financiers.
- Variables comportementales.
- Variables qualitatives.

Remarque : les variables qualitatives ne sont pas nécessairement catégorielles, en effet nous désignons par qualitative les variables pouvant apporter une indication sur les qualités de la relation.

L'ensemble des variables indépendantes utilisées sont présentées au niveau du tableau ci-dessous.

Tableau n°5 : Liste des variables indépendantes :

Variable	Type	Signification	Formule	Signe attendu
Ratios financiers				
Equilibre financier				
R1	Métrique	Trésorerie nette (TN)	$FDR - BFR$	-
R2	Métrique	Ratio de l'équilibre financier	$\frac{FDR}{BFR}$	-
R3	Métrique	Fonds de roulement en jour du CA	$\frac{FDR}{CA} * 360$	-
R4	Métrique	Besoin en fonds de roulement en jour du CA	$\frac{BFR}{CA} * 360$	+
Rentabilité				
R5	Métrique	Rentabilité des capitaux propres (ROE)	$\frac{RN}{\text{Capitaux propres}}$	-
R6	Métrique	Rentabilité des actifs (ROA)	$\frac{RN}{\text{Total actif}}$	-
R7	Métrique	Capacité d'autofinancement (CAF)	EBE + produit encaissable - charge décaissable	-
R8	Métrique	Ratio de couverture des charges financières	$\frac{EBE}{\text{Charges financières}}$	-
Structure financière				

R9	Métrique	Ratio d'autonomie financière	$\frac{\text{Capitaux propres}}{\text{Capitaux permanent}}$	-
R10	Métrique	Ratio de levier	$\frac{\text{Total Dettes}}{\text{Capitaux propres}}$	+
Solvabilité				
Ratios d'endettement				
R11	Métrique	Ratio d'endettement global	$\frac{\text{Total Dettes}}{\text{Total Bilan}}$	+
R12	Métrique	Ratio d'endettement structurel	$\frac{\text{Total passifs non courants}}{\text{Total bilan}}$	+
R13	Métrique	Ratio d'endettement net	$\frac{\text{Dettes} - \text{liquidités}}{\text{Capitaux propres}}$	+
R14	Métrique	Ratio de la capacité de remboursement	$\frac{\text{Dettes nettes}}{\text{CAF}}$	+
Ratios de liquidité				
R15	Métrique	Ratio de liquidité générale	$\frac{\text{Actifs Circulants}}{\text{Dettes CT}}$	-
R16	Métrique	Ratio de liquidité immédiate	$\frac{(\text{Actifs Circulants} + \text{stock})}{\text{Dettes CT}}$	-
Ratio Solvabilité				
R17	Métrique	Ratio de solvabilité générale	$\frac{\text{Capitaux propres}}{\text{Total bilan}}$	-
Activité				

R18	Métrique	Variation de l'EBE	$\frac{EBE_N - EBE_{N-1}}{EBE_{N-1}}$	-
R19	Métrique	Variation du CA	$\frac{CA_N - CA_{N-1}}{CA_{N-1}}$	-
R20	Métrique	Marge bénéficiaire nette	$\frac{RN}{CAHT}$	-
R21	Métrique	Ratio d'activité	$\frac{CAF}{CAHT}$	-
Variables qualitatives				
Q1	Métrique	Taux des provisions sectorielles	$\frac{Total\ provisions\ par\ secteur}{Total\ provisions\ BNA}$	+
Q2	Catégorielle (binaire)	Appartenance à un groupe	$\begin{cases} 0 & \text{si l'entreprise ne fait pas partie d'un groupe} \\ 1 & \text{si l'entreprise fait partie d'un groupe} \end{cases}$	
Q3	Métrique	Taille de la relation	$\log(Total\ bilan)$	
Q4	Métrique	Ancienneté de la relation	$N - \text{année d'ouverture du compte}$	
Variables comportementales				
C1	Métrique	Taux d'impayé par engagement	$\frac{Montants\ impayés}{Total\ engagements\ BNA}$	+
C2	Métrique	Mouvement débiteur	-	+

Source : élaboré par l'auteur.

Remarque : la BDD des clients a été obtenue auprès de la direction « Modèle et Analyse des Risques » de la BNA.

Section III : Etude des Pouvoirs Discriminants : Sélection Préalable des Variables

Pour construire un modèle de scoring permettant la détection préalable des mauvaises relations, il est indispensable d'effectuer un certain nombre de tests statistiques sur les variables initialement retenues.

Cette étape est la plus importante, il s'agit principalement de savoir quelle est la batterie de variables qui va être utilisée pour la classification et la séparation entre les deux groupes d'entreprises.

De ce fait, nous allons identifier, en premier lieu, les variables significatives qui expliquent plus la défaillance des relations étudiées et nous allons éliminer, en second lieu, les variables corrélées entre elles.

I. Etude d'association des variables : analyse bivariée :

1. Variables quantitatives et le passage en défaut : Test de Student :

Notre objectif est d'identifier les variables quantitatives ayant une relation de dépendance avec la défaillance des entreprises, en d'autres termes, préciser les variables qui participent à la discrimination entre les deux groupes d'entreprises étudiées.

Puisque notre variable qualitative est dichotomique (présentant deux modalités : entreprise défaillante et entreprise non-défaillante), nous utilisons, pour tester la différence significative des moyennes des deux groupes, le test de Student afin de vérifier la présence ou l'absence d'une relation de dépendance entre une variable qualitative : « le passage en défaut » et d'autres quantitatives : les ratios financiers et les variables comportementales. Le test d'hypothèse est comme suit :

$$\begin{cases} \mathbf{H_0} : \bar{X}_0 = \bar{X}_1 \rightarrow \text{Les moyennes sont égales.} \\ \mathbf{H_a} : \bar{X}_0 \neq \bar{X}_1 \rightarrow \text{Les moyennes ne sont pas égales.} \end{cases}$$

Avec :

- \bar{X}_0 désigne la moyenne du groupe des entreprises saines et \bar{X}_1 celle des entreprises défaillantes telle que :

$$\bar{X}_k = \frac{\sum_{p=1}^{n_k} X_p}{n_k}$$

- n_k représente le nombre des relations appartenant au groupe k : {0 ; 1}.

Le résultat du test repose sur la valeur de « t-statistique », en effet, plus les moyennes sont éloignées les unes des autres, plus la valeur $|t|$ est grande donc plus la probabilité de rejeter l'hypothèse nulle d'inexistence d'une relation significative entre le ratio et le « passage en défaut » est importante.

$$t - \text{statistique} = \frac{\bar{X}_0 - \bar{X}_1}{\sqrt{\frac{S^2}{n_0} + \frac{S^2}{n_1}}}$$

Avec :

- S^2 représente la variance commune aux deux groupes, dont la formule de calcul est la suivante :

$$S^2 = \frac{\sum_{i=1}^{n_0} (x - \bar{X}_0)^2 + \sum_{j=1}^{n_1} (x - \bar{X}_1)^2}{n_0 + n_1 - 2}$$

Les résultats de ce test effectué par le logiciel SPSS V26, avec un risque d'erreur de 5%, montrent la présence d'une relation d'association entre les variables présentées dans le tableau ci-dessous et notre variable qualitative (passage en défaut).

Tableau n°6 : Résultats du test de Student⁷³

	T-statistique	P-value
R1	3,063	0,004
R7	2,685	0,008
R11	-2,176	0,034
R12	-2,158	0,035
R15	-5,755	0,000
R21	3,321	0,001
C1	-6,129	0,000
C2	-2,674	0,010

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Parmi les conditions d'utilisation de cette analyse, nous trouvons la normalité des variables. En effet les ratios suivent rarement des lois connues, telle que la loi normale. Donc, la

⁷³ Voir annexe n°2.1.

comparaison des moyennes par catégorie est insuffisante pour juger l'existence ou non d'une relation d'association entre les variables.

Pour cette raison, nous suivons la démarche de Bardos et Zhu (1997) en renforçant notre analyse par le test de Mann-Whitney, aussi appelé Wilcoxon-Mann-Whitney, qui représente une alternative non-paramétrique au test de Student dans le cas où l'hypothèse de normalité n'est pas vérifiée.

2. Variables quantitatives et le passage en défaut : Test de Mann-Whitney :

Ce test nous permet de déterminer le nombre de fois qu'un score du groupe 0 est supérieur à celui du groupe 1, et inversement. Il nous permet de vérifier si, sur la base des rangs des échantillons, nous pouvons estimer que les échantillons sont identiques en termes de position ou non. Le test d'hypothèse est comme suit :

$$\left. \begin{array}{l} \text{H}_0: P(X_0 > X_1) = P(X_1 > X_0). \\ \text{H}_a: P(X_0 > X_1) \neq P(X_1 > X_0). \end{array} \right\}$$

En d'autres termes, sous H_0 , nous supposons que la probabilité qu'une observation X_0 de la population 0 soit supérieure à une observation X_1 de la population 1 est égale à la probabilité qu'une observation de la population X_1 soit supérieure à une observation de la population X_0 .

Autres que les variables identifiées à travers la comparaison des moyennes, le test de Mann-Whitney montre qu'il existe ainsi une relation d'association entre le passage en défaut et les ratios présentés au niveau du tableau suivant avec un niveau de confiance de 95%.

Tableau n°7 : Résultats du test de Mann-Whitney⁷⁴

	Z	P-value
R14	-1,995	0,046
R15	-3,097	0,002
R18	-3,865	0,000
R20	-2,607	0,009
Q1	-2,948	0,003

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

⁷⁴ Voir annexe n°2.3.

A partir d'une batterie de 26 variables métriques initialement traitées, nous gardons seulement au niveau de cette phase 13 indicateurs répartis entre 10 ratios financiers, 2 variables comportementales et une qualitative que nous jugeons les plus discriminants entre les deux groupes d'entreprises.

Nous passons maintenant à vérifier l'existence d'une relation d'association entre la variable qualitative « appartenance à un groupe » et « le passage en défaut » à l'aide du test de Khi-2.

3. Variable qualitative et le passage en défaut : Test de Khi-2 :

La dépendance entre les deux variables catégorielles signifie que la valeur prise par la variable « appartenance à un groupe » nous donne une information sur le passage en défaut de la relation étudiée. Le test d'indépendance de khi-2 sert à déterminer l'existence d'une telle relation.

- ⌈ **H₀** : absence de relation : indépendance des deux variables
- ⌋ **H_a** : existence de relation : dépendance des deux variables.

Tableau n°8 : Résultats du test de Khi-2⁷⁵

	Khi-2 statistique	P-value
Q2	0,218	0,641

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

A partir du résultat présenté dans le tableau ci-dessus, nous ne pouvons pas rejeter l'hypothèse nulle d'indépendance des deux variables étudiées (P-value > 5%). Donc, nous concluons que le passage en défaut des entreprises ne dépend pas de l'appartenance à un ou de la non-appartenance de ces dernières à un groupe.

Après avoir identifié les variables financières, comportementales et quantitatives ayant une forte relation statistiquement significative avec le passage en défaut, nous passons maintenant à l'étape suivante pour étudier la corrélation entre les variables retenues.

⁷⁵ Voir annexe n°2.5.

II. Etude de dépendance des variables :

1. Matrice de corrélation :

Il convient de signaler que l'une des conditions requises pour la construction d'une fonction de score est que les variables retenues ne soient pas corrélées. En effet, des variables liées apportent en réalité la même information et sont redondantes.

De ce fait, nous choisissons le test de corrélation pour étudier la significativité ainsi que la force de la relation de corrélation entre les variables retenues afin de garder que des variables non corrélées.

En effet, le coefficient de corrélation Rho « ρ » prend une valeur comprise entre 1 et -1, lorsque $|\rho|$ tend vers 1, nous estimons qu'il existe une forte relation (positive ou négative) entre les deux variables. Par conséquent, nous devons éliminer celle que nous considérons comme moins importante pour prédire la défaillance des relations. Le test d'hypothèse est comme suit :

$$\left. \begin{array}{l} \mathbf{H}_0: \rho = 0 \rightarrow \text{il n'existe pas de corrélation entre les deux variables.} \\ \mathbf{H}_a: \rho \neq 0 \rightarrow \text{il existe de corrélation entre les deux variables.} \end{array} \right\}$$

Règle de décision : une relation de corrélation est considérée comme étant forte, à partir d'une valeur égale à ± 0.5 .

Tableau n°9 : Ratios financiers corrélés⁷⁶

		R7	R11	R18
R11	Rho Spearman	0,708***	1,000	-
	P-value	0,000	-	-
R12	Rho Spearman	0,696***	0,991***	-
	P-value	0,000	0,000	-
R15	Rho Spearman	0,683***	-	-
	P-value	0,000	-	-
R20	Rho Spearman	-	-	0,561***
	P-value	-	-	0,000

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

⁷⁶ Voir annexe n°3.1.

Le tableau n°9 regroupe les couples des ratios financiers ayant statistiquement une relation de corrélation significative.

2. Eta² : Mesure d'association :

Notre objectif est d'éliminer les sources de corrélation tout en maximisant l'information expliquée par les ratios retenus finalement pour la modélisation d'une fonction de score.

Pour arriver à cet objectif, nous essayons de connaître l'origine de la variation du passage en défaut à travers la mesure d'association avec les ratios financiers corrélés. Nous utilisons l'effet taille « Eta² », cette mesure prend des valeurs comprises entre 0 (aucune association) et 1 (forte association) dont la formule de calcul est la suivante :

$$Eta^2 = \frac{SSB}{SST}$$

Avec :

- **SSB** (sum of squares between) représente la somme des carrés entre les moyennes de chaque groupe et la moyenne totale de l'échantillon. En d'autres termes, elle mesure la variabilité entre les groupes étudiés.
- **SST** (sum of squares total) représente la somme des carrés entre chaque point de données et la moyenne totale de l'échantillon. Autrement dit, elle quantifie la variabilité totale des données observées.

Donc, Eta² peut être formulée comme suit :

$$Eta^2 = \frac{\sum_{p=1}^{n_k} (\bar{X}_k - \bar{X})^2}{\sum_{i=1}^{n_0} \sum_{j=1}^{n_1} (X_{ij} - \bar{X})^2}$$

Pour résumer, Eta² représente la proportion de variance de la variable de groupe expliquée par la variable testée. Cela nous permet de comparer les effets entre les différents ratios.

Les résultats du test sont présentés au niveau du tableau suivant :

Tableau n°10 : Eta² : Mesure de l'effet de taille⁷⁷

	R7	R8	R11	R12	R15	R20
Eta ² en %	46,65	92,2	84,09	41,99	76,91	56,1

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

En comparant la valeur d'Eta² en pourcentage de R11 et R12 nous remarquons que R11, représentant « le ratio d'endettement global », explique près de 84% de la variabilité du « passage en défaut » ce qui reflète une forte relation d'association par rapport à R12 « le ratio d'endettement structurel ».

→ Donc nous décidons d'éliminer R12 afin de garder R11.

Pour R7, R11 et R15, nous remarquons que les deux derniers ratios ont une importante relation avec le passage en défaut par rapport à R7.

→ De ce fait, parmi les 3 ratios étudiés, nous décidons de retenir R11 et R15.

Finalement, entre R8 et R20, nous remarquons que R8 représentant « le ratio de couverture des charges financières » explique environ 92% du passage en défaut des relations étudiées, ce qui est largement supérieur à la valeur d'Eta² de R20 représentant « le rapport entre le résultat net et le chiffre d'affaires ».

→ Donc, nous décidons d'éliminer R20 afin de garder R8.

En conclusion, à travers l'une étude comparative entre les variables corrélées, nous gardons les ratios R8, R11 et R15 parmi les variables retenues finalement pour la modélisation d'une fonction de score.

Après avoir identifié les variables satisfaisantes les conditions prédéfinies, il nous reste qu'une seule hypothèse à vérifier avant de passer à la modélisation d'une fonction de score, c'est la multi-colinéarité.

