

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

ECOLE SUPERIEURE DE COMMERCE

Pôle Universitaire de Kolea

**Mémoire de fin d'études en vue de l'obtention du diplôme de Master en
Sciences Financières et Comptabilité**

Spécialité : FINANCE, MONNAIE ET BANQUE

THEME :

Gestion de risque crédit par la méthode de scoring

Elaboré par :

Achemoun Nedjma

Encadreur :

Pr.Gliz Abdelkader

Lieu du stage :La Direction de Surveillance du Risque Crédit –DSRC- ; CPA-Banque

Période du stage : du01/04/2020 au 30/6/2020

2019/2020

REMERCIEMENTS

Je remercie tout d'abord Dieu de m'avoir donné la force, la volonté et le courage pour la réalisation de ce travail

Je voudrais adresser mes sincères remerciements à mon encadreur Pr Gliz Abdelkader pour sa patience, sa disponibilité, ses orientations, et surtout ses judicieux conseils qui ont contribué à alimenter ma réflexion.

Je tiens à remercier chaleureusement mon encadreur de l'établissement de stage Mr. Merabti Rabah pour son aide, ses orientations ainsi que pour son chaleureux accueil.

Je remercie également tous les enseignants et travailleurs de l'Ecole Supérieure de Commerce pour ces cinq agréables années.

Je voudrais aussi exprimer ma reconnaissance envers les amis et collègues qui m'ont apportés leur soutien moral et intellectuel.

Enfin je remercie toutes les personnes qui ont contribué à l'élaboration de ce mémoire de près ou de loin.

Dédicaces

Avec l'expression de ma reconnaissance, je dédie ce modeste travail à ceux qui quelquesoit les termes que je dirai ne seront jamais assez pour exprimer mon amour pour eux :

A mon précieux offre de Dieu, à qui je dois ma vie, ma réussite et tout mon respect : mon cher papa, ta bienveillance m'a guidée et ta présence à mes cotés a toujours été ma source de force pour affronter les différents obstacles.

A la femme qui a toujours crue en moi et a su comment me motiver et me soutenir, à ma chère mère qui m'a appris à être patiente dans tout ce que j'entreprends.

A ma grande mère, ma deuxième mère, tes prières et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études. Puisse Dieu, le tout puissant, te préserver et t'accorder santé et longue vie.

A vous, mon frère « Hakim » et mes sœurs « Amel », « Mélissa » et « Tinhinane » qui m'avez toujours soutenu et encouragé durant ces années d'études.

A mon fiancé qui n'a pas cessé de me conseiller, encourager et soutenir, que Dieu le protège et lui offre tout le bonheur.

A tout les membres de ma grande famille.

Liste des tableaux

Tableau 1 : Les pondérations d'engagements de bilan

Tableau 2 : Synthèse des exigences en fonds propres de Bâle 3

Tableau 3 : Classement et provisionnement des créances en ex-post

Tableau 4 : Table d'affectation

Tableau 5: Les formes juridiques

Tableau 6 : Les variables sélectionnées

Tableau 7 : Test de spécification du modèle

Tableau 8 : Résultat de la régression logistique

Tableau 9 : Tableau récapitulatif

Tableau 10 : Défaillances observées et prédites

Tableau 11 : Test de Hosmer-Lemeshow

Tableau 12 : Surface sous la courbe ROC

Tableau 13: Tableau de classement

Liste des figures

Figure 1: Description des étapes à suivre pour la sélection des variables lors de la création du modèle.

Figure 2 : Taux de VP et de FP pour différents seuils de classification

Figure 3: Organisation du CPA

Figure 4 : Organisation de la DSRC

Figure 5 : Secteurs d'activité des entreprises de notre échantillon

Figure 6 : La courbe ROC de la régression logistique

Liste des annexes

Annexe 1 : Base de données des variables significatifs

Annexe 2 : Tableau de classement avant la régression logistique

Annexe 3 : Tableau de corrélation

Liste des abréviations

ADL:Analyse discriminante linéaire

BFR:Besoin en fond de roulement

CA: Chiffre d'affaire

CAF: Capacité d'autofinancement

CMC: Conseil de la monnaie et du crédit

FP: Fonds propres

FR: Fond de roulement

IRB: Internal ratings based

LCR: Liquidity coverage ratio

NSFR: Net stable funding ratio

RL: Régression logistique

VA: Valeur ajoutée

MVC: Mouvement de chiffres d'affaires confies

CAF:Capacité d'autofinancement

CPA:Crédit populaire d'Algérie

DLMT: Dette a long et moyen termes

PNC: Passif non courant

PD: Probabilité de défaut

PDG: Président directeur général

EURL: Entreprise unipersonnelle à responsabilité limitée

SARL: Société à responsabilité limitée

SNC: Société en nom collectif

SPA: Société par actions

Sommaire

Liste des tableaux

Liste des figures

Liste des annexes

Liste des abréviations

Introduction général

Chapitre I : Le risque de crédit

Section1 : Présentation de risque de crédit

1	Définition de risque de crédit	16
2	Typologie de risque credits	17
2.1	Le risque de défaut.....	17
2.2	Le risque de dégradation de Spread.....	17
2.3	Le risque de recouvrement.....	17
3	L'impact de risque credit sur l'activite bancaire	18

Section 2 : La réglementation prudentielle internationale

1	Définition de la reglementation prudentielle	18
2	Présentation des recommandations de bale I.....	19
2.1	Le comité de Bâle I.....	19
2.2	Recommandations de Bâle I :	19
2.3	Amendement de Bâle I	20
3	L'accord de bale 2	21
3.1	Calcul de risque crédit	22

3.2	Insuffisances de Bâle II	23
4	Les accords de bale iii :	24
4.1	Renforcement des fonds propres	24
4.2	Instauration de ratio de liquidité :	25
4.2.1	Le liquidity coverage ratio (LCR).....	25
4.2.2	Le Net Stable Funding Ratio (NSFR)	25
4.2.3	Mise en place d'un ratio effet de levier :	25

Section 3 : Les outils de gestion de risque crédit

1	Les garanties.....	26
1.1	Les garanties réelles.....	26
1.1.1	L'hypothèque	26
1.1.2	Le nantissement.....	27
1.1.3	Le gage.....	27
1.2	Les garanties personnelles	27
1.2.1	Le cautionnement	27
1.2.2	L'aval	28
2	Les garanties financieres (societes de contre garanties).....	28
3	Les assurances	28
4	Strategie de provisionnement du risque credit	29
4.1	Le provisionnement des pertes subies (ex-post).....	29
4.1.1	Présentation de provisionnement ex-post	29
4.1.2	Provisionnement ex-post selon les normes algériennes.....	29
4.2	Présentation de provisionnement ex-ante	30
7.2.1	Méthode de provisionnement ex-ante	31
5	Les dérivés crédit : Instrument de transfert du risque	31
5.1	Les dérivés de crédit	32
5.2	Paramètre du dérivé de crédit	32
6	La titrisation :.....	33

Chapitre II : Modèle de prévision de la défaillance: Le Crédit Scoring

Section 1: Présentation de crédit scoring

1	Définition de credit scoring	37
2	Historique de credit scoring	37
3	Avantages et inconvénients de credit scoring	39
3.1	Les avantages du crédit scoring	39
3.1.1	Atouts spécifiques à l'outil	39
3.1.2	Atouts spécifiques à l'établissement : Un établissement peut choisir de mettre en place un crédit scoring pour différentes raisons :	39
3.2	Les inconvénients du credit scoring	40
3.2.1	Les problèmes pratiques.....	40
3.2.2	Les problèmes de méthodologie	41

Section 2: Méthodes d'élaboration de modèle scoring

1	Les types de modele de crédit scoring	42
1.1	Le modèle déductif	42
1.2	Le modèle empirique	42
2	Démarches de construction d'un modèle scoring	43
2.1	Collection des données	43
2.1.1	Choix de critère de défaut et construction de l'échantillon	43
2.2	Spécificités des données	44
2.2.1	Définition de l'horizon temporel	44
2.2.2	Répartition des données	45
2.3	Traitement des valeurs manquantes et des valeurs aberrantes.....	45
2.3.1	Les valeurs manquantes	45
2.3.2	Les valeurs aberrantes	48
2.4	Le choix des variables explicatives	48
2.5	Choix de la méthode statistique.....	48
2.5.1	L'approche paramétrique	49
2.5.2	L'approche non paramétrique	49

2.6	Modélisation et tests	50
2.7	La validation	50

Section 3: Techniques de classification et validation de modèle de crédit scoring

1	L'analyse discriminante	50
1.1	Définition.....	50
1.2	Construction de modèle	51
1.2.1	L'approche géométrique	51
1.2.2	L'approche probabiliste	51
1.3	Sélection des variables	51
1.4	Test de validation	53
1.4.1	Méthode de resubstitution	53
1.4.2	Méthode de l'échantillon-test.....	53
2	La régression logistique :	54
2.1	Définition :.....	54
2.2	Construction du modèle régression logistique binaire :	55
2.3	La sélection des variables :	56
2.3.1	La selection ascendante : (Forward selection)	56
2.3.2	La selection descendante : (backward selection)	57
2.3.3	La selection pas à pas : (stepwise)	57
2.4	Interprétation des paramètres :.....	57
2.5	Validation du modèle :.....	58
2.5.1	Le test de Wald :	58
2.5.2	Test de significativité globale :	59
2.5.3	Test de Hosmer et Lemeshow :	59
2.5.4	Le coefficient de détermination généralisé :	60
3	La régression probit :.....	60

Section 4: Performance de modèle de crédit scoring

1	le taux de bon classement :.....	61
2	La courbe : ROC	62

3	La méthode de l'échantillon-test	63
---	--	----

Chapitre III: La mise en œuvre d'un crédit scoring au niveau du CPA

Section 1:Présentation de crédit populaire d'Algérie

1	Présentation générale.....	66
2	Axes stratégiques.....	67
3	Organisation de la banque	67
4	La direction de surveillance de risque crédit –DRSC-	69

Section 2 : Elaboration de modèle de crédit scoring

1	Construction de la base de données :	70
2	Critère de défaut :	71
3	Construction de modèle :	71
4	Validation :	75
4.1	Test de Significativité globale :	75
4.2	Test de Hosmer-Lemeshow :	76
5	Performance du modèle:.....	77
5.1	La courbe ROC: (Receiver Operating Characteristic):.....	77
5.2	Le taux de bon classement :.....	78

Conclusion

Conclusion générale

Annexes

Bibliographie

Résumé

Introduction générale

Les établissements bancaires jouent un rôle très important au sein du système financier. Ils font partie de ses principaux intermédiaires financiers. Ils interviennent entre les prêteurs et les emprunteurs, par la collecte de l'épargne des agents ayant une capacité de financement puis la distribuer aux agents ayant un besoin de financement.

Les établissements de crédit par la nature de leurs activités se trouvent face à différents risques. Leurs préoccupations sont d'arriver à maîtriser ces risques et minimiser les pertes qui peuvent en résulter, afin d'assurer leur pérennité dans un environnement de plus en plus incertain.

La gestion du risque a fait l'objet de plusieurs recherches mettant en évidence la multiplicité des risques bancaires. La banque s'expose à une variété de risques mais le risque de crédit est considéré comme le risque principal auquel une banque est confrontée. Il correspond à l'incapacité de l'emprunteur à payer l'intérêt dû et à rembourser le principal selon les termes spécifiés dans la convention de crédit. Il se manifeste par l'importance des crédits non performants ou des créances douteuses.

Pour prendre leurs décisions de manière rationnelle, les prêteurs doivent mesurer avec précision le risque de crédit des emprunteurs, autant avant de leur accorder un crédit (pour en fixer les conditions : montant, maturité, taux et garanties) que durant la phase d'exploitation. La gestion du risque bancaire correspond à l'ensemble des techniques, outils et dispositifs mis en place par la banque pour identifier, mesurer et surveiller les risques auxquels elle est confrontée.

Le Scoring en tant que technique d'évaluation des risques, a été développé dans cette optique. Bien que les définitions attribuées à la méthode de Scoring sont différentes d'un auteur à un autre, il est généralement admis que cette méthode est un outil de mesure des risques qui vise à prédire la probabilité de défaut d'un nouveau prêt en utilisant l'information des prêts précédents.

Notre mémoire de fin d'étude s'inscrit dans le contexte du développement de l'une des méthodes destinées à la réduction du risque de crédit et du taux de défaut de remboursement, « Le Crédit Scoring ». Cela nous ramène à poser la problématique suivante :

**Comment les banques peuvent-elles évaluer d'une manière efficace le risque crédit ?
Et quelle démarche à suivre pour mettre en place un système de score performant ?**

Dans le souci d'apporter des éléments de réponse, notre raisonnement répond progressivement aux questions suivantes :

-)] Qu'est-ce qu'un crédit bancaire et quels sont les conséquences provenant de ses risques ?
-)] En quoi consiste un modèle de Credit Scoring ? Et quelle est la méthodologie de son élaboration ?
-)] Comment obtenir un meilleur modèle de crédit scoring à partir des différentes méthodes statistiques ? et selon quels critères choisit-on celle qui donne de meilleurs résultats en matière de performance ?

Les hypothèses

- Le Scoring propose une base méthodologique de mesure du risque de crédit applicable aux emprunteurs à l'intention de la banque.
- Le crédit scoring sert à prévenir les risques de défaillances en mettant en évidence des variables qui influencent négativement sur le comportement des clients.
- Le taux d'endettement est un indicateur de défaillance des entreprises.

Méthodologie de recherche

Notre étude, en s'appuyant sur une méthodologie à la fois quantitative (modélisation statistique) et qualitative se propose de fournir un modèle de prédiction des risques de crédit. Ce modèle sera développé par la modélisation mathématique (régression logistique). A terme, le système d'estimation attendu permettra aux gestionnaires et aux analystes de crédit de prévoir, avec un très bon seuil de confiance, le niveau de risque lié à chaque nouveau demandeur de crédit.

Plan de travail :

Le présent travail est scindé en trois chapitres :

Le premier chapitre traite des aspects théoriques relatifs aux risques bancaires et les instruments de sa gestion.

Quant au deuxième chapitre, il est consacré à la méthodologie d'élaboration de modèle scoring et aux différentes méthodes statistiques qui peuvent être utilisées.

Le dernier chapitre constituera notre cas pratique où on mettra en application les différents aspects que nous avons traités dans les parties théorique

Chapitre I : Le risque de crédit

Introduction

La banque en tant qu'intermédiaire financier est au cœur du système de financement de l'économie est le principal offreur de crédit sur le marché de la dette. Elle fait face à une grande variété de demandeurs :: particuliers, professionnels, entreprises, collectivités. Tous demandent à leur banque de financer leur effet de levier afin d'obtenir des fonds pour acquérir un bien ou un actif qu'ils n'ont pas les moyens de financer par leur seuls fonds propres.

De par son activité, la banque est exposée à une multitude de risques qui peuvent entraîner sa défaillance. Le risque de crédit est le plus ancien des risques, il est aussi le plus important et le plus présent dans l'activité bancaire. La non maîtrise de ce dernier peut mettre la banque en péril.

Ce chapitre sera divisé en trois sections, la première portera sur le risque de crédit et son impact sur la banque. Dans la deuxième section, nous évoquerons l'ensemble des réglementations prudentielles qui lui sont destinées et dans la dernière section, nous citons quelques outils de gestion du risque de crédit.

Chapitre I : Le risque crédit

Section 1 : Présentation de risque crédit

L'environnement bancaire international a été caractérisé cette dernière décennie par une montée sérieuse du risque de crédit. Ce dernier est sûrement le risque le plus délicat pour une banque. C'est pourquoi, il est crucial de bien le cerner et le gérer afin de minimiser son impact sur la banque et sur l'économie toute entière.

1 Définition de risque de crédit

Le risque de crédit, également appelé risque de défaut ou de contrepartie, survient lorsque l'emprunteur se trouve dans l'incapacité de payer ses dettes et remplir ses engagements vis-à-vis de ses créanciers à l'échéance convenue. Il se compose de deux notions : le risque et le crédit

Le risque

NALLEAU.G et ROUACH.M désignent le risque comme « un engagement portant une incertitude dotée d'une probabilité de gain et de préjudice, que celui-ci soit une dégradation ou une perte »¹

Le risque est la probabilité qu'un événement ou un ensemble d'événements ait lieu, qui sont généralement indésirable. Concernant les établissements financiers la survenance de cet événement entraînerait plutôt leur faillite que leurs succès.

Le crédit

Le mot crédit provient du latin « cred tum » qui signifie croire ou avoir confiance².

Selon Charles Petit –Dutaillis « faire crédit,c'est faire aussi donner librement la disposition affective et immédiate d'un bien réel ou d'un pouvoir d'achat ,contre la promesse que le même bien ou l'équivalent vous sera restitué dans un certain délai,le plus souvent avec la rémunération du service rendu et du danger couru,danger de perte partielle ou totale que comporte la nature même de ce service »³

G.PRANSON donne une définition précise pour le risque crédit, « le risque crédit ou le risque de contrepartie représente la perte potentielle réalisée par la banque dans l'hypothèse d'une défaillance future de sa contrepartie. Le risque de crédit peut être défini comme la perte totale enregistrée sur une opération suite à la défaillance de la contrepartie.On l'appelle aussi parfois risque de signature.Il est courant d'employer le terme de risque de contrepartie pour désigner exclusivement le risque de crédit ».⁴

¹Naulleau Gérard et Rouach Michel (1998), le contrôle de gestion et financier, revue bancaire.

²<https://www.researchgate.net>

³Charles Petit-Dutaillis, Le risque de crédit bancaire, Edition scientifique Ribier, Paris, 1967.

⁴<http://www.pandat.fr> G.Pranson,Gestion du risque de contrepartie bancaire.

2 Typologie de risque crédits

On distingue trois types de risque crédit : le risque de défaut, le risque de dégradation du Spread et le risque lié à l'incertitude du recouvrement une fois le défaut survenu.

2.1 Le risque de défaut⁵

Cette forme de risque est associée à l'occurrence d'un défaut caractérisée par l'incapacité de la contrepartie à assurer le paiement de ses échéances. Le comité de Bâle dans son second document consultatif considère un débiteur est en défaut lorsque l'un ou plusieurs des événements suivants est constaté :

- ❖ L'emprunteur ne remboursera vraisemblablement pas en totalité ses dettes (principal, intérêts et commissions) ;
- ❖ La constatation d'une perte portant sur l'une de ses facilités : comptabilisation d'une perte, restructuration de détresse impliquant une réduction ou un rééchelonnement du principal, des intérêts ou des commissions ;
- ❖ L'emprunteur est en défaut de paiement depuis quatre-vingt dix(90) jours sur l'un de ses crédits ;
- ❖ L'emprunteur est en faillite juridique.

Pour sa mesure il existe trois paramètres :⁶

- ❖ Probabilité de défaut (PD) : qui est la probabilité qu'une contrepartie soit défaillante dans un horizon temporel déterminé, un an en général ;
- ❖ Exposition au défaut (EAD) : elle représente le montant des risques sur une contrepartie au moment de la défaillance ;
- ❖ Perte en cas de défaut (LGD) : représente les montants qui ne pourront pas être récupérés.

2.2 Le risque de dégradation de Spread

Le Spread de crédit est la prime de risque qui lui est associée. Sa valeur est déterminée en fonction du niveau de risque encouru (plus le risque est élevé, plus le Spread est élevé aussi). Le risque de dégradation du Spread est le risque de voir se dégrader la qualité de la contrepartie (dégradation de sa note) et donc l'accroissement de sa probabilité de défaut. Cela conduit à une hausse de sa prime de risque, d'où la baisse de la marge sur intérêts.

Ce risque peut être mesuré d'une façon séparé pour chaque contrepartie ou globalement sur tout le portefeuille de crédit.

2.3 Le risque de recouvrement

Le taux de recouvrement permet de déterminer le pourcentage de la créance qui sera récupéré en entreprenant des procédures judiciaires suite à la faillite de la contrepartie. Le recouvrement portera sur le principal et les intérêts après déduction du montant des garanties préalablement recueillis. Le taux de recouvrement constitue une source d'incertitude pour la banque dans la mesure où il est déterminé à travers l'analyse de plusieurs facteurs :

- ❖ La durée des procédures judiciaires qui varient d'un pays à un autre et d'un cas à l'autre ;

⁵<https://www.academia.edu>

- ❖ La valeur réelle des garanties ;
- ❖ Le rang de la banque dans la liste des créanciers.

3 L'impact de risque crédit sur l'activité bancaire

Les risques trouvent leur traduction comptable dans les dotations aux comptes de provisions et /ou les passages à pertes et profits des établissements. Ces provisions et pertes dégradent le résultat net de la banque. Elles provoquent aussi une série d'effets en chaîne qui affaiblissent la situation financière de la banque et la rendent plus sensible. Donc le risque crédit n'est pas sans effet sur la banque et peut provoquer ⁷ :

- ❖ La dégradation de ses résultats : cette dégradation est due aux provisionnements et aux pertes liées au non remboursement des créances ;
- ❖ La dégradation de sa solvabilité : en cas d'insuffisances des résultats pour absorber les pertes constatées, la banque sera dans l'obligation de les imputer sur ses fonds propres, ce qui peut conduire à son insolvabilité définitive à long terme ;
- ❖ La baisse de son rating : les résultats de la banque font partie des paramètres pris en considération par les agences de notation. Donc la diminution de ses résultats peut ramener ces agences à revoir à la baisse le rating de cette banque.
- ❖ Un risque systémique : le risque crédit peut entraîner par effet de contagion une crise systémique. En effet, la faillite d'une banque peut provoquer celle d'une autre. La banque en faillite ne remboursant pas celle qui lui a prêtée des fonds sur le marché interbancaire peut lui causer des problèmes de liquidité et donc entraînera sa chute.