III. Test de multi-colinéarité :

En présence de multi-colinéarité, les coefficients individuels de chaque ratio financier ne peuvent être ni stables ni interprétés de façon fiable. Pour détecter les variables multi-colinéaires, nous étudions les facteurs d'inflation de la variance **VIF** (Variance Inflation Factor), telle que :

⁷⁷ Voir annexe n°3.2.

$$\text{VIF} = \frac{1}{\text{Tolérance}} \quad \text{et} \quad \text{Tolérance} = 1 - R_i^2$$

Avec : R_i^2 représente le coefficient de détermination linéaire de la variable i.

Règle de décision : en règle générale, le VIF de chaque variable doit être inférieur à la valeur critique de 5. Par conséquent, un VIF supérieur à 5 ou une valeur de tolérance inférieure à 0,2, signifient que le ratio est une combinaison linéaire d'autres variables indépendantes.

Tableau n°11 : Mesure de multi-colinéarité

	Tolérance	VIF	Inflation en %
R1	0,822	1,216	21,6
R8	0,990	1,010	1
R10	0,937	1,067	6,7
R11	0,772	1,296	29,6
R14	0,912	1,096	9,6
R15	0,718	1,392	39,2
R21	0,777	1,287	28,7
Q1	0,939	1,064	6,4
C1	0,858	1,166	16,6
C2	0,831	1,203	20,3

Source : élaboré par l'auteur à partir du logiciel XLSTAT.

Comme les coefficients de tolérance sont largement supérieurs à 0,2, les VIF sont inférieurs à 5 ainsi qu'aucune variable n'a pas présenté une inflation de variance jugée importante. Alors, le problème de la multi-colinéarité ne se pose pas dans notre échantillon.

Au niveau de la présente partie, nous établissons la liste des variables prédictives, considérées comme les plus discriminantes entre le groupe des entreprises défailtantes et celui des entreprises saines ainsi que les moins corrélées, afin de procéder à la modélisation telle que présentée ci-dessous :

Tableau n°12 : Liste des variables retenues

Variable	Formule	Signe attendu
R1 : Trésorerie nette	$FDR - BFR$	-
R8 : Ratio de couverture des charges financières	$\frac{EBE}{Charges\ financières}$	-
R10 : Ratio de levier	$\frac{Total\ Dettes}{CP}$	+
R11 : Ratio d'endettement global	$\frac{Total\ Dettes}{Total\ Bilan}$	+
R14 : Ratio de la capacité de remboursement	$\frac{Dettes\ nettes}{CAF}$	+
R15 : Ratio de liquidité générale	$\frac{Actifs\ Circulants}{Dettes\ CT}$	-
R21 : Ratio d'activité	$\frac{CAF}{CAHT}$	-
Q1 : Taux des provisions sectorielles	$\frac{Total\ provisions\ par\ secteur}{Total\ provisions\ BNA}$	+
C1 : Taux d'impayés par engagements	$\frac{Montants\ impayés}{Total\ engagements\ BNA}$	+
C2 : Mouvement débiteur	-	+

Source : élaboré par l'auteur.

Conclusion :

Au niveau de ce chapitre, nous avons mis l'accent sur l'importance de l'élaboration d'un modèle de scoring au sein de la BNA permettant d'évaluer les relations du segment « Corporate ». De ce fait, notre étude empirique sera consacrée pour ce segment.

De plus, nous avons adopté toute une démarche méthodologique pour étudier, comparer et sélectionner les variables indépendantes ayant statistiquement un pouvoir discriminant entre le groupe des entreprises saines et le groupe des entreprises défailtantes.

En effet, parmi un ensemble de 26 variables réparties entre 21 indicateurs financiers, 4 variables qualitatives et 2 comportementales initialement retenues, nous n'avons retenu que 13 variables indépendantes pour la construction du modèle, telles que :

- ✓ 7 variables reflétant la situation financière,
 - ✓ 1 variable qualitative,
 - ✓ 2 variables comportementales.
-

Introduction :

La modélisation du risque de crédit consiste à évaluer la probabilité qu'un emprunteur ne respecte pas ses obligations de remboursement. Elle permet aux établissements de crédit d'évaluer et de gérer efficacement le risque de défaut de leurs emprunteurs.

Par ailleurs, l'établissement d'un modèle de scoring permet de standardiser et d'automatiser le processus d'évaluation du risque de crédit, donc de prendre des décisions de crédit plus pertinentes, plus rapides et moins coûteuses par rapport à d'autres méthodes d'évaluation du risque telle que l'analyse financière traditionnelle.

Pour parvenir à un résultat plus fiable, nous assistons à une panoplie des méthodes d'évaluation du risque de crédit visant le même objectif, à savoir l'identification des relations défaillantes, par conséquent, la prise d'une décision adéquate d'accord ou de refus d'octroi de crédit.

Au niveau de ce chapitre, nous allons modéliser le risque de passage en défaut des entreprises par les techniques de scoring. Par la suite, nous allons employer de multiples critères de performance en vue d'identifier le modèle le plus appropriés aux données de la banque.

Section I : Costruction d'un modèle de « Crédit Scoring »

Il existe plusieurs méthodes de crédit scoring, chacune utilise différentes techniques et algorithmes. Cependant, le principe est le même, il s'agit de modéliser une fonction de score permettant de déterminer la probabilité de défaut des demandeurs de crédit.

Ces techniques de crédit scoring sont mises en place à partir de trois principales formes de modélisation multivariée :

- ✓ Scoring par le modèle d'analyse discriminante.
- ✓ Scoring par la régression logistique.
- ✓ Scoring par les réseaux de neurones.

Au niveau de la présente section, nous allons utiliser ces trois fameuses techniques pour déterminer une fonction de score permettant d'évaluer le risque de crédit de notre échantillon.

I. Scoring par l'analyse discriminante factorielle (ADF) :

Pour la construction du modèle basé sur l'analyse discriminante, nous suivons une démarche de deux étapes : au début, nous allons estimer les coefficients de chaque variable et déduire une fonction de score permettant de discriminer entre les deux groupes d'entreprises. Ensuite, nous allons vérifier la validité globale des résultats de l'analyse discriminante.

1. Construction du modèle ADF :

1.1. Fonction discriminante⁷⁸ :

Le traitement de notre base de données par le logiciel SPSS V26 nous a conduit à identifier la fonction de discrimination de sorte que les scores soient à la fois très différents entre les groupes et très similaires au sein du même groupe.

$$\text{Score (X)} = -0,37 - 0,34 R1 - 0,08 R8 + 0,14 R10 + 0,51 R11 - 0,09 R14 - 0,30 R15 \\ - 0,42 R21 - 0,05 Q1 + 0,85 C1 + 0,33 C2$$

Cette fonction permet de déterminer le score discriminant individuel de chaque relation X. Nous remarquons que plus le score de la relation étudiée est élevé plus elle est risquée pour la banque.

Pour chaque groupe de notre échantillon, nous déterminons les centroïdes représentant les points des scores moyens obtenus en remplaçant les moyennes de chaque groupe de toutes les variables de notre fonction discriminante. Ces points, présentés au niveau du tableau suivant, sont appelés centres de gravité ou encore barycentres des groupes.

Tableau n°13 : Fonction 1 aux centroïdes des groupes

0 (entreprises saines)	-0,833
1 (entreprises défaillantes)	+1,980

Source : élaboré par l'auteur, à partir du logiciel SPSS V26.

Pour affecter les entreprises à l'un de deux groupes, nous devons comparer le score (X) obtenu aux barycentres 0 et 1. Par la suite, nous allons classer les observations dans le groupe où elles sont les plus proches.

⁷⁸ Voir annexe n° :4.4.

1.2. Fonctions de classification⁷⁹ :

De plus, nous pouvons déterminer les fonctions de classification relatives à chaque groupe de notre échantillon. En effet, chaque entreprise est classée dans le groupe où elle a obtenu le score le plus élevé. Les fonctions de classement sont les suivantes :

$$\begin{aligned} \text{Score } (X_0) &= -2,44 - 1,16 R1 + 0,14 R8 + 0,72R10 - 0,03 R11 + 0,04 R14 + 2,81 R15 \\ &\quad - 18,75 R21 + 0,18 Q1 - 0,04 C1 - 0,05 C2 \\ \text{Score } (X_1) &= -6,01 + 3,19 R1 - 0,03 R8 + 0,85 R10 + 0,32 R11 + 0,07 R14 + 1,63 R15 \\ &\quad - 12,46 R21 + 0,06 Q1 + 0,28 C1 + 0,1 C2 \end{aligned}$$

Cependant, il est à noter qu'il existe une zone d'incertitude entre les deux centroïdes des deux groupes (-0,833 et +1,980). Au niveau de cette zone, nous ne pouvons pas prononcer définitivement sur le passage en défaut ou non des relations, c'est le cas des dossiers tangents.

2. Analyse d'efficacité du modèle :

Pour garantir une bonne qualité prédictive de la fonction de score, nous procédons à une double validation globale du modèle en utilisant les tests hypothèses probabilistes d'une part, la matrice de confusion et le taux du bon classement d'autre part.

2.1. Tests d'hypothèses probabilistes :

Concernant les premiers tests, nous utilisons principalement la corrélation canonique et Lambda de Wilks. En effet, nous présentons, Au niveau du tableau suivant, les différentes statistiques relatives à la validité globale du modèle.

Tableau n°14 : Statistiques de validité globale de la fonction de score

Corrélation canonique	Lambda de Wilks	Khi-2	P-value
0,791	0,374	116,114	0,000

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

La corrélation canonique représente une des mesures du pouvoir discriminant des facteurs, elle prend des valeurs comprises entre 0 et 1. En effet, plus sa valeur est proche de 1, plus la qualité du modèle est jugée bonne. Dans notre cas, la corrélation canonique est égale à 0,791, cette valeur confirme un pouvoir discriminant assez important de notre fonction discriminante.

⁷⁹ Voir annexe n° :4.7.

De même, le lambda de Wilks permet de tester si les groupements obtenus ayant des moyennes significativement différentes. C'est le rapport du déterminant de la matrice de variances covariances intra-groupe au déterminant de la somme des matrices de variances covariances inter et intra-groupe.

Il prend ainsi des valeurs entre 0 et 1, en fait, il n'existe pas un niveau bien déterminé qu'à partir duquel notre lambda indique que les deux classes sont suffisamment différentes, tout dépend du nombre de groupes, de l'échantillon et du nombre de variables utilisées. Mais une faible valeur du lambda statistiquement significative, telle que 0,374, reflète une faible variation intra-classe, par conséquent de forte variation inter-classes, d'où une différence significative des moyennes des classes.

2.2. Matrice de confusion⁸⁰ :

Pour s'assurer que notre fonction discriminante permet de classer les entreprises en sous-groupes avec précision, nous employons la matrice de confusion afin de déterminer le taux de bon classement de notre modèle.

Cet outil d'évaluation présente les résultats de réaffectation de notre échantillon de construction à l'aide de la fonction déduite comme suit :

Tableau n°15 : Matrice de confusion (ADF échantillon d'apprentissage)

De \ Vers	0	1	Total	% Correct	% Faux
0	110	2	112	98,2%	1,8%
1	16	28	44	63,6%	36,4%
Total de classement			156	88,5%	11,5%

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Au total, le modèle réalise un taux de bon classement de 88,5%, donc un taux d'erreur de 11,5%, ce résultat reflète une qualité moyenne de précision de la fonction utilisée.

Cependant, en décortiquant ces 11,5% nous trouvons qu'ils sont dus, principalement, aux faux positifs. En effet, le pourcentage d'erreur de type I représente 36,4% par contre celui du type II est de 1,8%.

Nous pouvons conclure donc que notre modèle offre une forte discrimination entre les deux groupes d'entreprises, mais, il montre une difficulté au niveau de détermination

⁸⁰ Voir annexe n° :4.8

des relations défailtantes. Cette limite peut être expliquée par le fait que notre échantillon d'apprentissage est composé seulement de 28,2% des dossiers défailtants donc, il n'a pas appris à spécifier les vrais négatifs.

II. Régression logistique (RL) :

Du fait que notre variable de réponse est catégorielle binaire, nous utilisons au niveau de cette partie la régression logistique binaire afin d'élaborer un modèle prédictif permettant de mesurer l'association entre la survenance de l'événement « passage en défaut » et les variables préalablement sélectionnées. Par la suite, nous jugeons la significativité globale du modèle établi et des variables retenues en faisant appel à des tests statistiques. Finalement, nous évaluons la qualité de prédiction du modèle à travers la matrice de confusion et le taux de bon classement.

1. Construction du modèle :

1.1. Fonction de score :

Le traitement de notre base de données sous le logiciel SPSS V26, nous permet d'estimer les coefficients des variables indépendantes de la régression logistique.

Pour évaluer la significativité statistique de ces coefficients, nous nous basons sur la statistique de Wald, dont le test d'hypothèse est comme suit :

$$\begin{cases} H_0 : \beta_p = 0 \\ H_a : \beta_p \neq 0 \end{cases}$$

Avec : β_p représente le coefficient estimé de la variable indépendante p .

Remarque : afin de maximiser l'information expliquée par les variables et d'obtenir un modèle couvrant les différentes caractéristiques des entreprises, nous acceptons le taux de provision sectoriel avec un niveau de confiance de 90%.

Tableau n°16 : Variables significatives de la RL⁸¹

	R1	R10	R11	R15	R21	Q1	C1	Constante
Coefficient	-0,13	0,08	0,26	-0,31	-0,24	0,22	0,69	-2,68
P-value	0,019	0,016	0,05	0,002	0,002	0,073	0,02	0,038

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

⁸¹ Voir annexe n° : 5.2.

Nous trouvons que parmi les 10 variables initialement traitées seulement 7 permettent de prédire significativement la probabilité du passage en défaut des entreprises avec le terme constant.

Donc, sous le critère de maximum de vraisemblance, notre fonction s'écrit comme suit :

$$\text{Logit } P(y = 1) = -2,68 - 0,13 R1 + 0,26 R11 - 0,31 R15 + 0,08 R10 - 0,24 R21 + 0,22 Q1 + 0,69 C1$$

A partir de cette fonction, nous déduisons la fonction de probabilité de défaut :

$$PD = \frac{1}{1 + e^{-\text{Logit } P(y=1)}}$$

Dont la règle d'affectation de l'individu i est la suivante :

$$\begin{cases} Y_i = 1 \text{ si } PD_i > 0,5 \\ Y_i = 0 \text{ sinon} \end{cases}$$

1.2. Analyse de force des prédicteurs :

Chaque variable explicative retenue dispose d'un pouvoir explicatif de la probabilité de survenance du défaut des relations. Pour étudier ce dernier, nous utilisons les « Odds ratio OR » ou encore les rapports de cotes.

Ce ratio mesure la force de l'association entre chaque variable explicative et la variable de réponse à travers la comparaison de la probabilité du passage en défaut $P(Y=1)$ et la probabilité $P(Y=0)$ lorsque nous introduisons la variable étudiée dans le modèle prédictif.

$$\text{Odds} = \frac{P(Y = 1)}{P(Y = 0)} = e^{\beta_p}$$

Règle de décision :

- **OR<1** : la PD diminue lorsque la variable augmente.
- **OR=1** : la variation de la variable n'affecte pas la PD.
- **OR>1** : la PD augmente lorsque la variable augmente.

Le tableau suivant regroupe les rapports de cotes de chaque variable explicative.

Tableau n°17 : Odds Ratio par Variables

	R1	R10	R11	R15	R21	Q1	C1
OR	0,879	1,083	1,298	0,733	0,787	1,245	1,994
Δ PD en %	-12,1	8,3	29,8	-26,7	-21,3	24,5	99,4

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Parmi les variables étudiées, C1 (taux d'impayés par engagements) représente le facteur de risque le plus important dans le modèle, suivi par R11 (ratio d'endettement global) et Q1 (taux des provisions sectorielles).