Section 2 : La réglementation prudentielle internationale

L'activité bancaire demeure fortement réglementée du fait du rôle particulier joué par les établissements financiers et des risques qu'ils font courir à l'économie.

1 Définition de la réglementation prudentielle :

La réglementation prudentielle est l'ensemble des contraintes imposées aux établissements de crédits. Elle regroupe tous les dispositifs légaux et éthiques, permettant de protéger les épargnants contre les risques excessifs lors de leur placements et investissements, notamment au sein d'un capital de sociétés de fonds, banques, sociétés d'assurances, sociétés de placements, entreprises d'investissement, etc. ⁸

⁷ Boussoul Lamia « la notation interne : nouvel outil de gestion du risque de crédit », mémoire de fin d'étude, Ecole supérieure de banque, Alger, 2007, P06.

⁸ Adli Rahma et Habbi Hassina « Application de la réglementation prudentielle dans le Système bancaire algérien », mémoire de master en sciences économique, FSESG, UMMTO, 2017, P56.

2 Présentation des recommandations de Bâle I

2.1 Le comité de Bâle I

Le Comité de Bâle est un organisme de réflexion et de proposition sur la supervision bancaire, créé en 1974. Il est domicilié à la Banque des Règlements Internationaux (BRI) à Bâle en Suisse d'où son nom de « Comité de Bâle ». Il est composé des dix pays les plus industrialisés (G10). Aujourd'hui, beaucoup plus de pays ont adhéré à cet accord.

Son objectif est d'intensifier la coopération entre les autorités nationales chargées du contrôle bancaire afin de renforcer la stabilité et la solidité du système bancaire international et atténuer les inégalités concurrentielles existantes entre les banques internationales en établissant des normes prudentielles et des méthodes de surveillance bancaire.

2.2 Recommandations de Bâle I :

Le comité de G10 a signé un accord en 1988 (entré en vigueur en 1992) pour réglementer les banques.

Ce premier accord est limité au risque de crédit. Chaque banque est requise de conserver une réserve de capital de 8 % (ratio de Cooke) de la valeur des titres représentant son portefeuille. Ce ratio sert à créer une réserve de solvabilité pour la banque. Il est donné par la forme suivante :⁹

$$R \quad C = \frac{F \quad p \quad n}{R \quad d \quad c \quad éd} \geq 8\%$$

En 1988 deux formes de capital ont été considérées :

Fonds propres de base + Fonds propres supplémentaires = Fonds propres réglementaires

Fonds propres de base : Type 1 (Tier 1) ou capital de base (capital de première qualité), composés d'actions, participations dans des filiales et certains réserves déclarées à l'agences de réglementation.

Fonds propres supplémentaires : Type 2 (Tier 2) ou capital supplémentaires, composés d'instruments financiers hybrides (actions et débetures avec très longues échéances), dettes subordonnées avec maturité plus de 5 ans, autres titres et autres réserves.

➤ 50% des fonds propres nets doit être couvert par le type 1 et la somme doit représenter au moins 8% des actifs des risques pondérés de la banque. En plus des réserves requises, l'accord imposait des restrictions sur le comportement de prise de risque excessive :

- Pas de participations supérieures à 25 % dans le capital d'une entreprise ;
- Le total de participation vis-à-vis des clients ayant bénéficié de crédits excédant 15% des fonds propres de la banque ne peut dépasser dix fois ces fonds propres.

Les encours pondérés de crédit concerne le bilan et le hors bilan sont décrits dans le tableau suivant :

⁹George Dionne, Gestion des risques : Histoire, définition et critiques, support de cours au centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport.

Tableau 1 : Les pondérations d'engagements de bilan

Classes d'actifs	Pondération
Créances sur les états de l'OCDE	0%
Créances sur les banques et les collectivités locales d'états de l'OCDE	20%
Créances garantis par une hypothèque ou crédit bail immobilier	50%
Tous les autres éléments d'actifs, dont les crédits à la clientèle	100%

Source : Elaboré par nous- même

Pour les engagements de hors bilan, deux types de pondérations peuvent être distingués :¹⁰

- Engagements classiques non liés au cours de change et au taux d'intérêt: la pondération consiste à convertir l'engagement en équivalent crédit par un coefficient de conversion allant de 0% à 100% en fonction de leur nature, puis pondérés selon le statut de la contrepartie.
- Engagement liés au cours de change et /ou au taux d'intérêt : l'équivalent risque est la somme : du coût de remplacement total des contrats représentant un gain, et du risque potentiel produit du nominal par un coefficient de majoration dépendant de la durée résiduelle et de la nature du contrat.

2.3 Amendement de Bâle I

La réglementation a évolué progressivement pour prendre en compte les risques de marché liés aux variations des prix de marché, cours des actifs financiers, des devises et des produits de base. Il était identifiée en 1994/95 et pris en considération dans le calcul de ratio de solvabilité en 1996. D'où le ratio Cooke (corrige de risque de marché) qui devient :

$$R C = \frac{F p n}{R d c \text{ éd} + R q d m \text{ é}} \geq 8\%$$

Contrairement au risque crédit, deux approches existent pour calculer les risques de marché. La première approche est une méthode forfaitaire, alors que la seconde approche autorise les banques à utiliser des modèles internes.¹¹

1.2.2.4 Insuffisances de Bâle I

L'accord de 1988 a été jugé insuffisant pour les raisons suivantes :¹²

- L'absence de flexibilité : non prise en considération des innovations financières (produits dérivés, ...)
- La prise en charge de risque de marché en 1996 n'a pas suffit.
- Etant un compromis, le niveau de 8% manque de fondement économique.

¹⁰ P.Dumontier et D.Dupre, Pilotage bancaire : les normes IAS et la réglementation Bale II, Ed. Revue banque.

¹²La réglementation prudentielle, Cours de Mr. Ilmane, Ecole supérieure de commerce.

- Les coefficients de pondérations ne prennent en considération que la nature de la contrepartie et néglige d'autres éléments : la durée des engagements, la notation de la contrepartie, ...
- Les risque des différents crédits sont simplement additionnés, c'est-à-dire qu'ils sont considérés comme mutuellement indépendants.
- Il est constaté enfin que pratiquement toutes les banques ayant fait faillite depuis auraient respectés ce ratio.

3 L'accord de Bâle 2

Le deuxième accord de Bâle a eu lieu en 1999, appelé « réforme » plutôt que « accord » puisque ces accords sont venus pour remédier aux faiblesses proclamés concernant la pratique du ratio Cooke.

Les accords de Bâle II reposaient sur trois piliers. :

- Le premier visait à définir les exigences minimales de fonds propres des banques.
- Le deuxième instaurait le principe d'une surveillance prudentielle individualisée.
- Enfin, le troisième se focalisait sur les notions de transparence et de discipline de marché.

Pilier I : exigence minimales de fonds propres des banques :

Bâle II a pris en considération trois types de risques : le risque de crédit, risque de marché, ainsi que le risque opérationnel qui est « ...le risque de pertes directes ou indirectes résultant d'une inadéquation ou d'une défaillance attribuable aux procédures, au facteur humain et aux systèmes ou à des facteurs externes »¹³.

Ce nouveau ratio appelé ratio MC Donough est donné par le rapport suivant :

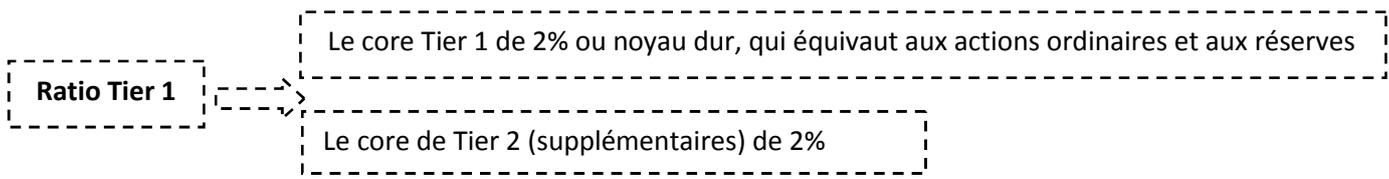
$$R C = \frac{R_{d c \acute{e}d} + R_{d m \acute{e}} + R_{o \acute{e}r}}{\frac{F}{d} + \frac{p}{m} + \frac{n}{\acute{e}}} \geq 8\%$$

Ce ratio affine le précédent (ratio Cooke) en imposant aux établissements de crédit de détenir un niveau de fonds propres minimum d'avantage en adéquation avec l'ensemble des risques encourus.¹⁴

Outre le ratio de solvabilité, Bâle II recommande d'autres ratios à respecter qui sont :

Tier I : Ratio de fonds propres de base qui se traduit par la formule suivante :

$$T 1 = \frac{R_{c \acute{e}d} + r_{d m \acute{e}} + r_{o \acute{e}r}}{\frac{F}{d} + \frac{p}{m} + \frac{d b}{\acute{e}}} \geq 4\%$$



¹³ Idem
¹⁴ www.lafinancepourtous.com

Tier II : Ratio de fonds propres complémentaires est donné par la formule :

$$T2 = \frac{F_{dcéd} + R_{dm} + R_{oér}}{R_{ém}} \geq 4\%$$

Tier III : sont les fonds propres destinés à couvrir le risque de marché :

$$T3 = \frac{F_{dcéd} + R_{dm} + R_{oér}}{R_{ém}} \geq 1\%$$

Fonds propres « surcomplémentaires » spécifiquement dédiés à la couverture du risque de marché. Y figuraient les fonds propres Tier II excédentaires (le montant éligibles au titre du Tier II mais qui dépassent 100% des fonds propres admis en Tier 1) ainsi que des titres de dettes subordonnées ayant une échéance initiale au moins.

3.1 Calcul de risque crédit

Par ailleurs, les méthodes de calcul du risque crédit ont également été modifiées. En effet le dispositif de Bâle II propose deux grandes méthodes pour calculer les exigences de fonds propres relatives au risque crédit :

L'approche standardisée est identique à celui de 1988, mais elle est plus différenciée en fonction du risque. Elle est basée sur la pondération de chaque poste de bilan et de hors bilan par des coefficients reflétant la qualité de la signature de la contrepartie. La matrice de pondération prend en compte deux critères :

-] Un critère institutionnel : les coefficients de pondération sont fixés par grandes catégories d'emprunteurs (souverains, banque et entreprises) ;
-] Un critère de qualité intrinsèque de la contrepartie appréciée à travers la note qu'attribue une agence externe d'évaluation du crédit (agence de notation reconnue, banque centrale, organisme de garanties des crédits à l'exportation).

Cette méthode présente l'inconvénient pour les banques de ne pas fournir de notes pour certains clients (généralement, ceux qui n'empruntent pas sur le marché des capitaux).

L'approche basée sur la notation interne ou IRB (internal rating based) permet aux banques d'estimer par elles-mêmes la charge en capital, c'est-à-dire le montant de fonds propres nécessaires pour couvrir ce risque de crédit. Pour ce faire, la banque note ses clients à partir d'informations recueillies en interne, pour les classer ensuite en portefeuilles homogènes. Cette note dérivera du calcul de la perte attendue (perte potentielle), qui se calcul par le rapport suivant :

$$PP = PD * ECD * PCD$$

- La probabilité de défaut « Probability At Default » (PD) : c'est la probabilité qu'une contrepartie devienne défaillante dans un horizon temporel déterminé, un an en général. Elle est unique pour chaque emprunteur quelque soient les produits souscrits.

- L'exposition en cas de défaut « Exposure At Default » (ECD) : elle représente le montant des risques sur une contrepartie au moment de la défaillance. Pour un prêt il s'agit du capital restant dû à l'horizon correspondant à celui utilisé pour la PD et éventuellement des intérêts courus non échus au même moment.
- La perte en cas de défaut « Loss Given Default » (PCD) : elle correspond aux pertes constatées en cas de défaillance de la contrepartie.

Dans le cadre de la notation interne, Bâle II prévoit deux approches :

- ❖ L'approche IRB simple : dans cette approche, la banque n'estimera que la probabilité de défaut du preneur du crédit. Les autres facteurs seront fournis par les autorités de tutelle.
- ❖ L'approche IRB avancée : dans celle-ci, la banque estimera en plus de la probabilité de défaut la perte en cas de défaut, l'exposition au risque et la maturité restante.

Pilier II : Principe d'une surveillance Prudentielle individualisé

La fonction de supervision est réservée aux autorités nationales, donc le comité de Bâle a exigé de ces derniers de s'assurer que chaque banque met en place un contrôle interne qui permet de maintenir le capital réglementaire quelle que soit l'évolution de l'activité de la banque ; il s'agit de :

- S'assurer que tous les risques soient pris en charge ;
- Contrôler la conformité des méthodes internes d'évaluation et de gestion de ces risques.

3.1.1 Pilier III : Notion de transparence et de discipline de marché

Ce pilier vise à instaurer des règles de transparence financière en améliorant la communication d'informations au grand public sur les actifs, les risques et leur gestion.

L'objectif sous-jacent est d'uniformiser les pratiques bancaires en matière de communication financière et de faciliter ainsi la lecture des informations comptables et financières des banques d'un pays à l'autre.¹⁵

3.2 Insuffisances de Bâle II

La crise financière de 2007/2008 a mis en évidence le fait que les fonds propres des institutions financières étaient insuffisants ou de mauvaise qualité. Certains risques avaient été peu ou mal identifiés et ce, pour deux raisons :

- 1 la complexification des opérations réalisées sur les marchés financiers (produits structurés, titrisation) ;
- 2 la défaillance du contrôle interne et de la gouvernance des établissements bancaires (dans de nombreux cas) ;
- 3 les insuffisances du contrôle exercé par les régulateurs dans un univers où on faisait une confiance sans doute excessive à l'autorégulation.

¹⁵ www.lafinancepourtous.com

4 Les accords de Bâle III :

Après les événements récents, le comité de Bâle a publié dès juillet 2009 un ensemble de documents pour améliorer les trois piliers de Bâle et afin d'encadrer d'avantage l'activité des banques. Ces mesures sont entrées définitivement en application au début de l'année 2013.

Les grandes mesures de Bâle III sont les suivantes :

4.1 Renforcement des fonds propres

- Le noyau dur de Tier1 devient de 4.5% et le noyau supplémentaire devient 1.5% ;
- Constitution d'un volant de conservation de fonds propres égal à :

$$\frac{F \quad p \quad d \quad b}{\sum R.C + R.O + R.M}$$

Le rôle de VCS est de faire en sorte que les banques constituent en dehors des périodes de tensions, des marges de fonds propres qu'elles peuvent mobiliser lorsqu'elles enregistrent des pertes.

- Le Tier II devient 2%.
- Le troisième type de capital a été éliminé (Tier III).
- En plus des modifications apportées au ratio de solvabilité, deux coussins de sécurité ont été introduits :

$$R \quad c_1 \quad - \quad c \quad = \quad \frac{F \quad p \quad d \quad b}{\sum R.C + R.O + R.M} \leq 2.5\%$$

Le rôle de VCC est de faire en sorte que les exigences de fonds propres tiennent compte de l'environnement macro financier, dans lequel les banques évoluent. Il vise à doter le secteur bancaire d'un volant de fonds propres lui permettant de se protéger contre des pertes potentielles futures.¹⁶

- Et un coussin de risque systémique : c'est le risque de matérialisation d'une rupture dans le fonctionnement des services financiers susceptible d'affecter l'ensemble du secteur ainsi que l'économie réelle. ce coussin pourra atteindre les 5%.¹⁷

Tableau 2 : Synthèse d'exigence en fonds propres de Bâle 3

<i>Ratio</i>	<i>Valeur</i>	<i>Total</i>
Core Tier I	4,5%	8,5%
Core supplémentaires	1,5%	
Ratio de sécurité	2,5%	
Tier II	2%	2%
Ratio systémique	0% à 5%	Coussins = 0% à 7.5%
Ratio contra-cyclique	0% à 2.5 %	
Ratio minimale de solvabilité		10.5% à 18%

¹⁶La réglementation prudentielle, Cours de Mr. Ilmane, Ecole supérieure de commerce.

¹⁷ Idem

Source :Elaboré par nous- même

4.2 Instauration de ratio de liquidité :

4.2.1 Le liquidity coverage ratio (LCR)¹⁸

- Le LCR est un ratio qui permet de mesurer la capacité d'un établissement à faire face à un choc de liquidité inattendu qui se traduirait par une augmentation soudaine des flux de paiements sortants dans un horizon de 30 jours.
- Afin d'amortir ces chocs, l'établissement doit disposer de suffisamment d'actifs liquides de haute qualité¹⁹ (High quality liquid assets - HQLA) pouvant être mobilisés en cas d'urgence nécessaire pour faire face aux engagements à très court terme.

$$L = \frac{E}{T} \frac{d a l}{d s} \frac{d h q}{n s} \frac{é}{l 3 j} \geq 1 \%$$

4.2.2 Le Net Stable Funding Ratio (NSFR)

- Le NSFR est un ratio qui permet de mesurer la qualité de la transformation exercée par l'établissement assujetti en réduisant l'asymétrie des échéances entre les actifs et les passifs à un horizon d'un an.
- Pour se faire, les besoins de financements stables (Required Stable Funding, ASF) de l'établissement doivent être entièrement et à tout moment couverts par ses ressources stables (Available Stable Funding, ASF). A l'inverse du LCR qui ne prend en compte que certains éléments du bilan, le NSFR a une approche plus globale en termes de prises en compte d'éléments de bilan.

$$N = \frac{R}{B} \frac{s}{d f} \frac{s}{s} \geq 1 \%$$

4.2.3 Mise en place d'un ratio effet de levier :

Une des caractéristiques de la crise a été l'accumulation dans le système bancaire d'un effet de levier excessif, qui correspond à l'utilisation de l'endettement pour doper la capacité d'investissement d'un organisme.

Nouvel élément de Bâle 3 est l'instauration d'un ratio de levier. Il est défini par le rapport des fonds propres de base (Tier1) sur le total de bilan, c'est-à-dire, total actif ou passif. Il a été convenu de tester un ratio de levier minimal de 3 %.

Ce ratio va permettre de limiter la part des actifs de la banque qui est financée par la dette.

¹⁸ Des actifs sont considérés comme actifs liquides de haute qualité s'ils peuvent être facilement et immédiatement transformés en liquidité sans perdre (ou en perdant très peu) de leur valeur.

Section 3 : Les outils de gestion de risque crédit

Pour se protéger contre les pertes et le risque de crédit, les banques ont mis en place des moyens de prévention. Le risque pris lors de l'octroi d'un crédit doit être acceptable pour la banque. C'est pour cela que les établissements de crédit mettent en pratique des mesures lors de la mise en œuvre du prêt ou de son suivi.

Nous présenterons dans cette section les méthodes les plus couramment utilisées.

1 Les garanties

Pour améliorer la sécurité de ses engagements et surtout pour se couvrir du risque de non remboursement, il faut que le banquier recueille des garanties.

La garantie est en quelque sorte un besoin de sécurité que la banque doit prendre. Elle sert à la protéger contre le risque de défaillance de son client et à atténuer des provisions à constituer.²⁰ Cette garantie peut être de nature très variée, un bien ou une personne.

On distingue deux sortes de garanties : les garanties réelles et les garanties personnelles.

1.2 Les garanties réelles

Les garanties sont dites réelles lorsqu'elles portent sur un bien mobilier²¹ ou immobilier que le débiteur s'engage à mettre à la disposition de sa banque.

On distingue trois types de garanties réelles : l'hypothèque, le nantissement et le gage.

1.1.1 L'hypothèque

Elle est définie par l'article 882 du code de commerce, comme suit :

« ...le contrat par lequel le créancier acquiert sur un immeuble affecté au paiement de sa créance, un droit réel qui lui permet de rembourser, par préférence aux créanciers inférieure en rang, sur le prix de cet immeuble en quelques mains qu'il passe ... »²²

L'hypothèque porte sur des biens immeubles que détient le débiteur et les met en garanties pour pouvoir bénéficier du crédit de la banque.

L'hypothèque peut être sous trois formes :

- Légal : elle est prévue par la loi au profit de certains créanciers privilégiés
- Conventionnelle : elle est inscrite de la signature d'un contrat. Elle doit obligatoirement être constatée par écrit sous une forme authentique (acte notarié) et publié au livre foncier au niveau de la conservation foncière.
- Judiciaire : elle résulte d'un jugement, car la banque peut procéder à l'inscription d'hypothèque sur les immeubles du débiteur par décision de justice et ce après une poursuite judiciaire.

²⁰ Mr. Azzaoui Khaled, Le financement des PME et le mécanisme de garantie en Algérie, thèse de doctorat, 2016/2017, ESC.

²¹ Les biens mobiliers ou meubles constituent une catégorie de biens caractérisés par le fait qu'ils peuvent être déplacés, ils sont divisés en deux catégories juridiques distinctes, «meubles corporels » Et « meubles incorporels. », ils s'opposent aux biens immobiliers ou immeubles qui ne peuvent pas être déplacés comme un terrain

²² L'article 409 du code de commerce.