Donc l'augmentation de ces trois facteurs permet d'augmenter considérablement la probabilité de défaut de la relation. Par contre, une amélioration des deux ratios R15 et R21 réduit l'exposition des entreprises au risque de défaut, toute chose étant égale par ailleurs.

2. Analyse d'efficacité du modèle :

Une fois le modèle de scoring par la régression logistique est estimé, il convient de valider ses résultats et d'analyser sa performance. De ce fait, nous utilisons les tests d'hypothèse pour évaluer la qualité d'ajustement du modèle et nous employons la matrice de confusion afin de juger le degré de précision prédictive de ce dernier.

2.1. Tests d'hypothèse probabilistes :

Pour vérifier la significativité globale du modèle nous utilisons le test de Hosmer et Lemeshow. Ce test permet de juger la qualité d'ajustement globale du modèle à travers la comparaison entre l'affectation réelle et estimée des observations de notre échantillon et d'évaluer la distance entre les deux. Le test d'hypothèse de Hosmer et Lemeshow est comme suit :

- H_0 : Distance faible : il n'existe pas de différence significative entre l'observé et l'estimé.
- H_a : Distance élevée : il existe une différence significative entre les deux.

Tableau n°18 : Test de Hosmer et Lemeshow⁸²

Khi-2	P-value
7,545	0,273

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

La P-value est supérieure à 5%, donc nous acceptons l'hypothèse nulle d'inexistence d'une différence statistiquement significative entre l'affectation observée de notre échantillon et l'affectation estimée par le modèle retenu.

De plus, la valeur du coefficient de R² de Cox et Snell du modèle estimé est de 69%.

Cette mesure du pouvoir prédictif des facteurs est basée sur le rapport du log de vraisemblance de notre modèle (avec les prédicteurs choisis) au log de vraisemblance du modèle de la ligne de base (avec seulement une constante et sans variables explicatives).

Par rapport à des valeurs appartenant à [0 ; 1[, une valeur de 69% signifie qu'il existe une différence statistiquement significative lorsque nous utilisons les variables explicatives pour prédire le défaut futur des entreprises.

Une valeur de R² de Nagelkerke de 93% confirme ainsi nos résultats. Ce coefficient représente une mesure ajustée de R² de Cox et Snell expliquant la proportion de variance dans la variable dépendante associée aux variables indépendantes retenues dans le modèle. En d'autres termes, 93% des variabilités du risque du passage en défaut des entreprises étudiées sont expliquées par les variables retenues dans le modèle.

➔ Nous pouvons conclure qu'à un seuil de 95% de significativité, il s'agit d'une bonne qualité d'ajustement et que le modèle est adéquat pour prédire la probabilité du passage en défaut des relations.

2.2. Matrice de confusion⁸³ :

La matrice de confusion représente un outil pertinent pour approuver les résultats présentés au niveau de la partie précédente. Par rapport aux autres techniques de modélisation du risque de crédit, le traitement de la régression logistique sous SPSS

⁸² Voir annexe n°5.5.

⁸³ Voir annexe n°5.7.

V26 nous permet de comparer le rendement du modèle estimé et le modèle de la ligne de base pour avoir une idée sur l'apport des variables choisies.

Tableau n°19 : Matrice de confusion (RL échantillon d'apprentissage)

Modèle de la ligne de base (terme constant seulement)					
De \ Vers	0	1	Total	% Correct	% Faux
0	112	0	112	100%	0%
1	44	0	44	0%	100%
Taux de classement				71,8%	28.2%
Modèle estimé (prédicteurs retenus)					
De \ Vers	0	1	Total	% Correct	% Faux
0	110	2	112	98,2%	1,8%
1	5	39	44	88,6%	11,4%
Taux de classement				95,5%	4,5%

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Selon les résultats présentés ci-dessus, notre modèle affiche un important taux des prédictions correctes de 95,5%. Les variables retenues nous permettent d'améliorer le taux de bon classement du modèle de la ligne de base de 23,7% en général et le taux des prédictions correctes du sous-groupe 1 de 88,6% en particulier.

Cependant, le fait que le pourcentage d'erreurs de type I atteint 11,4% est considéré comme significatif et représente une limite pour notre modèle. En effet, il est toujours plus coûteux pour la banque de ne pas identifier les relations défaillantes que de ne pas saisir les opportunités des relations saines.

III. Scoring par les Réseaux de neurones artificiels (RNA)⁸⁴ :

L'objectif de cette partie est de développer une nouvelle technique de classification par les réseaux de neurones artificiels afin d'améliorer la prise de décision des « credit men » au sein de la banque.

Le principal avantage de cette méthode, par rapport à celles paramétriques traditionnelles présentées ci-dessus, c'est le fait qu'elle n'exige aucune hypothèse a priori de normalité, de linéarité ni que de type de relation entre les variables d'entrée et les variables de sortie.

⁸⁴Voir annexes n° :

Dans cette sous-partie, nous présentons, en premier lieu, l'architecture du réseau de neurones construit. Nous examinons, en second lieu, les divers paramètres à interpréter afin de vérifier la validité globale du modèle mis en œuvre.

1. Construction du modèle :

Parmi les différentes formes de RNA, nous choisissons le modèle le plus utilisé pour l'exploration prédictive des données : le « MultiLayer Perceptron » MLP. Comme son nom l'indique, il s'agit d'un réseau « feedforward » avec éventuellement une ou plusieurs couches cachées (fonction d'activation : hidden layer), une couche d'entrée (variables retenues dans le modèle : input layer) et une couche de sortie (passage en défaut : output layer), totalement interconnectées regroupant des unités dites neurones.

1.1. Architecture choisie :

Bien que le réseau neurone représente un des outils informatiques puissants d'analyse des données, il nécessite également une analyse exploratoire. En effet, il n'existe aucun théorème, aucune règle, aucune loi qui permette de choisir une architecture bien déterminée.

De ce fait, nous réalisons plusieurs itérations sur le réseau en faisant varier les différents paramètres, notamment :

- ✓ Le nombre de couches masquées,
- ✓ Le nombre de neurones cachés au niveau de chaque couche,
- ✓ Les variables de couches d'entrée.

Afin de sélectionner l'architecture optimale présentant à la fois le plus faible taux d'erreur ainsi que la plus simple structure à interpréter, Bardos et Zhu (1997).

Remarque : par rapport aux autres paramètres du modèle, tels que les fonctions d'activation de différentes couches, l'algorithme d'apprentissage ainsi que d'optimisation, nous choisissons l'option d'identification automatique du SPSS V26 qui nous permet d'identifier le meilleur modèle possible pour chaque combinaison des couches (d'entrée et masquée).

Étant donné que cette technique n'annonce aucune instruction sur les variables d'entrée, nous réintroduisons donc les ratios financiers R7, R12 et R20 éliminés à cause du problème de corrélation afin d'élargir nos choix et d'améliorer nos résultats.

Pour garantir un certain niveau de simplicité, nous optons pour un modèle à une seule couche masquée. L'architecture du modèle finalement retenue est comme suit :

Tableau n°20 : Architecture du RNA retenue⁸⁵

Couche d'entrée	Couche masquée	Couche de sortie
13 variables retenues telles que : <ul style="list-style-type: none"> ▪ 10 Ratios financiers, ▪ 2 variables comportementales, ▪ 1 variable qualitative Unité de biais ⁸⁶ .	7 neurones cachés. Unité de biais.	0 entreprise saine. 1 entreprise défailante.

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Le choix d'une telle architecture est pris selon le critère de minimisation d'erreur de prévision en appliquant le modèle sur notre échantillon. Le tableau n°21 présente l'erreur de prévision par sous-échantillons en pourcentage.

Tableau n°21 : Pourcentage de prévisions incorrectes par échantillon

Apprentissage	Validation
3,4%	0%

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Notre modèle montre une excellente qualité de prévision avec une certitude près de 97% pour l'échantillon d'apprentissage et de 100% pour l'échantillon de validation.

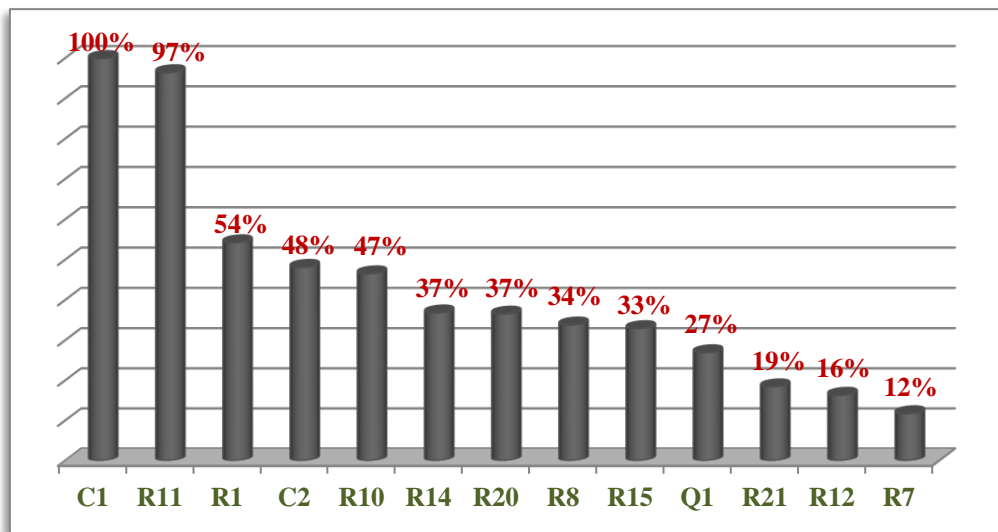
1.2. Analyse d'importance des prédicteurs :

Une des intéressantes analyses effectuées par le RNA, nous trouvons l'analyse d'importance des variables retenues dans la construction du modèle. Nous pouvons déterminer, alors, l'apport de chaque variable dans la prédiction de défaillance des entreprises.

⁸⁵ Voir annexe n°6.3.

⁸⁶ Le biais représente une unité d'auto-apprentissage du RNA dont la fonction d'activation égale toujours à 1.

Graphique n°5 : Importance en % par prédicteur



Source : élaboré par l'auteur à partir du logiciel SPSS V26.

En observant le graphique ci-dessus, nous trouvons que les principales variables participantes à la construction du modèle sont C1 (taux d'impayés par engagements) et R11 (ratio d'endettement global) par contre R7 (capacité d'autofinancement) et R12 (ratio d'endettement structurel) représentent les ratios les moins importants participants dans la construction du modèle.

Cette analyse confirme les résultats obtenus par la mesure d'association Eta^2 et celle effectuée à travers les OR de la régression logistique.

2. Analyse d'efficacité du modèle :

Afin de juger la validité du modèle, nous utilisons à la fois la matrice de confusion et le graphique des gains cumulé.

2.1. Matrice de confusion⁸⁷ :

Comme nous avons précisé au niveau de la partie précédente, notre modèle dispose d'une bonne capacité prédictive du défaut de 98,2%. Pour étudier plus cette capacité et afin de valider globalement le modèle retenu, nous étudions le taux de bon classement en faisant recours à la matrice de confusion pour l'échantillon d'apprentissage.

⁸⁷ Voir annexe n° :6.6.

Tableau n°22 : Matrice de confusion (RNA échantillon d'apprentissage)

De \ Vers	0	1	Total	% Correct	% Faux
0	84	0	84	100%	0%
1	2	31	33	93,9%	6,1%
Taux de classement			117	98,2%	1,8%

Source : élaboré par l'auteur à partir du logiciel SPSS V26.

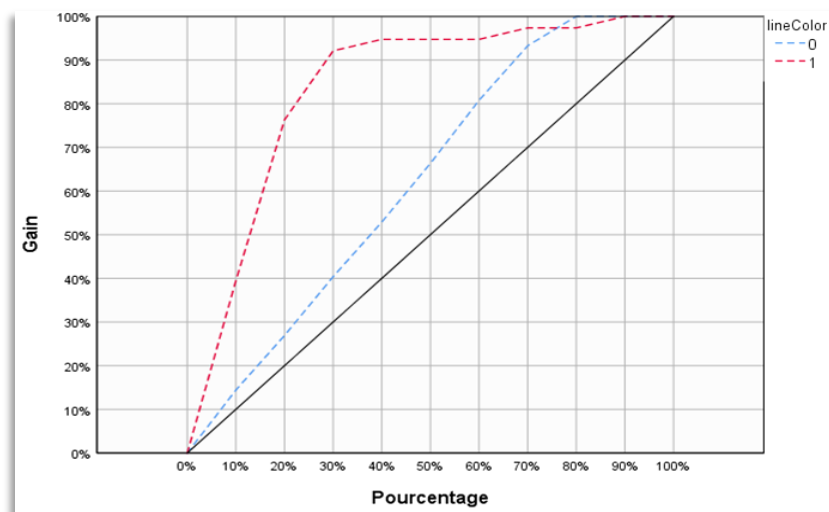
En observant la ligne de prévision des entreprises saines, nous pouvons estimer que notre modèle arrive à identifier ces dernières avec certitude. Par contre la ligne des entreprises défaillantes montre un pourcentage d'erreur type I de 6,1%. En conclusion, le modèle offre un important taux de bon classement, mais le fait de ne pas identifier les relations défaillantes parmi les relations saines représente un risque pour la banque.

2.2. Graphique des gains cumulés :

La courbe des gains cumulés représente une des outils d'évaluation de la performance de discrimination des modèles neuronaux. Elle indique le pourcentage des prévisions correctes par groupe d'entreprises par rapport au nombre des entreprises observées appartenant à ce groupe.

A l'aide de la 26^{ème} version du logiciel SPSS, nous pouvons visualiser les pourcentages des gains cumulés par centiles d'échantillon utilisé comme suit :

Graphique n°6 : Gains cumulés en %



Source : logiciel SPSS V26.

Remarque : la diagonale représente la performance obtenue si nous utilisons un jeu de hasard pour la classification des entreprises, c'est-à-dire un modèle ne délivrant aucune information (à 50% de précision).

D'une manière générale, la courbe d'un bon modèle présente une hausse rapide vers la valeur 100% puis elle se stabilise, c'est exactement le cas de notre modèle dont le pourcentage des gains cumulés est plus de 70% au niveau du 2^{ème} décile, en outre, à partir du 3^{ème}, il se stabilise entre la valeur de 90% et 100%.

En se basant sur les deux critères d'évaluation des modèles de classification présentés ci-dessus, nous pouvons conclure que malgré l'incapacité de prédire avec certitude les relations défailtantes, l'architecture du perceptron multicouche obtenue est jugée très performante pour la discrimination entre les bons et les mauvais clients.

Après avoir établi les modèles de scoring par les différentes approches :

- ✓ Paramétrique par l'analyse discriminante factorielle,
- ✓ Semi-paramétrique à travers la modélisation d'une régression logistique,
- ✓ Non-paramétrique en employant l'intelligence artificielle du MultiLayer perceptron.

Maintenant nous devons identifier celui le plus fiable.

———— Section II : Etude Comparative de Performance des Modèles ————

L'objectif de cette présente partie est d'évaluer la performance des différentes techniques utilisées précédemment en précisant les limites ainsi que les points forts de chacune.

Afin de bien apprécier la capacité des modèles à discriminer avec précision les entreprises de notre échantillon, nous allons évaluer leur performance dans le temps et dans l'espace. En d'autres termes, nous allons utiliser des échantillons d'entreprises différents de ceux qui ont servi à la construction des modèles pour tester leur stabilité de précision « out of sample » et « out of sample ».

Ainsi, nous allons recourir aux différents critères d'évaluation de la qualité de prédiction en vue de retenir l'approche offrant la meilleure qualité de prédiction de la défailtance.

I. Performance des modèles :

Au niveau du tableau ci-dessous, nous présentons les résultats de classement affichés par chaque modèle utilisé.

Nous trouvons que le RNA multilayer perceptron enregistre le taux de bon classement le plus important de 98,2% suivi par la RL avec un taux de 95,5% et l'ADF par le taux 88.5% occupe la dernière place.