1.1.2 Le nantissement

« le nantissement est un contrat par lequel une personne s'oblige ,pour la garantie de sa dette ou celle d'un tiers ,à remettre au créancier ou à tiers personne choisie par les parties ,un objet sur lequel elle constitue ,au profit du créancier ,un droit réel en vertu duquel celui-ci peut retenir l'objet jusqu'au paiement de sa créance et peut se faire payer sur le prix de cet objet en quelques mains qu'il passe ,par préférence aux créanciers chirographaires et aux créanciers inférieurs en rang »²³

Le terme nantissement désigne la situation juridique où un propriétaire affecte sans en perdre la possession, des biens incorporels (fonds de commerce, fonds agricoles, valeurs mobilières ...) à la garantie d'une ou plusieurs dettes dont il est redevable.

1.1.3 Le gage

Un contrat qui accorde à la banque le droit de se faire payer par préférence aux autres créanciers sur un bien mobilier en cas de défaillance d'un emprunt à condition qu'in s'agisse de bien corporels.

On distingue deux types de gage :

Avec dépossession	Sans dépossession
Le créancier doit veiller à la bonne conservation du gage	Le constituant ne peut pas vendre le bien gagé.
Il peut se faire rembourser les dépenses nécessaires à cette conservation.	Il peut percevoir les fruits du bien gagé.
Le créancier peut percevoir les fruits du bien gagé.	Le constituant doit veiller à la bonne conservation de bien gagé. En cas de choses fongibles, elles peuvent être remplacées.
Il doit imputer ces fruits sur les intérêts ou sur le capital de la dette.	Il doit remettre le bien gagé au créancier s'il ne peut honorer sa dette.
Le créancier peut retenir le bien jusqu'au paiement de la créance.	

1.2 Les garanties personnelles

Elles désignent l'engagement d'une personne ou plusieurs à rembourser un créancier en cas de défaut de paiement de l'emprunteur.

On distingue deux types de garanties personnelles : le cautionnement et l'aval

1.2.1 Le cautionnement

L'article 644 du code civil algérien définit le cautionnement comme « un contrat par lequel une personne garantit l'exécution d'une obligation, en s'engageant, envers le créancier, à satisfaire à cette obligation, si le débiteur n'y satisfait pas lui-même. »

On distingue deux types de cautionnement :²⁴

²³ L'article 948 du code civil.

²⁴ Mr. Azzaoui Khaled, Le financement des PME et le mécanisme de garantie en Algérie, thèse de doctorat ,2016/2017, ESC.

- Le cautionnement simple : lors d'un cautionnement simple, la caution peut bénéficier de deux allègements :
 - Le bénéfice de discussion : il donne à la caution poursuivie par le créancier d'éviter la poursuite et lui demander, sans aucune obligation de discuter d'abord les biens du débiteur principal
 - Le bénéfice de division : en cas de pluralité des cautions, le créancier ne peut poursuivre chacune d'elle que pour sa part dans le cautionnement, autrement dit, la dette est subdivisée entre les cautions.
- Le cautionnement solidaire : lors d'un cautionnement solidaire, il n'y a ni bénéfice de discussion, ni bénéfice solidaire. Le créancier au choix de poursuivre indifféremment le débiteur ou la caution.

1.2.2 L'aval²⁵

L'aval est une forme de cautionnement qui s'applique uniquement aux lettres de change, billets à ordre et chèques.

L'aval est une garantie qui a pour but d'assurer la bonne fin du paiement de la lettre de change, du billet à ordre ou du chèque, pour tout ou partie de son montant, cette sureté personnelle est une garantie qui ne peut être donnée que par une tierce personne autre que le tiré.

2 Les garanties financières (sociétés de contre garanties)

Les PME n'ont généralement pas de suretés suffisantes à offrir à leurs banques pour pouvoir bénéficier d'un financement et pour se faire elles doivent avoir des garanties complémentaires. Donc c'est dans ce contexte que les pouvoirs publics ont créé le FGAR et CGCI.

FGAR ²⁶ : le FGAR a été créé par le décret N°02-373 du 11/11/2002 en application de la loi d'orientation 01//18 pour la promotion de la PME. Il est placé sous la tutelle du ministre de la PME et de l'artisanat. Son objectif principal est de faciliter l'accès au financement bancaire à moyen terme, afin de supporter le démarrage et l'extension des PME, en accordant des garanties de crédits aux banques commerciales.

CGCI ²⁷ : c'est une société créée à l'initiative des pouvoirs publics par décret présidentiel 04-134 du 19 Avril 2004 portant statuts de la caisse pour soutenir la création et le développement de la PME en lui facilitant l'accès au crédit. Elle apporte aux banques sa garantie financière qui les met à l'abri du risque de non remboursement et aux investisseurs des conditions d'emprunt favorables.

3 Les assurances

L'assurance crédit est une assurance qui garantit la prise en charge de tout ou partie des échéances de remboursement ou du capital restant dû d'un crédit, en cas de survenance de certains événements, le

²⁵ Idem

²⁶ Fonds de garantie des crédits aux PME

²⁷ Caisse de garantie des crédits d'investissement.

plus souvent le décès ,la perte totale ou irréversible d'autonomie ,incapacité temporaire de travail et la perte d'emploi.²⁸

L'assurance « décès-IAD-décès-invalidité absolue et définitive. » :

Cette assurance offre à la banque la garantie de payer le capital restant dû si l'emprunteur décède ou devient invalide de façon définitive et permanente.

Selon la réglementation,un assuré en IAD s'il est dans l'impossibilité absolue et irréversible de se livrer à une occupation ou un travail quelconque pouvant lui procurer un gain et un profit et qu'il est dans l'obligation de faire appel à l'assistance d'autrui pour accomplir tous les actes habituels du quotidien.

L'assurance « perte d'emploi /insolvabilité » :

L'assurance perte d'emploi est une assurance qui prend en charge le remboursement du prêt si l'emprunteur perd son emploi .L'assurance perte d'emploi est toute fois assez restrictive :

- Elle n'est proposée qu'aux emprunteurs salariés en contrat à durée indéterminée ;
- Elle n'entre pas en application si l'emprunteur démissionne ou est licencié pour faute grave ;
- Elle ne couvre souvent qu'une partie du prêt.

Assurance « CAT-NAT-catastrophe naturelles » :

La banque exige cette assurance dans le cas de crédit immobiliers .Quatre événements sont pris en charge par cette assurance : les tremblements de terre, les inondations et coulées de boue, les tempêtes et vents violents, les mouvements de terrain.

4 Stratégie de provisionnement du risque crédit

Comme nous l'avons vu, le fait de prendre des risques est inhérent à l'activité de banquier, notamment le risque de crédit qui lui est le plus nocif

Si exiger des garanties et des polices d'assurances est un moyen de s'en prévenir, cela ne suffit pas. Afin de maîtriser encore plus ce risque, les banques sont dans l'obligation d'affecter une provision pour chacune de leurs créances afin de couvrir les pertes potentielles.

La provision est définie par la doctrine comme étant la fraction de bénéfice que l'entreprise met de coté en vue de faire face ultérieurement, soit à une dépréciation d'un élément de l'actif, soit à une perte ou charge, dont l'objet est nettement précisé et qui n'est pas encore effective à la clôture de l'exercice mais que des événement en cours rendent probable.²⁹

4.1 Le provisionnement des pertes subies (ex-post)

4.1.1 . Présentation de provisionnement ex-post

C'est une évaluation effectuée à la fin ou après la fin d'une intervention publique. Elle vise à tirer des enseignements rétrospectifs sur une politique parvenue à maturité.³⁰

4.1.2 Provisionnement ex-post selon les normes algériennes

²⁸Revue de l'assurance N°24, janvier à mars 2019, éditée par le conseil national des assurances.

²⁹<http://www.dge.gov.dz>

³⁰ Dictionnaire Larousse

Selon le règlement N°14-03 du 16 février 2014 relatif aux classements et provisionnement des créances et des engagements par signature des banques et établissements financiers, les créances sont classées en créances courantes et créances classées et c'est selon leurs classements qu'on les approvisionne. Les classements et leurs provisionnements respectifs sont présentés dans le tableau suivant :

Tableau 3 : Classement et provisionnement des créances en ex-post

Classement		Nombre de retard	Taux de provisionnement
Créances courantes C 0		aucun	1% à 3%
Créances classées	C 1	90 à 180 jours	30 %
	C 2	180 à 360 jours	50 %
	C 3	Plus de 360 jours	100 %

Source :Elaboré par nous même

Les créances courantes sont des créances dont le recouvrement intégral dans les délais contractuels apparaît assuré.

Créance de catégorie 1 (C1) : créance à problème potentiel.

Créance de catégorie 2 (C2) : créances très risquées

Créances de catégorie 3(C3) : créances compromises, c'est celles dont le recouvrement partiel ou total est compromis.

Remarque :

Le provisionnement se fait à chaque trimestre. En clôturant son bilan, le banquier doit calculer les provisions nécessaires pour l'année suivante.

Le provisionnement s'effectue sur le montant brut, hors intérêts non recouverts.

4.2 Présentation de provisionnement ex-ante

Suite à la crise des Subprimes de 2007-2008 qui a mis en évidence cette reconnaissance tardive des pertes de crédit de la part des banques et leurs caractères pro cycliques, les dirigeants du G20, les investisseurs, les organes de régulation et les autorités prudentielles ont appelées les instances de normalisation comptable à prendre des mesures pour améliorer les règles et les pratiques en matières de provisions pour pertes sur prêts. En réaction, l'International Accounting Standards Board (IASB)³¹ a publié en 2014 le document IFRS³² 9 Financial Instruments, qui comprend une nouvelle norme sur

³¹Est l'organisme international chargé de l'élaboration des normes comptables internationales IAS/IFRS.

³²L'international Financial Reporting Standards (IFRS): une série de normes comptables internationales. Elles définissent la manière dont certains types de transactions et autres événements doivent apparaître dans les états financiers. Les IFRS sont émis par le bureau des standards comptables internationaux.

les provisions pour pertes fondées sur les « pertes de crédit attendues. ». De son côté, l'US Financial Accounting Standards Board (FASB)³³ a publié sa norme finale de provisionnement fondée sur les « pertes de crédit actuellement attendues. » en 2016. Ces nouvelles normes entreront en vigueur entre 2018 et 2021.³⁴

Le provisionnement ex-ante se fait par une analyse des faits économiques effectués avant qu'ils ne soient produits, résultant des projets et des anticipations des agents économiques, par opposition à ex-post.³⁵

4.2.1 Méthode de provisionnement ex-ante

Elle prévoit la classification des instruments en trois « Bucket. », avec des règles de calcul spécifiques à chacun d'eux :

Le Bucket 1 : abrite les créances n'ayant pas enregistré de détérioration significative de la qualité du crédit depuis leur comptabilisation initiale. Cette catégorie fait l'objet d'une provision correspondant à la perte attendue sur 12 mois.