Tableau n°23 : Résultats de classement des modèles

	Erreur de type I	Erreur de type II	Taux de bon classement
ADF	36,4%	1,8%	88,5%
RL	11,4%	1,8%	95,5%
RNA	6,1%	0%	98,2%

Source : élaboré par l'auteur.

→ Par rapport aux autres méthodes, le MLP affiche les taux d'erreur type I et type II les plus faibles. En effet, il prouve son excellente capacité prédictive des entreprises saines ainsi que défailtantes.

Enfin, il ne faut pas fonder notre évaluation de performance des modèles seulement sur le taux de bon classement. D'autres critères, tels que la stabilité des modèles dans le temps et dans l'espace, peut également être utiles pour avoir une vision complète de la performance du modèle.

Mais avant de passer aux tests de stabilité, nous allons employer, au niveau de la partie suivante, l'analyse ROC.

II. Analyse ROC :

Il est à noter que l'analyse ROC est un outil particulièrement fiable dans les cas où les 2 classes ne sont pas équilibrées et où le taux de bonne classification ne représente pas un indicateur de performance adéquat.

Cette partie regroupe à la fois l'analyse de la courbe ROC et la mesure de l'aire sous cette courbe (AUC : area under the curve) comme mesures d'évaluation de performance.

1. Courbe ROC⁸⁸ :

La courbe ROC (Receiver Operating Characteristic) est un outil couramment utilisé pour évaluer la performance d'un modèle à détecter correctement les individus qui appartiennent à l'une des deux classes possibles.

Ce graphique représente :

⁸⁸ Voir annexe n°7.1.

- En abscisses : la sensibilité (ou taux de vrais positifs VP) : qui représente la proportion de prédictions correctes parmi l'ensemble des individus qui appartiennent à la classe cible. Elle mesure la capacité du modèle à détecter correctement les individus qui appartiennent à la classe cible.

$$\text{Sensibilité} = \frac{VP}{VP + FN}$$

Avec : FN désigne les faux négatifs.

- En ordonnées : la spécificité (ou taux de vrais négatifs VN) : qui représente la proportion de prédictions correctes parmi l'ensemble des individus qui n'appartiennent pas au groupe k. Elle mesure la capacité du modèle à éviter les faux positifs, c'est-à-dire les individus qui sont prédits comme appartenant au groupe k alors qu'ils n'en font pas partie.

$$\text{Spécificité} = \frac{VN}{VN + FP}$$

Avec : FP désigne les faux positifs.

La sensibilité ainsi que la spécificité de chaque modèle sont comme suit :

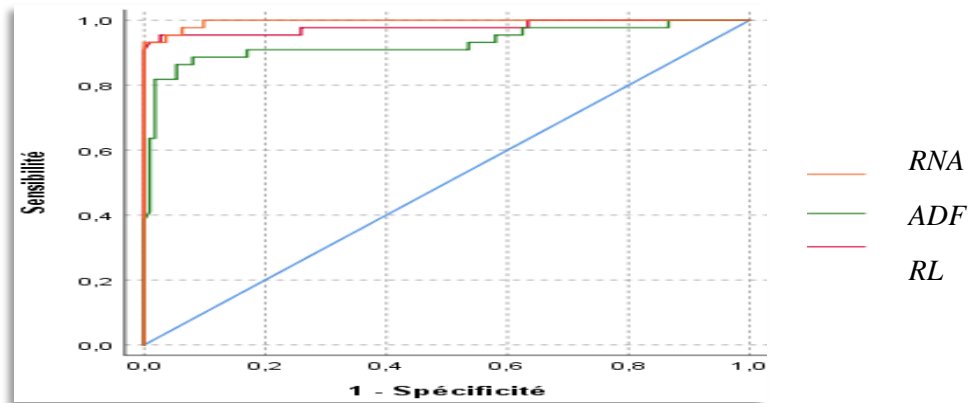
Tableau n°24 : Sensibilité-spécificité par modèle

	Sensibilité	Spécificité
ADF	98,2%	63,6%
RL	98,2%	88,6%
RNA	100%	93,9%

Source : élaboré par l'auteur.

Pour mieux comparer visuellement les modèles, nous présentons les trois courbes ROC sur le même graphique.

Graphique n°7 : Courbe ROC de trois modèles



Source : logiciel SPSS V26.

Nous pouvons remarquer que la courbe ROC relative au RNA (orange) est la plus haut par contre celle de l'ADF (verte) est la plus basse. Quant à la courbe de la RL (rouge), elle se situe entre les deux autres.

Ce graphique reflète les résultats de classement affichés par chaque modèle. En effet, la position du ROC dépend du taux de vrais positifs et du taux de vrais négatifs du modèle étudié.

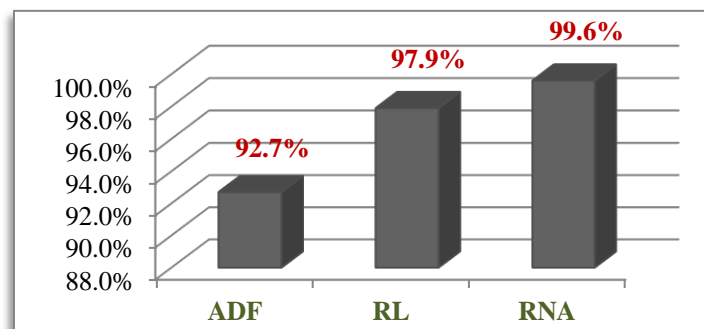
Les trois modèles sont loin de la ligne de référence (la diagonale en bleu) et ils présentent de bons résultats.

➔ Selon ce critère, le RNA est jugé le plus performant.

2. AUC (area under the curve) :

Comme son nom l'indique, l'AUC est calculé en déterminant l'aire sous la courbe ROC de chaque modèle. Le graphique suivant représente l'AUC en pourcentage par méthode de scoring :

Graphique n°8 : AUC par modèle de scoring



Source : élaboré par l'auteur à partir du logiciel SPSS V26.

Plus l'AUC est élevé, plus le modèle est spécifié à détecter les entreprises défaillantes ainsi que sensible à identifier les entreprises saines.

➔ Les trois modèles affichent une valeur d'AUC supérieure à 0.9, donc ils sont considérés comme excellents. Mais par rapport aux autres, le RNA est jugé plus performant.

III. Courbe PR (Précision-Rappel) :

La courbe PR est un outil couramment utilisé pour évaluer la performance d'un modèle à détecter correctement les individus qui appartiennent à l'un des deux groupes possibles.

Il s'agit d'un moyen utile pour évaluer la performance d'un classificateur dans les cas où il est important de minimiser les faux négatifs. Cet outil d'évaluation, nous permet de fixer le seuil « trad-off » entre précision et rappel de chaque modèle qu'à partir duquel le modèle est considéré précis à X%.

Ce graphique représente :

➤ En abscisses : le rappel est la proportion de prédictions correctes parmi l'ensemble des individus qui appartiennent au groupe k. Il représente la sensibilité du modèle.

$$\mathbf{Rappel} = \frac{\text{Nombre de prédictions correctes du groupe } k}{\text{Nombre total d'individus appartenant au groupe } k} = \mathbf{Sensibilité}$$

➤ En ordonnées : la précision est la proportion de prédictions correctes parmi l'ensemble des prédictions faites par le modèle. Elle reflète la capacité du modèle à prédire correctement la classe de l'individu.

$$\mathbf{Précision} = \frac{VP}{VP + FP}$$

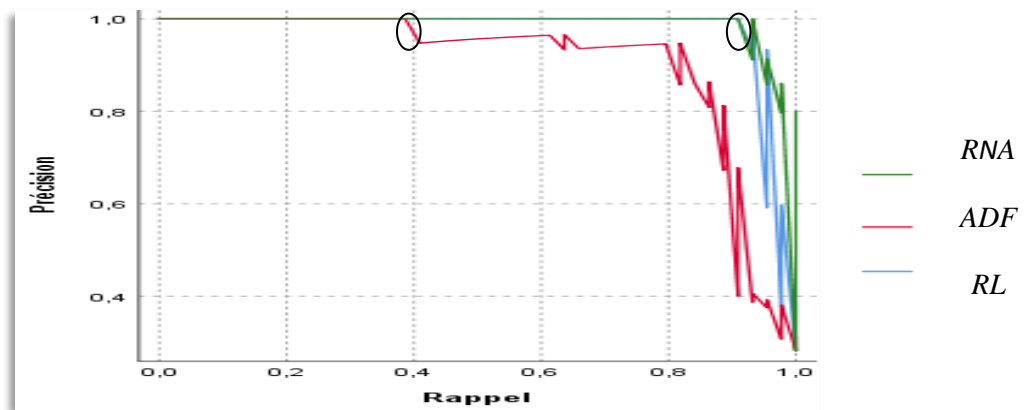
Tableau n°25 : Précision-Rappel par modèle

	Précision	Rappel
ADF	87,3%	98,2%
RL	95,6%	98,2%
RNA	97,7%	100%

Source : élaboré par l'auteur.

Pour mieux comparer visuellement les modèles, nous présentons les trois courbes PR sur le même graphique.

Graphique n°9 : Courbe PR de trois modèles



Source : logiciel SPSS V26.

Ce graphique reflète les résultats de classement affichés par chaque modèle. En effet, la RL (courbe en bleu) et le RNA (courbe en vert) présentent à peu près la même allure. En effet, pour afficher une parfaite précision (100%) les deux modèles présentent un taux de rappel de presque 90%.

Quant à la courbe de l'ADF (rouge), elle se situe au-dessous des deux autres. Pour un rappel d'environ 40%, l'ADF est précise à 100%, ce qui est jugé faible.

→ En se référant sur l'allure de la courbe PR, le RNA et la RL présentent la même qualité de classification. Mais les valeurs affichées au niveau du tableau n°25, le multilayer perceptron est jugé le plus performant.

IV. Généralisation du modèle :

Une des importantes conditions de validations du modèle de scoring au sein de l'établissement bancaire est la capacité de ce dernier d'enregistrer le même degré de précision indépendamment de l'échantillon et du temps.

En effet, la validation, dans le temps ou dans l'espace, est l'une des techniques utilisées pour évaluer la performance des modèles.

Dans le contexte de la prédiction de défaut futur des entreprises, la validation peut être utilisée pour évaluer la capacité du modèle à prédire les défaillances de nouvelles entreprises qui n'ont pas été incluses dans l'ensemble d'apprentissage du modèle.

Cela nous permet de vérifier que le modèle est à la fois robuste et généralisable, et qu'il ne se limite pas à mémoriser les données utilisées pour sa construction sans apprendre à prédire de manière générale.

Pour chaque sous-échantillon réservé à la validation des modèles, nous allons présenter la matrice de confusion pour étudier les taux de classement global et par groupe.

Remarques :

- Grâce à l’option de la validation croisée⁸⁹ offerte par le logiciel SPSS V26 parmi les rubriques du menu « analyse discriminante factorielle », nous ne sommes pas appelés à décortiquer l’échantillon de [2019-2020] pour la validation « out of sample » de la fonction discriminante.
- En outre, lors de la modélisation des RNA, le logiciel propose une partition de l’échantillon de [2019-2020] en trois sous-groupes : échantillon d’apprentissage et deux échantillons pour la validation en deux phases pour s’assurer de la stabilité du réseau construit.

1. Validation « out of sample »⁹⁰ :

Les tests hors échantillon consistent à utiliser de nouvelles données qui ne font pas partie de la base de données utilisée pour la construction du modèle.

Cette démarche est souvent considérée comme étant la meilleure méthode pour évaluer la capacité du modèle à anticiper les résultats sur des nouvelles bases de données non utilisées, Bardos et Zhu (1997).

En effet, la validation hors échantillon présente l’avantage, par rapport à la validation dans l’échantillon, de prémunir contre le sur-ajustement.

Tableau n°26 : Résultats de classement : validation « out of sample »

		Erreur de type I	Erreur de type II	Taux de bon classement
ADF		38,6%	1,8%	87,8%
RL		11.1%	3.3%	94,9%
RNA	Test 1	0%	0%	100%
	Test 2	33.3%	0%	90%

Source : élaboré par l’auteur.

⁸⁹La phase de validation croisée consiste à découper les données en plusieurs sous-ensembles, appelés « folds », et à entraîner le modèle sur chacun de ces sous-ensembles. Le modèle est alors évalué sur le sous-ensemble qui n'a pas été utilisé pour l'entraînement. Cette étape est répétée plusieurs fois, en utilisant chaque sous-ensemble comme échantillon de test à tour de rôle. La performance du modèle est alors moyennée sur l'ensemble des folds.

⁹⁰ Voir annexe n°8.

En observant la 3^{ème} colonne du tableau ci-dessus, nous remarquons que les trois modèles affichent pour l'échantillon du test un taux de bon classement acceptable et largement supérieur à 50% (le taux du modèle de classification aléatoire). Cela signifie que les modèles testés sont capables de prédire la classe d'appartenance d'un individu avec précision.

Le réseau de neurones affiche le taux le plus important de 100% pour le 1^{er} test de validation. Cependant, il ne garde pas ce degré de précision lors du 2^{ème} test et il enregistre un taux de 90% supérieur au taux de l'ADF mais inférieur à celui de la régression logistique qui montre des résultats constants et stables avec un taux de 94,9% pour la validation par rapport à 95,5% pour les données de construction.

De plus, le RNA multiplie le taux d'erreur type I fois 5 pour l'échantillon de test 1 (33,3% contre 6,1% pour l'échantillon d'apprentissage). Sous cette dégradation de qualité de modèle multilayer perceptron, nous mettons en évidence la limite de sur-apprentissage de cette famille de modèles.

→ Par conséquent, nous considérons, selon les résultats du 1^{er} test de validation, que les résultats de la RL ainsi que de la technique de l'ADF sont stables dans l'espace.

2. Validation « out of time »⁹¹ :

Pour s'assurer de la stabilité dans le temps des modèles élaborés, nous exploitons l'échantillon de [2020-2021]. Nous utilisons les modèles retenus pour estimer la probabilité de défaut des entreprises afin de réaffecter ces dernières en entreprises saines et autres défaillantes.

Tableau n°27 : Résultats de classement : validation « out of time »

	Erreur de type I	Erreur de type II	Taux de bon classement
ADF	0%	36,4%	74,3%
RL	0%	4%	97,1%
RNA	39%	4%	85,7%

Source : élaboré par l'auteur.

En termes du taux de bon classement, du pourcentage d'erreur de type I ainsi que de type II, la régression logistique enregistre les meilleurs résultats par rapport aux autres méthodes.

⁹¹ Voir annexe n°9.

➔ Selon les résultats de ce test, nous considérons que la RL est capable de fournir des résultats cohérents et précis dans le temps.

Au niveau des parties précédentes nous avons essayé de comparer les méthodes selon plusieurs critères séparément. Nous avons trouvé qu'en termes de performance le réseau de neurones de type perceptron multicouches est le modèle à retenir. Cependant, en termes de stabilité des résultats à la fois dans le temps et dans l'espace, la régression logistique représente le modèle à choisir.

Pour arriver à un choix définitif, nous exploitons une mesure générale d'évaluation de performance des modèles.

3. F1-Score :

Le F1-score est une mesure générale d'évaluation de performance des modèles qui permet de prendre en compte à la fois la précision et le rappel.

De ce fait, nous utilisons cette mesure afin d'étudier la stabilité des résultats de chaque modèle en changeant l'échantillon d'entreprises, dont la formule de calcul du F1-score est comme suit:

$$F1 - score = 2 * \frac{Précision * Rappel}{Précision + Rappel}$$

La démarche adoptée est la suivante, nous calculons pour chaque échantillon (apprentissage et tests) le F1-score du modèle, puis le F1-score moyen pour déterminer la performance moyenne des modèles dans le temps et dans l'espace. Finalement, nous choisissons le modèle le plus optimal.

Les résultats de calculs sont présentés au niveau du tableau suivant :

Tableau n°28 : F1-Score des modèles par échantillon

	Apprentissage	Testing « out of sample »		Testing « out of time »	F1-score moyen	F1-score Ecart-Type	F1-score CV
ADF	0,92	0,92		0,78	0,87	0,08	9,3
RL	0,97	0,95		0,98	0,97	0,02	1,6
RNA	0,99	1	0,93	0,96	0,96	0,05	5

Source : élaboré par l'auteur.

Sur la base des résultats affichés sur les 3 échantillons, la régression logistique enregistre à la fois :

- ✓ Le F1-score moyen le plus élevé,
 - ✓ L'écart-type du F1-score le plus faible,
- ➔ Par conséquent, elle affiche le coefficient de variation (CV) le plus faible par rapport aux autres méthodes utilisées.

Cela nous permet de vérifier que le modèle est généralisable et robuste et qu'il ne se contente pas de mémoriser les données d'apprentissage sans apprendre à prédire de manière générale.

Pour conclure, la régression logistique représente une méthode de crédit scoring dotée d'une bonne capacité de prédiction du passage en défaut des entreprises.

Donc, elle peut être utilisée de manière fiable pour la prédiction de défaut des entreprises au sein de la banque.

V. Modèle retenu : Régression logistique :

Après avoir choisi le modèle de la régression logistique, nous présentons au niveau du tableau suivant l'interprétation financière des différentes variables.

Tableau n°29 : Interprétation des variables de la régression logistique

Variable	Formule	OR	Signe	Définition	Interprétation
<p>R1 : Trésorerie nette</p>	$FDR - BFR$	0,879	-	<p>C'est un ratio qui mesure la capacité de l'entreprise à faire face à ses dépenses courantes d'exploitation.</p> <p>→ Equilibre financier.</p>	<p>Une trésorerie nette positive signifie que l'entreprise dispose de disponibilité mobilisable à court terme.</p> <p>Donc elle reflète une situation financière saine de l'entreprise.</p> <p>→ Une relation négative avec le passage en défaut de l'entreprise.</p>
<p>R10 : Ratio de levier</p>	$\frac{Total\ Dettes}{Capitaux\ propres}$	1,083	+	<p>Ce ratio mesure le degré de financement de l'entreprise par des fonds empruntés par rapport à ses capitaux propres.</p> <p>→ Structure financière.</p>	<p>Un ratio de levier élevé signifie que l'entreprise a un endettement élevé et un faible coussin de sécurité pour faire face aux imprévus.</p> <p>Par conséquent, elle est plus susceptible de ne pas être en mesure de rembourser sa dette en cas de difficultés financières.</p> <p>→ Une relation positive avec le passage en défaut de l'entreprise.</p>

<p>R11 : Ratio d'endettement global</p>	$\frac{\text{Total Dette}}{\text{Total Bilan}}$	1,298	+	<p>C'est un ratio qui compare la dette totale de l'entreprise à son total des actifs, qui incluent tous les biens et ressources dont dispose l'entreprise pour mener son activité.</p> <p>Il reflète la solvabilité ainsi que le niveau de dépendance financière de l'entreprise.</p> <p>➔ Solvabilité : endettement.</p>	<p>Un ratio d'endettement global élevé indique que l'entreprise a un niveau de dette élevé par rapport à ses actifs.</p> <p>Donc, elle est considérée peu solvable, par conséquent, elle est incapable de rembourser un éventuel prêt accordé.</p> <p>➔ Une relation positive avec le passage en défaut de l'entreprise.</p>
<p>R15 : Ratio de liquidité générale</p>	$\frac{\text{Actifs Circulants}}{\text{Dettes CT}}$	0,733	-	<p>C'est un ratio qui mesure la capacité de l'entreprise à faire face à ses obligations de CT.</p> <p>➔ Solvabilité : liquidité.</p>	<p>Un ratio de liquidité générale élevé indique que l'entreprise a suffisamment d'actifs courants pour couvrir ses obligations à CT.</p> <p>➔ Une relation négative avec le passage en défaut de l'entreprise.</p>
<p>R21 : Ratio de rentabilité</p>	$\frac{\text{CAF}}{\text{CAHT}}$	0,787	-	<p>Ce ratio permet de déterminer la part du chiffre d'affaires destinée à la création de ressources internes afin d'assurer le financement de l'entreprise.</p> <p>➔ Activité.</p>	<p>Plus le ratio de rentabilité est important plus l'entreprise est en mesure de financer ses investissements à partir des ressources dégagées de sa propre activité.</p> <p>➔ Une relation négative avec le passage en défaut de l'entreprise.</p>

<p>Q1 : Taux des provisions sectorielles</p>	$\frac{\text{Total provisions par secteur}}{\text{Total provisions BNA}}$	1,245	+	<p>Cette mesure indique le niveau de risque de défaillance dans un secteur d'activité. Donc, elle reflète l'appartenance des entreprises à un secteur d'activité particulier. → Qualité.</p>	<p>Si la relation étudiée appartient à un secteur d'activité sinistré dot le taux de provisions du secteur est important. Elle peut avoir du mal à générer des revenus et à maintenir sa rentabilité, ce qui peut entraîner une augmentation de sa probabilité de défaut. → Une relation positive avec le passage en défaut de l'entreprise.</p>
<p>C1 : Taux d'impayés par engagements</p>	$\frac{\text{Encours impayés}}{\text{Total engagements BNA}}$	1,994	+	<p>Ce taux représente un indicateur d'évaluation de la solvabilité de l'entreprise, il reflète sa capacité de rembourser ses dettes. → Comportement.</p>	<p>Il y a deux cas de figures pour interpréter la situation d'une entreprise ayant un taux important d'impayés par engagements : Soit l'incapacité de régler ses engagements. Soit l'volonté de régler ses engagements. → Une relation positive avec le passage en défaut de l'entreprise.</p>

Source : élaboré par l'auteur.

Conclusion :

L'objectif du présent chapitre était l'élaboration d'un modèle de crédit scoring pour la discrimination entre les deux groupes d'entreprises saines et défaillantes de notre échantillon, et ce en employant les variables sélectionnées au niveau du chapitre précédent.

D'abord, nous avons exploité les différentes techniques de modélisation du risque de crédit, telles que la technique d'analyse discriminante factorielle, la régression logistique ainsi que la technique fondée sur l'intelligence artificielle à savoir le réseau de neurones multilayer perceptron.

Par ailleurs, en se basant sur les résultats empiriques relatifs à chaque modèle estimé, nous avons constaté que chacun d'entre eux a montré une performance acceptable en termes de précision et de qualité d'ajustement.

Cependant, en confrontant les trois modèles afin de préciser celui ayant le pouvoir explicatif et discriminant le plus important, nous avons remarqué que la régression logistique a fourni la meilleure capacité de généralisation. Donc, elle peut prédire de manière fiable le passage en défaut des nouvelles relations par rapport à l'analyse discriminante et le réseau de neurones.

Autre que la performance et la stabilité, notre modèle de prédiction du risque de crédit définitivement retenu est facile à comprendre et à mettre en œuvre. Dans la mesure où elle utilise une fonction de probabilité simple et facile à interpréter.

Conclusion Générale

Par l'exercice quotidien de son activité, la banque est exposée à différents types de risques qui peuvent entraîner des pertes financières considérables et donc affecter sa rentabilité et sa viabilité à long terme.

L'incertitude de l'environnement bancaire, associée à la forte concurrence mettent la gestion des risques bancaires, dont le risque de crédit fait partie, parmi les premières préoccupations des établissements de crédit.

A cet effet, afin de tirer profit de leurs activités tout en respectant un niveau acceptable de risque, les banques doivent mettre en place un système d'évaluation du risque de crédit efficace et adapté à leur profil de risque.

Dans ce travail de recherche, nous nous sommes penchés sur l'élaboration un modèle de « crédit scoring » fiable et assez robuste offrant une meilleure visibilité sur la situation financière de la relation, en général, et sur le degré de solvabilité de cette dernière, en particulier.

Pour effectuer ce travail, nous avons donc cité dans la première section du premier chapitre les différents types de risques liés à l'activité bancaire. Dans la deuxième section du même chapitre, nous nous sommes focalisés sur le risque de crédit. Finalement, nous avons présenté, au niveau de la troisième et la quatrième section, l'évolution de la réglementation prudentielle internationale ainsi que nationale en matière de risque de crédit.

Nous avons consacré le deuxième chapitre théorique de notre mémoire, à l'exposition de différentes approches d'évaluation du risque de crédit, en commençant par l'approche traditionnelle fondée sur l'analyse financière. Par la suite, nous avons présenté le principe du rating, les avantages et les limites de la méthode. Finalement, nous avons cité les diverses techniques de crédit scoring de l'approche statistique ainsi que celle basée sur l'intelligence artificielle.

Quant à la partie empirique, nous pouvons représenter notre démarche méthodique par trois grandes étapes fondamentales :

La première était réservée aux prétraitements et au nettoyage de la base de données des relations (saines et défaillantes) de la BNA nécessaires à la préparation des variables pour notre étude.

Après, nous procédons à la sélection des variables significatives les moins corrélées et les plus discriminantes entre les entreprises saines et défaillantes de notre échantillon en utilisant le test t de Student, le test de Mann-Whitney et le test du khi-2.

Parmi un ensemble de 26 variables initialement retenues, ces tests nous ont permis d'identifier seulement 13 pour passer à la deuxième étape relative à la construction du modèle. Nous avons modélisé à la fois la technique de l'analyse discriminante, de la régression logistique et des réseaux de neurones artificiels.

Chaque modèle estimé a montré une capacité importante de prédire le défaut des entreprises de l'échantillon d'apprentissage et une bonne qualité d'ajustement. Mais nous avons constaté que l'architecture du réseau de neurones représente le modèle le plus précis par rapport aux autres.

Pour arriver à une décision finale portant sur l'estimation à retenir, en se basant sur les deux caractéristiques d'un bon modèle de scoring tels que, le bon classement et la stabilité des résultats, nous avons conclu, au niveau de la troisième étape, que notre choix final est la régression logistique.

Cette régression regroupe sept variables discriminantes dont cinq ratios reflétant la situation financière de la contrepartie, une variable comportementale décrivant l'attitude historique de la relation et une autre qualitative relative au secteur d'activité de cette dernière.

Dans un souci de rigueur scientifique, comme toute recherche la nôtre comporte certaines limites qui n'affectent pas nos résultats mais qui peuvent les relativiser.

Une première limite porte sur la taille de l'échantillon et une deuxième concerne l'ensemble des variables utilisées. Il aurait été intéressant d'utiliser, au niveau de cette étude, un plus grand nombre de variables comportementales et qualitatives afin d'intégrer plus d'informations sur la relation, et par conséquent d'avoir une image plus claire de la capacité et de la volonté de cette dernière à honorer ses engagements contractuels.

Enfin, ce mémoire peut servir à d'autres études beaucoup plus approfondies. Pour optimiser le potentiel des données, nous recommandons la technique du « Stacking » ou encore L'empilement, il s'agit d'une technique d'assemblage de modèles, permettant de combiner les informations de trois modèles employés.

Pour finir ce mémoire, je vous souhaite bonne réception en vous remerciant vivement de l'attention que vous m'avez bien voulu me réserver et de votre précieux temps que vous m'avez aimablement accordé.

Bibliographie

- ABDELOUAHED, B., & EL GAZI, M. (2015). La transposition de l'approche IRB-A de b le II pour le secteur bancaire marocain
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4), 589-609.
- Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of banking & finance*, 1(1), 29-54.
- Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of banking & finance*, 18(3), 505-529.
- Altman, EI, & Saunders, A. (1997). Mesure du risque de cr dit :  volutions au cours des 20 derni res ann es. *Journal of banking & finance* , 21 (11-12), 1721-1742.
- Anandarajan, M., Lee, P., & Anandarajan, A. (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *Intelligent Systems in Accounting, Finance & Management*, 10(2), 69-81.
- Arbulu, P., & Lamarque,  . (2011). Chapitre introductif.
- Bardos, M. (2008). Scoring sur donn es d'entreprises: instrument de diagnostic individuel et outil d'analyse de portefeuille d'une client le. *Revue Modula*, 38, 159-177.
- Bardos, M., & Zhu, W. H. (1997). Comparaison de l'analyse discriminante lin aire et des r seaux de neurones. Application   la d tection de d faillance d'entreprises. *Revue de statistique appliqu e*, 45(4), 65-92.
- Beaver, WH (1966). Les ratios financiers comme pr dicteurs d' chec. *Revue de recherche comptable* , 71-111
- Bel s, J. & Cipovov , E. (2011), « Internal model of commercial bank as an instrument for measuring credit risk of the borrower in relation to financial performance (credit scoring and bankruptcy models) », *Journal of Competitiveness*, vol. 4. p. 104-120
- Bessis J ,(2015), risk management in banking, Fourth edition. Great Britain: Library of Congress Cataloging-in-publication Data.

- Bohn, J. R., & Stein, R. M. (2009). *Active credit portfolio management in practice*. John Wiley & Sons.
- BOUAZZARA, A., Riad, B. A. H. A., & BEKTACHE, F. (2020). Évaluation du risque de défaillance de solvabilité des PME: une application du modèle de la régression logistique. *Dirassat Journal Economic Issue*, 11(2), 491-505.
- BOUHLAL, N. (2021). Internal Rating: A Tool for Managing the Bank Credit Risk of SMEs. *La Notation Interne: Un Outil De Gestion Du Risque De Crédit Bancaire Des PME*. *Revue d'Economie et de Statistique Appliquée*, 18(2).
- Boujelbène, Y., & Khemakhem, S. (2013). Préviation du risque de crédit: Une étude comparative entre l'Analyse Discriminante et l'Approche Neuronale.
- BRUNEL, V. (2009). Gestion des risques et risque de crédit. *Article de recherche, Ecole Nationale des Ponts et Chaussées*.
- BRUNEL, V. (2009). Gestion des risques et risque de crédit. *Article de recherche, Ecole Nationale des Ponts et Chaussées*.
- Bruno, O. & Girod, A. (2008). Bâle II, réallocation des portefeuilles de crédits et incitation à la prise de risque: Une application au cas des pays émergents d'Asie du Sud-Est. *Revue économique*, 59, 1193-1213. <https://doi.org/10.3917/reco.596.1193>
- Carruthers, B. & Cohen, B. (2010). Noter le crédit : classification et cognition aux États-Unis. *Genèses*, 79, 48-73. <https://doi.org/10.3917/gen.079.0048>.
- Collard, F. (2012). Les agences de notation. *Courrier hebdomadaire du CRISP*, 2156-2157, 5-60. <https://doi.org/10.3917/cris.2156.0005>
- Collongues Y. (1977), "Ratios Financiers et Préviation des Faillites des Petites et Moyennes Entreprises", in *Banque*, vol. n°365 (septembre).
- Daoui, D., & Mouatassim, M. (2021). La réglementation prudentielle de Bâle au Maroc: état des lieux, et apports pour le secteur bancaire. *International Journal of Accounting, Finance, Auditing, Management and Economics*, 2(5), 707-719.
- Darbellay, A. & Partnoy, F. (2012). Agences de notation et conflits d'intérêts. *Revue d'économie financière*, 105, 309-318. <https://doi.org/10.3917/ecofi.105.0309>
- Darbellay, A. (2011). Régulation des notations : l'oligopole des agences de notation de crédit d'un point de vue réglementaire . *Schulthess*.