Le Bucket 2 : contient les créances ayant enregistré une détérioration qualifiée de sensible. Cette catégorie fait l'objet d'une provision correspondant à la perte attendue sur la durée restante jusqu'à la maturité.

Le Bucket 3 : contient les créances en souffrance. Cette catégorie fait l'objet, comme celle du Bucket 2, d'une provision correspondant à la perte attendue sur la durée restante jusqu'à la maturité.³⁶

Remarques

-) Ce modèle a pour objectif de mieux refléter la détérioration (ou l'amélioration) du risque de crédit au niveau des provisions, et ce durant toute la vie du contrat.
-) Au fil du temps et de la survenance d'indicateurs généraux (la conjoncture économique) et spécifiques (l'évolution de la valeur des garanties), un crédit peut se dégrader et migrer vers les Bucket suivants. À ce titre, l'IASB propose une liste d'indicateurs à prendre en compte pour évaluer en fin d'année, si le crédit reste dans son Bucket ou migre vers le Bucket supérieur.³⁷
-) Dans un système de provisionnement dynamique, les banques sont tenues de fournir aux autorités de surveillance des informations sur leurs méthodes statistiques avant que leurs provisions ne soient prises en compte.

5 Les dérivés crédit : instrument de transfert du risque

Les produits dérivés (derivative product) sont des instruments dont les caractéristiques découlent (dérivent) d'autres instruments, appelés « sous-jacents. » (Underlying). Il peut se définir de façon très générale comme un contrat financier négociable portant droit sur un actif.³⁸

³³ Est une association à but non lucratif, non gouvernementale, dont le but est de développer les principes comptables des États-Unis dans l'intérêt du public.

³⁴ Benjamin H. Cohen et Gerald A. Edwards JR, la nouvelle ère des provisions pour pertes de crédit attendues, rapport trimestriel BRI, mars 2017.

³⁵ Dictionnaire Larousse

³⁶ Lotfi Said et Bensaida Salma, Measuring the impact of IFRS 9 on Bank credit risk, revue du contrôle de la comptabilité et de l'audit, Maroc, Juin 2018.

³⁷ Idem

³⁸ Les produits dérivés, Cours Mr. BENIALLES, , ESC.

5.1 Les dérivés de crédit

Un dérivé de crédit est un produit dérivé dont le sous-jacent est un actif de type crédit, c'est-à-dire une créance ou un titre représentatif d'une créance (obligation). Son but est de transférer les risques relatifs au crédit, sans transférer l'actif lui-même.

5.2 Paramètre du dérivé de crédit

Quel que soit le crédit, il se caractérise par :

- le sous-jacent : le prêt que l'établissement cherche à couvrir ;
- L'acheteur de la protection ou le vendeur du risque : celui qui utilise le dérivé comme un instrument de couverture du risque de crédit ;
- L'entité de référence ou le débiteur contre lequel l'acheteur désire se couvrir ;
- Le vendeur de la protection ou l'acheteur du risque : l'établissement qui s'engage à verser un certain montant en cas de défaut du débiteur ;
- Le « credit event » ou l'événement de crédit qui correspond à la réalisation du risque de crédit, il peut prendre différentes formes qui sont prédéterminées à l'avance dans un contrat comme la défaillance de l'entité de référence, un défaut de paiement, la faillite, etc.
- La prime ou aussi appelée Spread du swap est exprimée en points de base ³⁹du nominal, c'est le montant payé par l'acheteur de protection au vendeur.

Le crédit default swap ou swap sur défaillance :

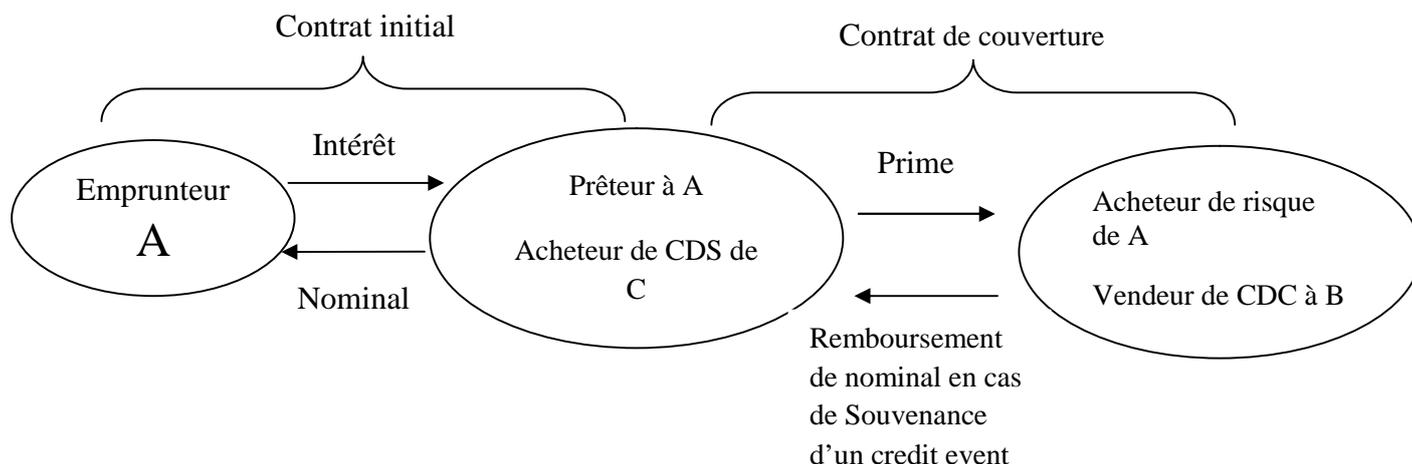
Il existe plusieurs dérivés de crédit, mais les crédits default swap(CDS) constituent le plus grand segment du marché mondial des produits dérivés de crédit.

« Il s'agit d'un contrat par lequel l'acheteur de la protection (A) verse une prime régulière (trimestrielle dans la majorité des contrats) au vendeur (C) lequel s'engage en cas de survenance d'un événement de crédit sur l'entité de référence (B) de la dédommager de la perte subie sur l'actif sous-jacent. »⁴⁰

Le fonctionnement d'un crédit default swap est schématisé comme suit :

³⁹ Un point de base est une unité de mesure fréquemment employée par les traders. Elle permet de quantifier les mouvements au niveau de taux de change ou autres pourcentage. Un point de base est l'équivalent de 0.01%.

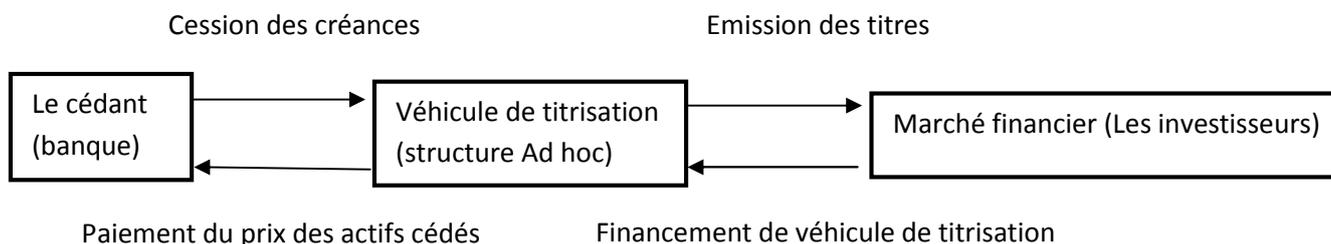
⁴⁰REY .N, Les dérivés de crédit : Instrument de couverture et facteurs d'instabilité. L'exemple des « crédits default swap. », Atelier dans le cadre de l'axe thématique « intermédiation et systèmes financiers. » du GDRE n°335, monnaie, banque et finance du CNRS, novembre 2007, université Paris 12.



6 La titrisation :

La titrisation ou la « securitization. » selon la terminologie anglo-saxonne, est une technique financière qui consiste classiquement à transférer à des investisseurs des actifs financiers tel que des créances (des prêts en cours) en transformant ces créances par le passage à travers d'une société ad-hoc en titres financiers émis sur le marché des capitaux. Elle consiste de ce fait en une innovation majeure du système financier international, tant par le volume mis en jeu que par la complexité des techniques financières développées à cet effet.⁴¹

Une opération de titrisation est schématisée comme suit :



La titrisation permet donc à la banque de « diminuer ses actifs à risque mais, à la différence des dérivés de crédit, elle opère une sortie de la créance de l'actif du bilan. »⁴²

En effet lorsque la banque conserve un crédit à l'actif de son bilan jusqu'à l'échéance, elle supporte seule le risque lié au défaut de l'emprunteur. Mais en ayant recours à la titrisation, ce risque de perte est transféré aux investisseurs, ce qui signifie que si le défaut se concrétise, c'est l'investisseur qui subira, le cas échéant, une perte financière.

⁴¹ Michèle Ceresoli et Michel Guillaud, « TITRISATION : gestion financière de la banque. », édition ESKA, paris.

⁴² Cous Sergues et d'autres, Gestion de la banque: Normes et réglementation à jour, nouvelles stratégies bancaires 8ème édition, Edition Dunod, 2017.

Conclusion :

Lorsque le débiteur ne répond pas à son obligation initiale qui est le remboursement de sa dette, la banque subit un risque de crédit.

Le risque de crédit est un élément indissociable de l'activité bancaire. Toutefois la réalisation de ce risque a toujours pour impact la baisse des résultats, à l'extrême, la réduction drastique des fonds propres, voire même la faillite de la banque.

Pour se protéger contre les pertes que peut engendrer le risque de crédit, les autorités de réglementation bancaire ont imposé plusieurs normes aux banques pour atténuer ses conséquences. Ces normes se basent principalement sur le capital comme instrument de régulation du système bancaire. En effet, tout établissement doit respecter un seuil minimum de fonds propres, qui servira à couvrir les pertes les plus improbables.

Le risque pris lors de l'octroi de crédit doit être acceptable pour la banque. C'est pour cela que les établissements de crédit mettent en pratique des mesures très différentes selon la politique de l'établissement lors de la mise en œuvre du prêt ou de son suivi. On distingue des méthodes classiques tels que les garanties, les provisions, l'assurance, et d'autres plus sophistiqués tel que le recours à la titrisation et les produits dérivés.

Chapitre II :

Modèle de prévision de la défaillance : Le crédit Scoring

Introduction

« Faire crédit, c'est faire confiance ; c'est donner librement la disposition effective et immédiate d'un bilan réel ou d'un pouvoir d'achat, contre la promesse que le même bien, ou un bien équivalent, vous sera restitué dans un certain délai, le plus souvent avec rémunération du service rendu et du danger couru, danger de perte partielle ou totale que comporte la nature même de ce service »⁴³

Comme nous l'avons vu dans le chapitre précédent plusieurs risques affectent les organismes de crédits mais le risque de contrepartie reste la principale préoccupation de ces derniers.