- De Coussergues, S., Bourdeaux, G., Péran, T. (2017). *Gestion de la banque*. Dunod. <https://doi.org/10.3917/dunod.decou.2017.01>
- De La Bruslerie H. (2014), *Analyse financière : Information financière, évaluation, diagnostic*, Paris, Dunod 5^{ème} éd.
- De La Bruslerie, H. (2014). Chapitre 12. L'analyse du risque crédit par le prêteur. Dans : , H. de La Bruslerie, *Analyse financière: Information financière, évaluation, diagnostic* (pp. 453-506). Paris: Dunod.
- De Servigny, A. & Zelenko, I. (2010). Chapitre 3. Défaut et migration de crédit. Dans : *Le risque de crédit: Face à la crise* (pp. 79-127). Paris: Dunod.
- Degos, J. G., Hmiden, O. B., & Henchiri, J. E. (2010). Credit rating agencies: accounting adjustments and econometrical analysis. *International Journal of Economics and Accounting*, 1(1-2), 88-106.
- Degos, J. G., Hmiden, O. B., & Henchiri, J. E. (2010). Credit rating agencies: accounting adjustments and econometrical analysis. *International Journal of Economics and Accounting*, 1(1-2), 88-106.
- Degos, J., Ben Hmiden, O. & Henchiri, J. (2012). Les agences de notation financières: Naissance et évolution d'un oligopole controversé. *Revue française de gestion*, 227, 45-65. <https://doi.org/10.3166/RFG.227.45-65>
- Degos, JG, Hmiden, OB et Henchiri, JE (2010). Agences de notation : ajustements comptables et analyse économétrique. *Revue internationale d'économie et de comptabilité* , 1 (1-2), 88-106.
- Delaite, M. (2012). Les fonds propres bancaires au cœur de la crise financière. *Mondes en développement*, 160, 117-132. <https://doi.org/10.3917/med.160.0117>
- Delaite, M. (2012). Les fonds propres bancaires au cœur de la crise financière. *Mondes en développement*, 160, 117-132. <https://doi.org/10.3917/med.160.0117>.
- Dietsch, M., & Petey, J. (2003). *Mesure et gestion du risque de crédit dans les institutions financières* (No. 2013/14375). ULB--Universite Libre de Bruxelles.
- Éric Paget-Blanc (2003), *Rating et probabilité de défaut des entreprises européennes : détermination par un modèle de régression logistique ordonné*. *Banques & marchés*. N° 65.

- Fekir, H. (2009). Presentation Du Nouvel Accord De Bale Sur Les Fonds Propres. arXivpreprint arXiv:0905.2546.
- Ferri, G., & Liu, L. G. (2005). Assessing the effort of rating agencies in emerging economies: Some empirical evidence. *European Journal of Finance*, 11(3), 283-295.
- Gabillon, J. (2007). Spreads de crédit et taux d'intérêt. *Finance*, 28, 121-160. <https://doi.org/10.3917/fina.282.0121>
- Gadhoun, Y., Gueyie, J. & Karim Siala, M. (2007). La décision de crédit : Procédure et comparaison de la performance de quatre modèles de prévision d'insolvabilité. *La Revue des Sciences de Gestion*, 224-225, 177-183. <https://doi.org/10.3917/rsg.224.0177>
- Galati, G. (2002). CLS Bank et le risque de règlement dans les opérations de change. Rapport trimestriel BRI.
- Gaultier-Gaillard, S., & Pratlong, F. (2011). Le risque de réputation: le cas du secteur bancaire. *Management Avenir*, (8), 272-288.
- Ghosh, A. (2012). *Managing risks in commercial and retail banking*. John Wiley & Sons.
- Giudici, P. (2003) *Applied Data Mining: Statistical Methods for Business and Industry*. Wiley, Hoboken.
- Grosjean, P. & Engels, G. (2022), *Science des données biologiques III*. <https://wp.sciviews.org/sdd-umons3>
- HABIBA MOUSSOUNI : « LES ACCORDS DE BÂLE ET REGLES PRUDENTIELLES DES BANQUES: DEFIS ET CONTRAINTES POUR LE SYSTEME BANCAIRE ALGERIEN » (Thèse de doctorat : 2013-2014).
- Horrigan, JO (1968). Une brève histoire de l'analyse des ratios financiers. *La revue comptable*, 43 (2), 284-294.
- Ishtiaq, M. (2015). Risk management in banks: determination of practices and relationship with performance.
- Kefi, M. & Maraghni, H. (2011). Respect des normes prudentielles et solvabilité des banques commerciales : étude empirique des banques tunisiennes. *Management & Avenir*, 48, 382-403. <https://doi.org/10.3917/mav.048.0382>

- Kefi, M. K., & Maraghni, H. (2011). Respect des normes prudentielles et solvabilité des banques commerciales: étude empirique des banques tunisiennes. *Management Avenir*, (8), 382-403.
- Kharoubi, C. & Thomas, P. (2016), *Analyse du risque de crédit : banque & marchés*. Paris: RB 2^{ème} éd.
- Koch, T. W., & MacDonald, S. S. (2014). *Bank management*. Cengage Learning.
- Levy, A., Baha, R. & Bouzar, S. (2021). Évaluation du risque de défaut de solvabilité des PME par une double approche: Une application du modèle Logit et du modèle *CreditRisk+*. *Gestion 2000*, 38, 51-70. <https://doi.org/10.3917/g2000.381.0051>
- Lopez, JA (2004). La relation empirique entre la corrélation moyenne des actifs, la probabilité de défaut de l'entreprise et la taille des actifs. *Journal de l'intermédiation financière*, 13 (2), 265-283.
- Lotfi, S., & Mesk, H. (2020). Prédiction du risque de crédit: étude comparative des techniques de Scoring. *International Journal of Accounting, Finance, Auditing, Management and Economics*, 1(2), 511-527.
- Makany, J. & GabsouboYienezoune, C. (2013). L'évaluation du risque de crédit des entreprises: cas de la banque congolaise de l'habitat. *Revue Congolaise de Gestion*, 17, 87-130. <https://doi.org/10.3917/rcg.017.0087>
- Makany, J. N. J., & Yienezoune, C. G. (2013). L'évaluation du risque de crédit des entreprises: cas de la banque congolaise de l'habitat. *Revue Congolaise de Gestion*, 17(1), 87-130.
- Matoussi, H., Mouelhi, R., & Salah, S. (1999, May). La prédiction de faillite des entreprises tunisiennes par la régression logistique. In 20ÈME CONGRES DE L' AFC (pp. CD-Rom).
- Mester, LJ (1997). Quel est l'intérêt de la cote de crédit. *Business review*, 3 (septembre/octobre), 3-16.
- Modrik, K. (2016). L'évaluation du risque de crédit des PME françaises internationalisées (Doctoral dissertation, Université d'Angers).

- NDAYIZEYE, L., & NIYUHIRE, P. (2018). IMPORTANCE DU RISQUE SYSTEMIQUE AU BURUNDI ET SA PREVENTION. *FINANCES, AGRICULTURE ET ASSAINISSEMENT*, 11.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109-131.
- Refait-Alexandre, C. (2004). La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux. *Économie & prévision*, 129-147. <https://doi.org/10.3917/ecop.162.0129>
- REGRAGUI, Y., & AL MERIOUH, Y. (2015). Gestion du risque de liquidité: Problématique du comportement de la clientèle bancaire marocaine [Risk management of liquidity: Problematic of the behavior of the Moroccan banking customers].
- Rochet, J. C. (2008). *Le futur de la réglementation bancaire*.
- Roncalli, T. (2009). *La gestion des risques financiers*. Paris: Economica.
- Rudd, M. P. H., GStat, J. M., & Priestley, J. L. (2017). A comparison of decision tree with logistic regression model for prediction of worst non-financial payment status in commercial credit.
- SADI, K. (2010). Un état des lieux des différentes techniques d'analyse statistiques pour la construction des modèles de prévision de défaillance. *Revue d'économie et de statistique appliquée*, Volume 7, Numéro 2, Pages 83-95
- Sardi (2004) « Bâle II », édition afdg, paris.
- Seffar, M., Olivero, B., & Luu, P. (2019). *Dynamique de la Notation et Performance Financière de l'Entreprise*.
- Vernimmen P. (1998), *Finance d'entreprise*, 3ème édition par Quiry P. & Ceddaha F., Dalloz.
- VISNOVSKY, F. (2017) Secrétaire général adjoint. : « Bâle 1, 2, 3 ... de quoi s'agit-il ? Banque de France.
- Wakeman M. (1998). "The real function of bond rating agencies", *The revolution in corporate finance*, Stern & Chew (Eds), Malden MA Blackwell Business.
- White, L. A (2009). Les agences de notations de crédit et la débâcle des subprimes, *Critical Review*, 21:2-3, 389-399, <https://doi.org/10.1080/08913810902974964>

Webliographie

Site officiel de la BNA : <http://www.bna.tn/>

Site officiel de la BCT : <https://www.bct.gov.tn/bct/siteprod/index.jsp>

Site officiel du CBCB : <https://www.bis.org/>

Site officiel de Fitch Rating : <https://www.fitchratings.com/>

Site web du Google Scholar : https://scholar.google.com/schhp?hl=fr&as_sdt=0,5

Site officiel de Moody's : <https://www.moodys.com/>

Site web de Science Direct : <https://www.sciencedirect.com/>

Site officiel de S&P : <https://www.spglobal.com/ratings/en/>

Annexes

Annexe 1 : Statistiques descriptives

Statistiques descriptives des variables quantitatives

	N	Minimum	Maximum	Moyenne	Ecart type
R1	156	-250169234	91794132	-7823428,28	38385439,237
R7	156	-52,00%	86,00%	32,1090%	25,32921%
R11	156	-15,08	34,30	3,0218	5,78049
R12	156	-2,50	5,48	,3212	,92415
R9	156	-21,62	73,43	4,0304	9,89658
R15	156	,25	5,25	1,4583	,86230
R16	156	,00	29,60	1,3272	3,32374
R5	156	-1,38	1,01	,1060	,28409
R6	156	-,17	,26	,0377	,06581
R13	156	-3,31	5,16	,0356	,90087
R10	156	-402,40	62,31	-2,5817	40,84590
R8	156	-11,46	1393,00	15,1654	116,24080
R18	156	-176,00%	1698,00%	128,3910%	345,50480%
R19	156	-62,00%	78,00%	4,4808%	22,94438%
R21	156	-0,06%	0,14%	0,0080%	0,02178%
R14	156	-40,42	34,54	,6117	66,1571
R20	156	-20,60%	38,00%	3,8521%	7,77101%
R17	156	-,33	,75	,1835	,22582
R2	156	-4,62	5,00	,4235	1,12574
R3	156	-736731,773	1059643,503	63961,129	162836,792
R4	156	-609169,909	1162523,990	108345,260	179896,175
Q1	156	0,00%	20,83%	0,6872%	2,09476%
Q3	156	0,602	6,602	4,100	0,957
Q4	156	1	30	14,97	8,504
C1	156	0,00%	50,00%	3,7756%	9,73426%
C2	156	0	6518	238,21	921,126

Statistiques descriptives des variables qualitatives

Variable Statistique	Y (Passage en Défaut)
Nb. d'observations	156
Nb. de valeurs manquantes	0
Somme des poids	156
Nb. de modalités	2
Mode	0

Mode (effectif)	112	
Modalités	0	1
Effectif par modalité	112	44
Fréquence par modalité (%)	71,795	28,205
Borne inf. des freq. (95%)	64,733	21,144
Borne sup. des freq. (95%)	78,856	35,267
Proportion par modalité	0,718	0,282
Borne inf. des proportions (95%)	0,647	0,211
Borne sup. des proportions (95%)	0,789	0,353

Variable Statistique	Q2	
Nb. d'observations	156	
Nb. de valeurs manquantes	0	
Somme des poids	156	
Nb. de modalités	2	
Mode	1	
Mode (effectif)	84	
Modalités	0	1
Effectif par modalité	72,000	84,000
Fréquence par modalité (%)	46,154	53,846
Borne inf. des freq. (95%)	38,331	46,023
Borne sup. des freq. (95%)	53,977	61,669
Proportion par modalité	0,462	0,538
Borne inf. des proportions (95%)	0,383	0,460
Borne sup. des proportions (95%)	0,540	0,617

Annexe 2 : Présélection des variables discriminantes

2.1. Test de Student

		Test de Levene		Test t pour égalité des moyennes		
		F	Sig.	t	ddl	Sig.
R1	Hypothèse de variances égales	51,633	0	4,602	154	0
	Hypothèse de variances inégales			3,063	45,219	0,004
R7	Hypothèse de variances égales	3,176	0,077	2,685	154	0,008
	Hypothèse de variances inégales			2,296	59,854	0,025

R11	Hypothèse de variances égales	12,035	0,001	-2,721	154	0,007
	Hypothèse de variances inégales			-2,176	54,633	0,034
R12	Hypothèse de variances égales	14,16	0	-2,728	154	0,007
	Hypothèse de variances inégales			-2,158	53,872	0,035
R9	Hypothèse de variances égales	8,149	0,005	-0,171	154	0,865
	Hypothèse de variances inégales			-0,132	52,535	0,895
R15	Hypothèse de variances égales	290,828	0	-9,228	154	0
	Hypothèse de variances inégales			-5,755	43	0
R16	Hypothèse de variances égales	2,053	0,154	-0,538	154	0,591
	Hypothèse de variances inégales			-0,504	69,327	0,616
R5	Hypothèse de variances égales	0,986	0,322	0,821	154	0,413
	Hypothèse de variances inégales			0,783	71,665	0,436
R6	Hypothèse de variances égales	1,019	0,314	1,379	154	0,17
	Hypothèse de variances inégales			1,319	72,167	0,191
R13	Hypothèse de variances égales	0,13	0,719	-0,171	154	0,865
	Hypothèse de variances inégales			1,988	107,424	0,049
R10	Hypothèse de variances égales	0,127	0,722	-0,713	154	0,477
	Hypothèse de variances inégales			-0,82	108,687	0,414
R8	Hypothèse de variances égales	2,183	0,142	0,909	154	0,365
	Hypothèse de variances inégales			1,451	111,562	0,15
R18	Hypothèse de variances égales	0,193	0,661	0,548	154	0,584
	Hypothèse de variances inégales			0,544	77,561	0,588
R19	Hypothèse de variances égales	1,893	0,171	1,736	154	0,084
	Hypothèse de variances inégales			2,034	62,799	0,046
R21	Hypothèse de variances égales	0,083	0,773	1,611	154	0,109
	Hypothèse de variances inégales			1,314	56,059	0,194
R14	Hypothèse de variances égales	0,22	0,639	1,196	154	0,234
	Hypothèse de variances inégales			1,032	60,697	0,306
R20	Hypothèse de variances égales	3,908	0,05	3,321	154	0,001
	Hypothèse de variances inégales			1,567	63,987	0,122
R17	Hypothèse de variances égales	0,038	0,846	1,002	154	0,318
	Hypothèse de variances inégales			0,978	74,921	0,331
R2	Hypothèse de variances égales	0,057	0,812	1,754	154	0,081
	Hypothèse de variances inégales			3,034	66,459	0,003
R3	Hypothèse de variances égales	0,012	0,914	0,967	154	0,335
	Hypothèse de variances inégales			1,017	87,802	0,312
R4	Hypothèse de variances égales	0,948	0,332	-0,918	154	0,36
	Hypothèse de variances inégales			-1,063	110,311	0,29

Q1	Hypothèse de variances égales	1,805	0,181	-0,571	154	0,569
	Hypothèse de variances inégales			-0,436	51,604	0,665
Q3	Hypothèse de variances égales	2,183	0,142	0,909	154	0,365
	Hypothèse de variances inégales			1,451	111,562	0,15
Q4	Hypothèse de variances égales	0,886	0,348	-0,405	154	0,686
	Hypothèse de variances inégales			-0,417	83,552	0,678
C1	Hypothèse de variances égales	221,905	0	-9,827	154	0
	Hypothèse de variances inégales			-6,129	43	0
C2	Hypothèse de variances égales	43,592	0	-3,921	154	0
	Hypothèse de variances inégales			-2,674	46,165	0,01

2.2. Test de normalité

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistiques	ddl	Sig.	Statistiques	ddl	Sig.
R1	0,383	156	0	0,268	156	0
R7	0,113	156	0	0,902	156	0
R11	0,29	156	0	0,423	156	0
R12	0,337	156	0	0,38	156	0
R9	0,379	156	0	0,341	156	0
R15	0,311	156	0	0,234	156	0
R16	0,449	156	0	0,1	156	0
R5	0,273	156	0	0,518	156	0
R6	0,165	156	0	0,902	156	0
R13	0,277	156	0	0,507	156	0
R10	0,483	156	0	0,052	156	0
R8	0,487	156	0	0,099	156	0
R18	0,31	156	0	0,436	156	0
R19	0,16	156	0	0,92	156	0
R21	0,461	156	0	0,117	156	0
R14	0,433	156	0	0,088	156	0
R20	0,267	156	0	0,607	156	0
R17	0,195	156	0	0,916	156	0
R2	0,318	156	0	0,475	156	0
R3	0,429	156	0	0,133	156	0
R4	0,449	156	0	0,11	156	0
Q1	0,477	156	0	0,289	156	0
Q3	0,379	156	0	0,341	156	0
Q4	0,08	156	0,001	0,955	156	0
C1	0,364	156	0	0,599	156	0
C2	0,405	156	0	0,361	156	0

2.3. Test de Mann-Whitney

	U de Mann-Whitney	W de Wilcoxon	Z	Sig.
R1	1549,500	2539,500	-3,602	0,000
R7	2050,500	3040,500	-1,629	0,103
R11	2148,000	8476,000	-1,245	0,213
R12	2215,500	8543,500	-0,988	0,323
R9	1999,000	2989,000	-1,831	0,067
R15	1961,500	2951,500	-1,979	0,048
R16	2248,500	3238,500	-0,849	0,396
R5	2374,500	3364,500	-0,353	0,724
R6	2267,000	3257,000	-0,778	0,437
R13	2068,500	3058,500	-1,558	0,119
R10	1957,500	8285,500	-1,995	0,046
R8	1677,500	2667,500	-3,097	0,002
R18	2172,000	3162,000	-1,151	0,250
R19	1940,500	2930,500	-2,088	0,037
R21	1482,500	2472,500	-3,865	0,000
R14	1802,000	2792,000	-2,607	0,009
R20	2364,500	3354,500	-0,392	0,695
R17	2270,000	3260,000	-0,765	0,445
R2	2038,000	3028,000	-1,678	0,093
R3	2043,500	3033,500	-1,656	0,098
R4	2059,500	8387,500	-1,593	0,111
Q1	1715,500	2705,500	-2,948	0,003
Q3	1961,500	2951,500	-1,979	0,048
Q4	2346,500	8674,500	-0,463	0,643
C1	672,000	7000,000	-10,003	0,000
C2	1163,000	7491,000	-5,824	0,000

2.4. Tableau croisé

Effectif				
		Q2		Total
		0	1	
Y (Passage en Défaut)	0	53	59	112
	1	19	25	44
Total		72	84	156

2.5. Test de Khi-2

	Valeur	ddl	Signification asymptotique (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
khi-carré de Pearson	,218 ^a	1	,641		
Correction pour continuité	,083	1	,773		

Rapport de vraisemblance	,218	1	,640		
Test exact de Fisher				,722	,387
Association linéaire par linéaire	,216	1	,642		
N d'observations valides	156				

Annexe 3 : Etude de dépendance des variables à retenir

3.1. Test de corrélation de Spearman

		R1	R7	R11	R12	R15	R21	C1	C2	R21	R14	Q1	R10	R8
R1	Rho	1												
	Sig	.												
R7	Rho	,237**	1											
	Sig	0,003	.											
R11	Rho	0,132	,708**	1										
	Sig	0,1	0	.										
R12	Rho	0,149	,696**	,991**	1									
	Sig	0,063	0	0	.									
R15	Rho	,334**	,683**	,409**	,409**	1								
	Sig	0	0	0	0	.								
R21	Rho	,174*	,431**	,265**	,278**	,260**	1							
	Sig	0,03	0	0,001	0	0,001	.							
C1	Rho	-,270**	-0,038	0,028	0,015	-0,081	0,053	1						
	Sig	0,001	0,641	0,727	0,851	0,316	0,509	.						
C2	Rho	-0,128	-0,013	0,122	0,121	-0,006	-0,092	,439**	1					
	Sig	0,113	0,876	0,129	0,131	0,942	0,255	0	.					
R21	Rho	,295**	,414**	0,126	0,137	,420**	,229**	-,260**	-,187*	1				
	Sig	0	0	0,117	0,087	0	0,004	0,001	0,019	.				
R14	Rho	-0,019	0,014	-0,118	-0,121	-0,082	-0,05	0,016	-0,081	-,214**	1			
	Sig	0,815	0,861	0,144	0,132	0,311	0,535	0,839	0,312	0,007	.			
Q1	Rho	0,037	0,001	-0,07	-0,06	0,091	0,009	-0,107	0,015	0,015	0,013	1		
	Sig	0,649	0,989	0,383	0,454	0,26	0,911	0,183	0,853	0,852	0,869	.		
R14	Rho	-,451**	-,319**	-,312**	-,326**	-,264**	-,349**	,216**	0,026	-,369**	,265**	0,08	1	
	Sig	0	0	0	0	0,001	0	0,007	0,747	0	0,001	0,318	.	
R8	Rho	,314**	,345**	,198*	,210**	,269**	,561**	-,163*	-,251**	,204*	-0,08	0,115	-,353**	1
	Sig	0	0	0,013	0,009	0,001	0	0,042	0,002	0,011	0,32	0,152	0	.

3.2.Mesures directionnelles : Eta2

			Valeur
Données nominales / intervalle	Eta	Dépendant de R7	,211
		Dépendant de Y (Passage en Défaut)	,683
Données nominales / intervalle	Eta	Dépendant de R11	,214
		Dépendant de Y (Passage en Défaut)	,917
Données nominales / intervalle	Eta	Dépendant de R12	,215
		Dépendant de Y (Passage en Défaut)	,648
Données nominales / intervalle	Eta	Dépendant de R15	,147
		Dépendant de Y (Passage en Défaut)	,877
Données nominales / intervalle	Eta	Dépendant de R20	,140
		Dépendant de Y (Passage en Défaut)	,749

Annexe 4 : Analyse discriminante factorielle

4.1.Somme des poids, probabilités a priori pour chaque classe

Classe	Somme des poids	Probabilités a priori
0	112	,718
1	44	,282

4.2.Valeurs propres

Fonction	Valeur propre	% de la variance	% cumulé	Corrélation canonique
1	1,675 ^a	100,0	100,0	,791

4.3.Lambda de Wilks

Test de la ou des fonctions	Lambda de Wilks	Khi-carré	ddl	Sig.
1	,374	116,114	10	,000

4.4.Coefficients des fonctions discriminantes canoniques

Fonction	(Constante)	C1	R11	R21	R1	C2	R15	R10	R14	R8	Q1
1	-,372	,848	,525	-,423	-,344	,327	-,304	,142	-,092	-,076	-,047

4.5.Fonctions aux centroïdes des groupes

Y (Passage en Défaut)	Fonction 1
0	-,833
1	1,980

Fonctions discriminantes canoniques non standardisées évaluées aux moyennes des groupes

4.6. Probabilités à priori pour les groupes

Passage en défaut	Probabilités à priori	Observations utilisées dans l'analyse	
		Non pondérées	Pondérées
0	,704	88	88
1	,296	37	37
Total	1	125	125

4.7. Coefficients de la fonction de classement

	Passage en défaut	
	0	1
R1	-1,158	3,186
R11	-,034	,318
R15	2,810	1,632
R10	,721	,852
R8	,138	-,030
R21	-18,749	-12,462
R14	,037	,068
Q1	,175	,062
C1	-,035	,284
C2	-,052	,096
(Constante)	-2,439	-6,094

4.8. Résultats du classement (ADF échantillon d'apprentissage)

		Y (Passage en Défaut)	Appartenance au groupe prévu		Total
			0	1	
Original	Effectif	0	110	2	112
		1	16	28	44
	%	0	98,2	1,8	100
		1	36,4	63,6	100

a. 88,5% des observations originales sont classées correctement.

Annexe n°5 : Régression Logistique

5.1. Bloc 0 : Bloc de début

Variables de l'équation

		B	E.S	Wald	ddl	Sig.	Exp(B)
Pas 0	Constante	-,396	,152	6,755	1	,009	,673

Bloc 1 : Méthode = Introduction

5.2. Variables de l'équation

	B	E.S	Wald	ddl	Sig.
R1	,129	,055	5,491	1	,019
R11	,261	,133	3,858	1	,050
R15	-,310	,099	9,922	1	,002
R10	,076	,032	5,752	1	,016
R8	-,123	,078	2,519	1	,112
R21	-,240	,077	9,754	1	,002
R14	-,136	,252	,293	1	,588
Q1	,491	3,616	3,222	1	,073
C1	,690	,030	5,416	1	,020
C2	-,001	,005	,012	1	,912
Constante	-2,677	1,288	4,322	1	,038

5.3. Tests composites des coefficients du modèle

		Khi-carré	ddl	Sig.
Pas 1	Pas	207,558	10	,000
	Bloc	207,558	10	,000
	Modèle	207,558	10	,000

5.4. Récapitulatif des modèles

Pas	Log de vraisemblance -2	R-deux de Cox et Snell	R-deux de Nagelkerke
1	33,700	,686	,927

5.5. Test de Hosmer et Lemeshow

Pas	Khi-carré	ddl	Sig.
1	7,545	6	,273

5.6. Tableau de contingence pour le test de Hosmer et Lemeshow

		Y (Passage en Défaut) = 0		Y (Passage en Défaut) = 1		Total
		Observé	Attendu	Observé	Attendu	
Pas 1	1	18	17,983	0	,017	18
	2	17	17,859	1	,141	18

	3	18	17,698	0	,302	18
	4	17	17,551	1	,449	18
	5	18	17,323	0	,677	18
	6	17	16,125	1	1,875	18
	7	2	2,462	9	8,538	11
	8	0	,000	60	60,000	60

5.7. Table de classification (RL : échantillon d'apprentissage) : modèle de la ligne de base a,b

Observé			Prévisions		
			Y (Passage en Défaut)		Pourcentage correct
			0	1	
Pas 0	Y (Passage en Défaut)	0	112	0	100,0
		1	44	0	,0
	Pourcentage global				
a. La constante est incluse dans le modèle.					
b. La valeur de coupe est ,500					

5.8. Table de classification (RL : échantillon d'apprentissage) : modèle retenu a

Observé			Prévisions		
			Y (Passage en Défaut)		Pourcentage correct
			0	1	
Pas 1	Y (Passage en Défaut)	0	110	2	98,2
		1	5	39	88,6
	Pourcentage global				
a. La valeur de coupe est ,500					

Annexe n°6 : Perceptron multicouche

6.1. Récapitulatif de traitement des observations

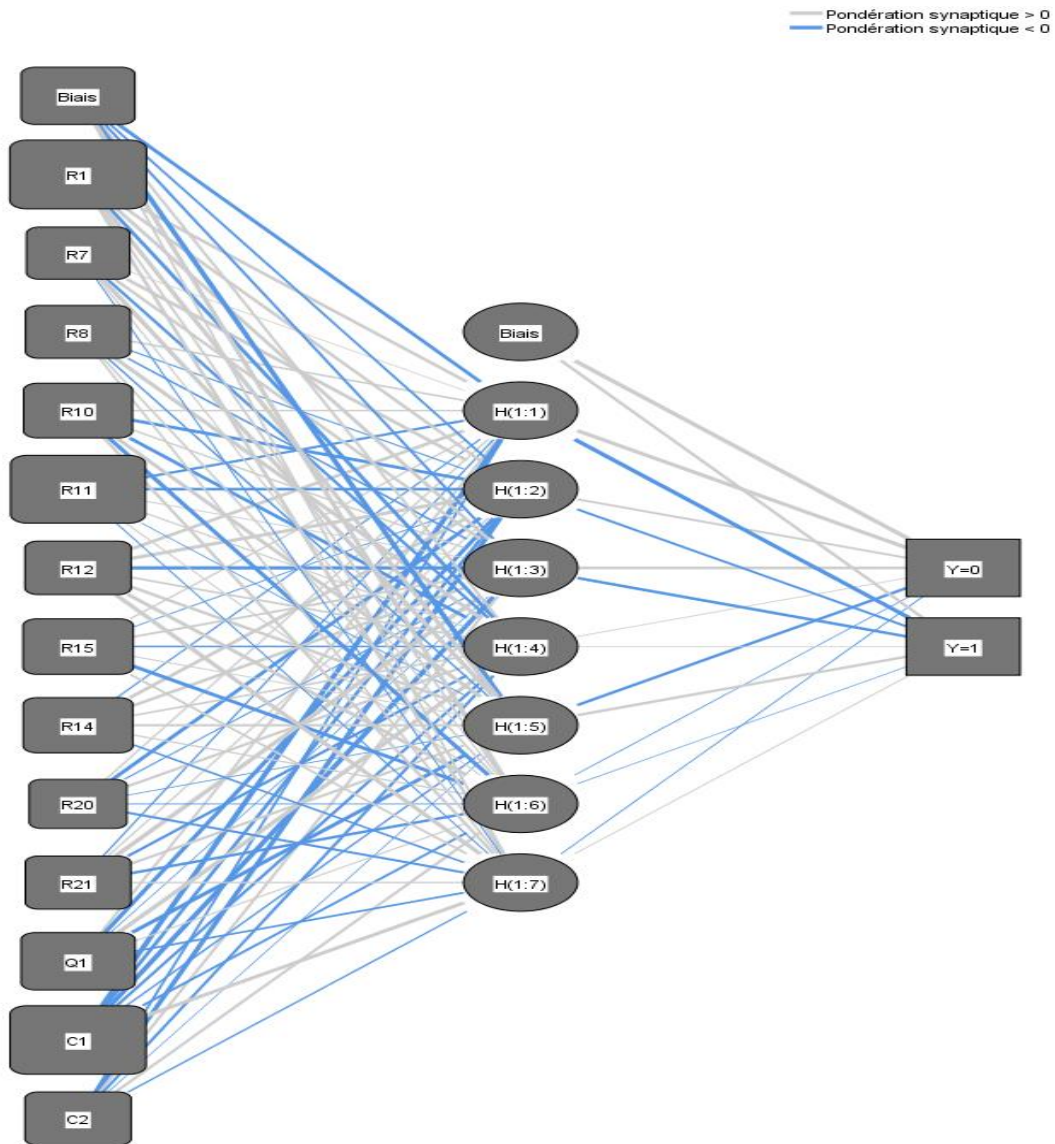
		N	Pourcentage
Echantillon	Apprentissage	117	75%
	Test	19	12,2%
	Élément traité	20	12,8%
Validées		156	100,0%

Exclue	0	
Total	156	

6.2. Informations réseau

Informations réseau			
Couche d'entrée	Covariables	1	R1
		2	R7
		3	R11
		4	R12
		5	R15
		6	R10
		7	R8
		8	R21
		9	R14
		10	R20
		11	Q1
		12	C1
		13	C2
		Nombre d'unités ^a	13
	Méthode de redimensionnement pour les covariables	Standardisé	
Couche(s) masquée(s)	Nombre de couches masquées		1
	Nombre d'unités dans la couche masquée 1 ^a		7
	Fonction d'activation		Tangente hyperbolique
Couche de sortie	Variables dépendantes	1	Y (Passage en Défaut)
	Nombre d'unités		2
	Fonction d'activation		identité
	Fonction d'erreur		Somme des carrés
a. Exclusion de l'unité de biais			

6.3. Architecture du réseau



6.4. Récapitulatif des modèles

Apprentissage	Erreur de somme des carrés	4,319
	Pourcentage de prévisions incorrectes	3,4%
	Règle d'arrêt utilisée	1 pas consécutifs sans diminution d'erreur ^a
	Durée d'apprentissage	0:00:00,05
Test	Erreur de somme des carrés	,205
	Pourcentage de prévisions incorrectes	0,0%
Elément traité	Pourcentage de prévisions incorrectes	10%
Variable dépendante : Y (Passage en Défaut)		
a. Calculs d'erreur basés sur l'échantillon de test.		