Afin d'éviter de réduire ce risque, les banques font recours à des méthodes classiques dont la principale est l'évaluation de diagnostic financier de l'entreprise, mais ce dernier nécessite beaucoup de temps et un personnel compétant. C'est pourquoi les banques ont cherché des méthodes plus pertinentes qui leur permettront de distinguer les bons clients des clients risqués et de prendre les décisions en adéquation avec leur évaluation.

Il existe plusieurs méthodes de modélisation du risque de crédit, et chacune à ses propres caractéristiques qui la diffère de l'autre. Parmi ces méthodes, on peut citer : L'analyse financière, la VAR (Value at Risk), la notation financière, les systèmes experts, les réseaux de neurones et la méthode des scores.

La méthode des scores est considérée comme l'un des outils de gestion de risque crédit qui a permis aux analystes de se prévenir d'un mauvais client en utilisant une approche probabiliste.

Dans ses accords de 2004, le comité de Bâle a souligné l'importance d'introduire le crédit scoring dans les systèmes de notation interne comme véritable technique de mesure du risque de crédit, car l'estimation des probabilités de défaut est une étape nécessaire à la construction de ces modèles de notation.

Dans le présent chapitre, on présentera d'une façon générale le crédit scoring, avec un aperçu de son développement, aussi la méthodologie d'élaboration d'un modèle de score. Et en dernier, nous exposons des outils statistiques utiles pour l'élaboration d'un meilleur modèle, ainsi qu'un état des avantages et faiblesses découlant de l'utilisation d'un tel modèle.

Chapitre II :Modèle de prévision de la défaillance : Le crédit Scoring

⁴³ Dutailis, G, Le risque de crédit bancaire, Edition scientifique Riber, p.19.

Section 1 : Présentation de crédit scoring

Le crédit scoring est une méthode importante pour les institutions financières, son but est d'arriver à distinguer entre les bons demandeurs de crédit et les mauvais, afin de prendre la décision de leurs octroyer ou non l'emprunt.

1 Définition de crédit scoring

Anderson(2007) a suggéré que pour définir le crédit scoring, le terme devrait être divisé en deux composantes : le crédit et le scoring.

)] Le mot crédit comme nous l'avons vu au premier chapitre désigne le fait d'acheter maintenant et de payer plus tard. Il est dérivé de mot latin «credo» qui signifie «je crois » ou «je fais confiance en ».

)] Le mot « scoring » fait référence à « l'utilisation d'un outil numérique pour classer les cas selon une qualité réelle ou perçue afin de les discriminer et garantir des décisions objectives et cohérentes ».

Par conséquent, les scores peuvent être présentés comme « nombres » pour représenter une seule qualité, ou « notes » qui peuvent être présentées sous forme de « lettres » ou « étiquettes » pour présenter une ou plusieurs qualités.

Le crédit scoring peut être simplement défini comme « l'utilisation de modèle statistique pour transformer les données pertinentes en des mesures numériques qui aide à prendre les décisions d'octroi de crédit »⁴⁴

« La notation du crédit consiste à utiliser des modèles statistiques pour déterminer la probabilité qu'un emprunteur potentiel fasse défaut sur un prêt. Les modèles de notation de crédit sont largement utilisés pour évaluer les prêts aux entreprises, à l'immobilier et à la consommation ».⁴⁵

« Le crédit-scoring est une méthode statistique utilisée pour prédire la probabilité qu'un demandeur de prêt ou un débiteur existant fasse défaut » (traduit de MESTER, 1997).

2 Historique de crédit scoring

Les techniques statistiques utilisées pour la notation du crédit reposent sur l'idée de discrimination entre plusieurs groupes dans un échantillon de données.

Ces procédures datent des années trente et quarante du siècle précédent (Fisher, 1936; Durand, 1941). A cette époque, certaines maisons de financement et sociétés de vente par correspondance éprouvaient des difficultés avec leur gestion du crédit.

La décision d'accorder des prêts ou d'envoyer des marchandises aux candidats a été prise avec jugement par les analystes de crédit. La procédure de décision était non uniforme et subjective, elle dépendait des règles de chaque maison financière et de la connaissance personnelle et empirique de chaque commis.⁴⁶

⁴⁵ Gup.B. E & Kolari.J.W, Commercial banking: The management of risk.

⁴⁶ Karel Komorad, "Credit Scoring Estimation", these de masters, Institute for Statistics and Econometrics Humboldt University Spandauer Str.1D-10178 Berlin.

Le premier cabinet de conseil a été créé à San Francisco par Bill Fair et Earl Isaac à la fin des années 50. Leur système s'est rapidement répandu lorsque les institutions financières ont découvert que l'utilisation de la notation de crédit était moins chère, plus rapide, plus objective et surtout bien meilleure prédictive que tout système de jugement.⁴⁷

La recherche dans le domaine de la notation du crédit a commencée dans les années 30. Après cette date, de nombreux travaux et méthodes différentes sont entrés dans la littérature de la notation des crédits.

Selon le type de méthodes, on peut scinder la période 1930-2005 en 4 sous-périodes :

- Période 1 : L'âge primitif de la notation de crédit : durant cette période, la recherche était basée uniquement sur une analyse de ratios. Au cours de ces années, les scientifiques ont comparés les ratios d'entreprises en défaut et non en défaut et ont tentés de se faire une idée des performances financières des entreprises. Comme on peut le deviner, ces types de méthodes n'avaient pas un pouvoir prédictif élevé et ne convenaient donc pas très bien.
- Période 2 : Application de l'analyse discriminante : grâce à cette application la recherche a acquis un pouvoir prédictif, mais cette méthode repose sur des hypothèses très fortes sur les variables et donc, le pouvoir de prédiction n'est pas très élevé. L'application de l'analyse discriminante est un tournant pour la notation de crédit, car elle a ouvert la porte aux méthodes statistiques.⁴⁸

Période 3 : L'âge de la régression : Après les années 70, les méthodes qui s'appliquaient à ce domaine ont changés rapidement. Les principaux types de méthodes étaient les approches basées sur la régression. La régression linéaire a été appliquée en premier, mais elle n'a pas donnée de bons résultats. Parce que les probabilités de défaut de crédit ne prennent que des valeurs comprises entre 0 et 1, mais la régression linéaire peut donner les résultats entre $-$ et $+$. Ensuite, deuxièmement, la régression probit est entrée en jeu. Puisqu'elle a également des hypothèses fortes de normalité, la fin de l'application de ces méthodes est rapide. En d'autres termes, dans la période 1970-1980, les méthodes de type régression n'étaient pas passées au premier plan de l'analyse discriminante.

Dans les années 1980, l'étude de la régression logistique a accrue l'intérêt pour la régression puisqu'elle n'a pas d'hypothèses de normalité sur les variables, permet de prédire, et d'interpréter des coefficients, et elle donne la sortie sur l'intervalle [0,1]. Bien qu'après les années 1980, de nombreuses autres méthodes statistiques aient également été appliquées, telles que le voisinage le plus proche, les arbres de classification et de régression, l'analyse de survie, etc., cette méthode a conservé son importance même de nos jours en tant que technique statistique la plus utilisée dans la recherche.

- Période 4 : L'âge de la machine : L'année 1990 est un autre tournant pour la notation du crédit. Cette année, les méthodes statistiques ont cédé leur place aux méthodes de type machine Learning avec les réseaux applicatifs.

⁴⁷ Idem

3 Avantages et inconvénients de crédit scoring⁴⁹

3.1 Les avantages du crédit scoring

3.1.1 Atouts spécifiques à l'outil

La simplicité: La note de score, qui est obtenue en général à partir de 6 à 12 informations, est calculable en très peu de temps. Cette rapidité de décision présente un double avantage :

-)] Un avantage interne de charge de travail car la tâche de l'exploitant et le processus administratif de décision sont considérablement accélérés.
-)] Un avantage commercial car le client obtient une réponse en quelques minutes. C'est un énorme progrès par rapport à certaines demandes de prêts qui constituent un véritable interrogatoire, parfois mal acceptés par le client.

L'homogénéité: Avec la décision traditionnelle, un client refusé aujourd'hui par l'exploitant pourrait être accepté demain ; de même un client accepté dans une agence aurait pu être refusé dans une autre. Il est difficile dans un tel contexte de définir une politique de crédit homogène. Par contre, insensiblement aux humeurs quotidiennes, le crédit scoring donne invariablement la même décision permettant une politique de sélection des homogènes dans le temps et dans l'espace. De plus, le crédit scoring permet de s'affranchir des difficultés causées par l'indisponibilité momentanée ou définitive (turn-over) du personnel spécialiste du risque.

La souplesse : Il est difficile de prendre des mesures de libération ou de restriction dans la distribution des crédits sans disposer de règles claires et de références chiffrées. En jouant sur les seuils de sélection du score, un établissement peut, en quelques jours, modifier sa politique de sélection des risques tout en mesurant avec exactitude les conséquences en taux de contentieux ou en taux de refus de ces modifications.

3.1.2 Atouts spécifiques à l'établissement : Un établissement peut choisir de mettre en place un crédit scoring pour différentes raisons :

La politique de cautionnement : Le cautionnement, qui est parfois une nécessité face au risque, est un procédé soit coûteux, soit anti-commercial, soit les deux. Le score permet d'accepter sans caution les dossiers sans problèmes et de ne faire intervenir une caution que sur les dossiers tangents, réduisant ainsi considérablement les avantages du cautionnement.

⁴⁹<https://detectionfaillite.fr>

La délégation des décisions : Pour la plupart des dossiers, le score autorise la délégation des décisions à un personnel beaucoup moins qualifié, et moins coûteux, que le personnel capable de mener à terme le processus traditionnel de décision. Les étapes administratives sont très raccourcies. C'est une forme de productivité.

La diminution des impayés : Sur l'analyse statistique et objective des critères de risques, le crédit scoring se révèle, dans bien des cas, d'une efficacité supérieure au processus de décision traditionnel ; cela ne signifie pas que le crédit scoring soit plus efficace qu'un bon expert ; mais les bons spécialistes étant parfois relativement peu nombreux, ils ne peuvent participer à toutes les décisions ou y consacrer le temps nécessaire. Ainsi les décisions sont parfois prises trop vite ou par un personnel peu expérimenté et se révèlent globalement moins justes que les conclusions d'un crédit scoring.

Le contrôle et la prévision des impayés : Chaque client accepté a été noté par le score ; la probabilité d'impayés ou de contentieux de chaque client est connue. En cumulant ces probabilités sur l'ensemble des encours, la masse totale des impayés ou des contentieux peut être estimée et, dans certains cas (si les fichiers de prêt de l'établissement le permettent), répartie dans le temps suivant un échéancier.

3.2 Les inconvénients du credit scoring

Les techniques du credit scoring présentent un certain nombre de faiblesses. En effet, l'avenir peut être différent du passé, alors que les prédictions fondées sur la notation s'appuient uniquement sur les événements passés. Les changements pouvant survenir sont susceptibles d'affecter la qualité de prédiction du modèle.