6.5. Estimations des paramètres

Prédicteurs		Prévisions								
		Couche masquée 1							Couche de sortie	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	[Y=0]	[Y=1]
Couche d'entrée	(Biais)	-0,399	0,814	-0,587	-0,37	-0,211	-0,284	1,446		
	R1	0,603	0,93	1,381	0,691	1,042	0,983	-0,397		
	R7	0,085	0,054	1,036	0,635	0,707	0,411	0,649		
	R11	-0,735	-0,004	-0,321	-0,532	-0,353	-0,622	-0,497		
	R12	-0,198	-1,314	-0,696	-0,134	0,399	-0,076	0,089		
	R15	0,055	0,119	0,622	0,319	0,386	0,616	0,709		
	R10	0,074	-0,509	-0,313	-0,384	-0,79	0,138	0,313		
	R8	0,286	0,174	-0,279	0,429	0,622	0,084	-0,517		
	R21	-0,112	1,219	1,131	0,611	0,924	0,317	-0,725		
	R14	0,039	0,562	0,7	0,073	0,13	0,361	-0,221		
	R20	0,217	0,126	-0,112	-0,204	-0,327	0,244	0,532		
	Q1	0,805	-0,661	0,035	0,906	0,315	-0,173	-0,178		
	C1	-2,712	-1,581	-3,729	-2,939	-3,226	-2,764	0,741		
	C2	0,056	-0,765	-1,673	-1,271	-1,365	-0,82	-0,693		
Couche masquée 1	(Biais)								-1,232	1,309
	H(1:1)								2,463	-2,643
	H(1:2)								-2,251	1,92
	H(1:3)								2,129	-2,189
	H(1:4)								3,123	-3,434
	H(1:5)								2,852	-2,39
	H(1:6)								1,195	-0,737
	H(1:7)								-3,248	3,253

6.6. Résultats de classification a (RNA échantillon d'apprentissage)

hantillon	Observé	Prévisions		
		0	1	Pourcentage correct
Apprentissage	0	84	0	100%
	1	2	31	93,9%
	Pourcentage global	98,2%		

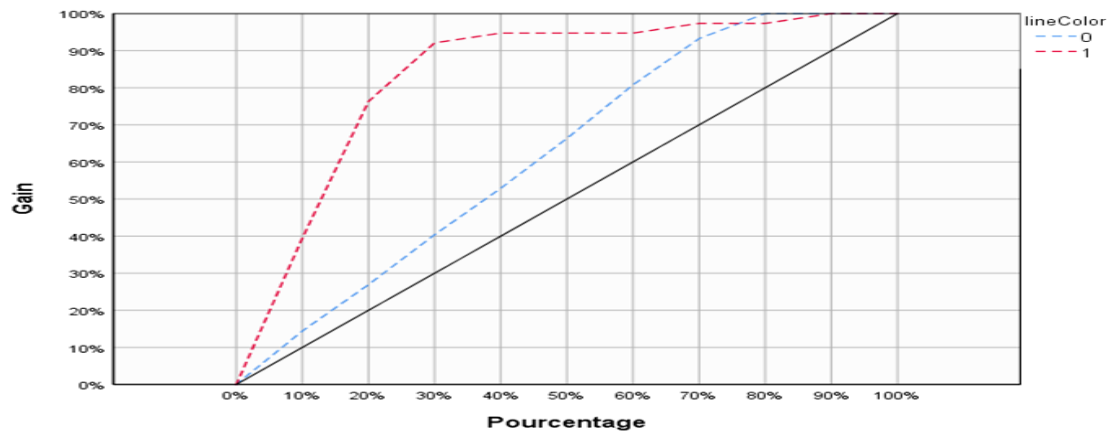
a. Variable dépendante : Y (Passage en Défaut)

6.7. Importance des variables indépendantes

	Importance	Importance normalisée
R11	,179	100%
C1	,172	96,50%

R15	,097	54,30%
C2	,086	48,20%
R14	,083	46,60%
R20	,066	36,90%
R8	,065	36,60%
R10	,061	33,90%
R1	,059	33%
Q1	,048	27%
R21	,033	18,50%
R12	,029	16,40%
R7	,021	11,80%

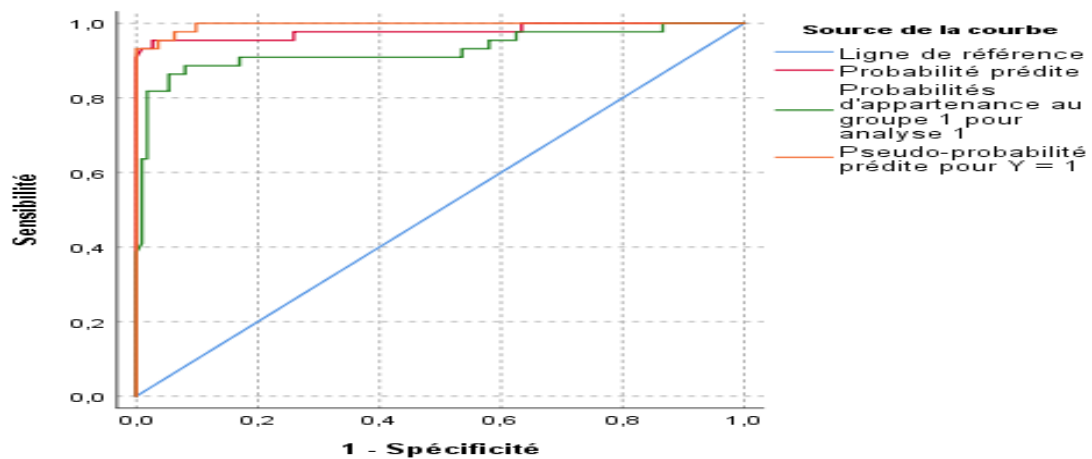
6.8. Graphique de gains cumulés



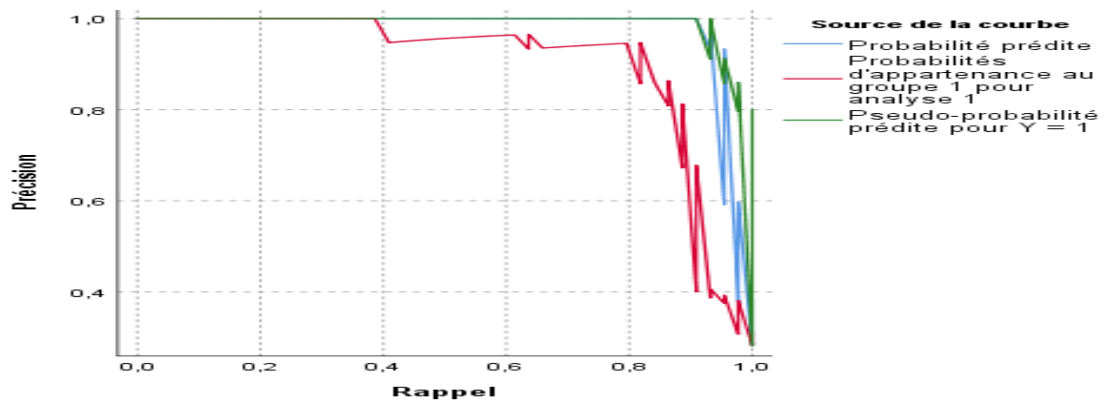
Annexe n°7 : Analyse de performance des modèles

Analyse ROC

7.1. Courbe ROC des 3 modèles



7.2. Courbe PR des 3 modèles



7.3. Zone sous la courbe ROC des 3 modèles

Variable(s) de résultats tests	Zone
Probabilités d'appartenance au groupe 1 pour analyse 1	,927
Probabilité prédite	,979
Pseudo-probabilité prédite pour Y = 1	,996

Les variables de résultats tests : Probabilités d'appartenance au groupe 1 pour analyse 1, Probabilité prédite comportent au moins une liaison entre le groupe d'état réel positif et le groupe d'état réel négatif.
Les statistiques peuvent être déformées.

Annexes n° 8 : Validation « out of sample » des modèles

8.1. Résultats du classement ^c (ADF échantillon de validation croisée)

		Y (Passage en Défaut)	Appartenance au groupe prévu		Total
			0	1	
Validé-croisé ^b	Effectif	0	110	2	112
		1	17	27	44
	%	0	98,2	1,8	100
		1	38,6	61,4	100

b. La validation croisée n'est effectuée que pour les observations de l'analyse. Dans la validation croisée, chaque observation est classée par les fonctions dérivées de toutes les autres observations.

c. 87,8% des observations validées-croisées sont classées correctement.

8.2. Résultats du classement ^c (RL échantillon de test out of sample)

Echantillon	Observé	Prévisions		
		0	1	Pourcentage correct
	0	21	1	97%
	1	1	11	88,89%
Validation	Pourcentage global	94,87%		

Variable dépendante : Y (Passage en Défaut)

8.3.Résultats de classification (RNA échantillon de Test 1 out of sample)

Echantillon	Observé	Prévisions		
		0	1	Pourcentage correct
Test	0	14	0	100%
	1	0	5	100%
	Pourcentage global	100%		

Variable dépendante : Y (Passage en Défaut)

8.4.Résultats de classification (RNA échantillon de Test 2 out of sample)

Echantillon	Observé	Prévisions		
		0	1	Pourcentage correct
Elément traité	0	14	0	100%
	1	2	4	66,7%
	Pourcentage global	90%		

Variable dépendante : Y (Passage en Défaut)

Annexes n° 9 : Validation « out of time » des modèles

9.1.Résultats du classement a (ADF échantillon test out of time)

		Y (Passage en Défaut)	Appartenance au groupe prévu		Total
			0	1	
Validation	Effectif	0	63	36	99
		1	0	41	44
	%	0	63,6	36,4	100
		1	0	100	100

a. 74,3% des observations originales sont classées correctement.

9.2.Résultats du classement a (RL échantillon test out of time)

Echantillon	Observé	Prévisions		
		0	1	Pourcentage correct
Validation	0	95	4	96%
	1	0	41	100%
	Pourcentage global	97,1%		

a. Variable dépendante : Y (Passage en Défaut)

9.3.Résultats du classement a (RNA échantillon test out of time)

Echantillon	Observé	Prévisions		
		0	1	Pourcentage correct
Validation	0	95	4	95,9%
	1	16	25	60,9%
	Pourcentage global	85,7%		
a. Variable dépendante : Y (Passage en Défaut)				

Table des Matières

Dédicaces	I
Remerciements	II
Sommaire	III
Liste des Abréviations	V
Liste des Tableaux	VII
Liste Des Figures	VIII
Liste Des Graphiques	VIII
Introduction Générale	1
Partie Théorique	3
Chapitre I : Risques Bancaires et Réglementation en Vigueur	4
Introduction	4
Section I : Risques Bancaires : Concepts et Typologies	5
I. Typologies des risques bancaires :	5
Section I : Généralités sur le Risque de Crédit	9
I. Risque de crédit : Définition	9
II. Composantes de risque de crédit :	9
III. Risque de crédit : Facteurs :	11
Section III : Réglementation Prudentielle Internationale : Normes Baloises	13
I. Création du comité de Bâle :	14
II. Bâle I : Ratio COOKE :	15
III. Bâle II :	17
IV. Bâle III :	22
V. Bâle IV :	25
Section IV : Réglementation Prudentielle Nationale	26
I. Division et couverture des risques :	26
Conclusion	29
Chapitre II : Méthodes d'Evaluation du Risque de Crédit	30

Introduction	30
Section I : Approche Traditionnelle Basée sur l'Analyse Financière	31
II. Principe de l'approche traditionnelle :	31
III. Avantages de l'approche traditionnelle :	32
IV. Limites de l'approche traditionnelle:	32
Section II : Notation Externe : « Rating »	34
I. Principe de la méthode :	34
II. Output du rating :	35
III. Avantages de la notation externe :	37
IV. Limites de la notation externe :	37
Section III : Approche Statistique de « Crédit Scoring » :	39
I. Crédit scoring : principe, méthodologie et qualité de prédiction :	40
II. Modèles paramétriques :	43
III. Modèles semi-paramétriques	49
Section IV : Approche Basée sur l'Intelligence Artificielle	55
I. Partitionnement récursif :	55
II. Réseaux de neurones artificiels :	58
Conclusion :	62
Partie Empirique	63
Chapitre I : Elaboration d'un Modèle de Scoring : Cas de la BNA	64
Introduction :	64
Section I : BNA en Chiffres : Indicateurs Clés de Performance	64
I. Indicateurs d'exploitation : produit net bancaire (PNB) :	65
II. Indicateurs de rentabilité :	65
III. Indicateurs de risque de crédit :	66
Section II : Méthodologie Empirique : Description des Données et Présentation des Variables	68
I. Méthodologie empirique :	68
Section III : Etude des Pouvoirs Discriminants : Sélection Préalable des Variables	75
I. Etude d'association des variables : analyse bivariée :	75
II. Etude de dépendance des variables :	79
III. Test de multi-colinéarité :	81
Conclusion :	84

Chapitre II : Modélisation du Risque de Crédit	85
Introduction :	85
Section I : Costruction d'un modèle de « Crédit Scoring »	85
I. Scoring par l'analyse discriminante factorielle (ADF) :	86
II. Régression logistique (RL) :	89
III. Scoring par les Réseaux de neurones artificiels (RNA) :	93
Section II : Etude Comparative de Performance des Modèles	98
I. Performance des modèles :	98
II. Analyse ROC :	99
III. Courbe PR (Précision-Rappel) :	102
IV. Généralisation du modèle :	103
V. Modèle retenu : Régression logistique :	107
Conclusion :	111
Conclusion Générale	112
Bibliographie	115
Webliographie	121
Annexes	122
Table des Matières	139
Résumé :	142
Abstract:	142

Résumé :

La nature de l'activité bancaire expose les établissements bancaires à de nombreux risques, entre autres, le risque de défaut. Cette exposition peut engendrer des pertes comme elle peut assurer des gains énormes.

Dans ce travail de recherche, nous nous sommes penchés sur l'élaboration un modèle de « crédit scoring » fiable et assez robuste offrant une meilleure visibilité sur la situation financière des relations de la BNA, en général, et sur le degré de solvabilité en particulier.

Pour effectuer ce travail, nous avons donc modélisé à la fois la technique de l'analyse discriminante, de la régression logistique et des réseaux de neurones artificiels.

Chaque modèle estimé a montré une capacité importante de prédire le défaut des entreprises.

Pour arriver à une décision finale portant sur l'estimation à retenir, en se basant sur les deux caractéristiques d'un bon modèle de scoring tels que, le bon classement et la stabilité des résultats, nous avons conclu que notre choix final est la régression logistique.

Cette régression regroupe sept variables discriminantes dont cinq ratios reflétant la situation financière de la contrepartie, une variable comportementale décrivant l'attitude historique de la relation et une autre qualitative relative au secteur d'activité de cette dernière.

Abstract:

The nature of the banking business exposes banking institutions to many risks, including the risk of default. This exposure can result in losses as well as huge gains.

In this research, we focused on the development of a reliable and robust model of "credit scoring" offering a better visibility on the financial situation of the relations of the BNA, in general, and on the degree of solvency in particular.

To carry out this work, we have therefore modeled both the technique of discriminant analysis, logistic regression and artificial neural networks.

Each estimated model showed a significant capacity to predict corporate default.

To arrive at a final decision on which estimation to retain, based on the two characteristics of a good scoring model such as, good ranking and stability of results, we concluded that our final choice is logistic regression.

This regression combines seven discriminant variables, including five ratios reflecting the financial situation of the counterparty, a behavioral variable describing the historical attitude of the relationship and another qualitative variable relating to the sector of activity of the counterparty.