3.2.1 Les problèmes pratiques

- **Coût onéreux de mise en place :** La collecte et le traitement des données nécessaires au calcul d'un score entraînent des surcoûts qui peuvent être importants pour les institutions financières les moins pourvues en moyens humains et matériels. Cette charge provient des coûts engendrés par la collecte d'un grand nombre de données de bonne qualité, des coûts de traitement automatisé, déstockage et de la maintenance du système d'informations nécessaires, et enfin des coûts relatifs à la formation des responsables de prêts, aux techniques de collecte des données supplémentaires, qui s'avèrent nécessaires. À ce titre le projet de notation induit, lui-même, un coût important lié au développement du score, à l'intégration au système d'information, à l'ajustement du système d'information, à la formation des utilisateurs finaux et au suivi. En particulier, l'ajustement du système d'information, pour le calcul et

l'édition automatique des prévisions du risque, peut s'avérer extrêmement longue et difficile et peut représenter une part considérable du budget prévu.

- **Complexité de mise en place :** La mise en place d'un système de scoring est souvent techniquement complexe. L'évaluation statistique nécessite du personnel spécialiste et expérimenté. De plus, ce personnel doit suivre des formations concernant ces nouvelles techniques pour se familiariser avec les résultats et pouvoir les interpréter.
- **L'accès aux données pertinentes :** La procédure d'archivage des dossiers acceptés et refusés est rarement informatisée, souvent les dossiers refusés, jugés inutiles, ne sont même pas conservés. Il est indispensable de prévoir une phase d'inventaire, de codification et de saisie de tous les dossiers. Il faut aussi mettre en place une phase de mise à jour des données pour l'alimentation de la base de données qui pourra servir pour des futures études de scoring.
- **Aspect réglementaire :** Les réglementations prudentielles mises en vigueur par le comité de Bâle n'avaient pas force de loi, donc certains établissements bancaires et agences de notation ne tiennent pas compte des mesures proposées. La crise des Subprimes illustre bel est fort ce problème. En effet, la sous évaluation des risques des crédits des «Subprimes», sortaient des bilans des banques via la titrisation, vers d'autres établissements financiers qui n'étaient forcément pas soumis à la réglementation prudentielle. Le modèle «originate and distribute» (octroi puis cession des crédits) n'incite pas les établissements financiers à donner une grande importance à l'évaluation du risque.

3.2.2 Les problèmes de méthodologie

- **La population et l'échantillonnage :** Les seules données disponibles pour construire un modèle de scoring sont relatives à l'historique des dossiers acceptés dont la variable à expliquer est connue (bon payeur ou mauvais payeur). La probabilité de défaut n'est ainsi estimée que pour les dossiers acceptés. Le modèle de scoring ne tient donc pas compte des demandeurs rejetés dès le départ, donc il ne pourra pas estimer la probabilité de défaut de ces derniers. Ses résultats sont donc biaisés car les dossiers refusés n'entrent pas dans l'étape d'apprentissage du modèle de score établi. Le problème de la mise à jour ou d'actualisation de l'échantillon doit aussi être effectué. En conclusion, deux types de redressements sont obligatoires :
 - Le premier a pour but de corriger la structure de l'échantillon pour remédier au problème des refusés.-
 - Le deuxième doit envisager de permettre de corriger les effets de vieillissement de l'échantillon pour lui redonner la structure de la demande

- **Le choix des variables :** Le choix des variables à introduire dans le modèle n'est pas évident. Il est nécessaire de mettre au point un algorithme de sélection afin de déterminer les variables les plus discriminantes et les plus significatives, pour la classification des dossiers de crédit. Une fois la sélection des variables faite, le modèle pourra alors être aisément construit.
- **Écarts entre prévisions et réalisations:** Une fois le modèle élaboré, il sera utilisé, en premier lieu, pour la sélection des dossiers basée sur la grille de score. D'autre part, il sera utilisé pour réaliser des prévisions d'impayés, compte tenu des politiques de sélection. Ses écarts entre prévisions et réalisations peuvent apparaître et doivent être expliqués puis maîtrisés. Ces écarts peuvent être dus à une évolution de l'environnement ou une évolution de la demande ou encore à des biais non pris en compte lors de la construction de l'échantillon.

Section 2 : Méthodologie d'élaboration de modèle scoring

1 Les types de modèle de crédit scoring⁵⁰

Pour qu'on puisse développer notre modèle score on doit tout d'abord savoir quel type de modèle on doit utiliser. Il existe deux types de modèle : Le déductif et l'empirique.

1.1 Le modèle déductif

Un système de Crédit Scoring déductif affecte des points à plusieurs attributs concernant le demandeur du crédit ; la somme de ces points constitue le score. Ces attributs et ces points sont prédéterminés par les décideurs qui se basent sur leurs propres expériences. Le fait de déduire ces points et ces attributs de cette façon implique qu'ils sont déterminés d'une façon plus ou moins subjective. C'est la raison pour laquelle on qualifie ce genre de modèles de «quasi-objectif », il est utilisé en grande partie pour aider dans la décision d'octroi de crédits aux particuliers (notamment pour la distribution des cartes de crédits).

1.2 Le modèle empirique

Les systèmes empiriques (ou statistiques) de Crédit Scoring sont construits sur la base de modèles techniques. La sélection des attributs (variables ou indicateurs) et la détermination de leurs poids, et donc le calcul des scores, sont basés sur des données de défaillances extraites du passé à l'aide d'algorithmes statistiques. L'avantage des modèles empiriques est non seulement qu'ils prennent en considération plusieurs critères simultanément, mais tiennent compte aussi de l'interdépendance entre ces critères et donc de la complexité du phénomène de la défaillance.

⁵⁰ Bensliman Amel et Lagagna Seif el islam, « Gestion du risque crédit lie au financement du commerce international par la méthode scoring », Mémoire de master en sciences commerciales, option: finance et commerce international.

Dans notre cas, notre travail portera sur le modèle empirique.

2 Démarches de construction d'un modèle scoring

Pour élaborer un modèle scoring on doit passer par les étapes suivantes :

- Collection des données : Choisir un critère de défaut et construire l'échantillon qu'on utilisera.
- Spécifier des données.
- Traitement des valeurs manquantes et des valeurs aberrantes.
- Choix des variables explicatives.
- Choix de la méthode statistique.
- Modélisation et test de validation.

2.1 Collection des données

2.1.1 Choix de critère de défaut et construction de l'échantillon

Le critère de défaillance

On peut considérer qu'il y a défaillance s'il y a faillite. Mais la faillite est le stade final de la défaillance. Des études considèrent des critères moins restrictifs comme le redressement judiciaire par exemple, mais ce dernier représente une mesure prononcée par un juge et qui est loin de constituer une mesure économique objective. Il est de fait problématique de définir comme entreprises saines, celles qui n'ont pas fait faillite, ou qui n'ont pas été engagées dans un redressement judiciaire.⁵¹

On distingue trois types d'entreprises de point de vue de la défaillance :

- Les entreprises saines : Ce sont celles qui remboursent leurs échéances le plus normalement.
- Les entreprises précontentieuses : Ce sont les entreprises qui ont des problèmes financiers graves et dont la situation n'est pas claire quant au remboursement.
- Les entreprises défaillantes : Ce sont les entreprises dont l'incapacité de remboursement est avérée et dont le seul moyen de remboursement est la procédure judiciaire.

Cette classification ne peut pas être utilisée dans le système de scoring classique car ce dernier distingue seulement deux types d'entreprises : Saine et Défaillante. Cependant, avec l'apparition du modèle de régression logistique multinomiale, le modèle de scoring peut être appliqué dans ce cas.

De point de vue juridique, dire qu'une entreprise est défaillante veut dire qu'elle est en face de liquidation, donc cette définition exclue les entreprises contentieuses qui sont considérées comme saines. Mais cela

⁵¹ Ouennoughi Sid-Ali, Modalités de détection des risques de défaillance au sein des PME, Essai d'analyse empirique.

pose problème aux banques car leurs objectifs c'est de récupérer l'argent prêté et les intérêts. C'est pour cela dans beaucoup de modèles, on préconise un critère arbitraire pour apprécier la défaillance, définie par des difficultés significatives.

Le comité de Bâle définit le défaut de la façon suivante : Un défaut de la part d'un débiteur intervient lorsque l'un des événements ci-dessous se produit :

- La banque estime improbable que le débiteur rembourse en totalité son crédit au groupe bancaire sans qu'elle ait besoin de prendre des mesures appropriées telles que la réalisation d'une garantie.
- L'arriéré du débiteur sur un crédit important dû au groupe bancaire dépasse 90 jours.
- Le report du paiement est associé à un événement de type abandon de créances.
- L'emprunteur est juridiquement en faillite.

Constitution de l'échantillon :

Une fois que le critère de défaillance est défini et les données historiques recueillies, on subdivisera notre échantillon en deux groupes : Entreprises saines et Entreprises défaillantes

Ces données historiques peuvent être :

- Des variables quantitatives : Soit des variables comptables (ratios de liquidité, ratios de structure financière, ratios de trésorerie, ratios de rentabilité...), ou bien des variables liées à l'information externes (transmissions des documents fiscaux et parafiscaux, engagements vis à vis de la centrale des risques, le mouvement de chiffres d'affaire confiés, ...)
- Des variables qualitatives : Comme l'âge, l'ancienneté, le statut juridique, etc.

2.2 Spécificités des données

2.2.1 Définition de l'horizon temporel⁵²

Le choix de l'horizon temporel peut être ramené en fait, à un choix entre un horizon unique pour tous les crédits, qui est généralement l'année et un horizon variable pour chaque type de crédit :

- Le premier choix est essentiellement guidé par des considérations pratiques, c'est pourquoi la plupart des modèles de risque de crédit développés à ce jour sont construits avec un horizon d'un an.
- Le second se fonde sur l'idée selon l'horizon temporel doit coïncider avec la durée de vie des crédits. Cela permet de mesurer plus fidèlement l'ensemble des pertes potentielles.

1 ⁵²<https://wikimemoires.net>, Construction et Paramètres du modèle de risque de crédit,

2.2.2 Répartition des données

La vérification de la performance d'un modèle score nécessite un échantillon de validation en plus de l'échantillon de construction, et cela requière un échantillon principal assez grand pour qu'on puisse le subdivisé, dans ce cas 60% à 80% de l'échantillon principal est utilisé pour la construction de modèle et 20% à 40% est utilisé pour valider le résultat. Et dans le cas où notre échantillon n'est pas assez grand on fait recours à des méthodes statistiques adéquates, comme la procédure Leave One Out Cross Validation - LOOCV.

2.3 Traitement des valeurs manquantes et des valeurs aberrantes

2.3.1 Les valeurs manquantes

Définition⁵³

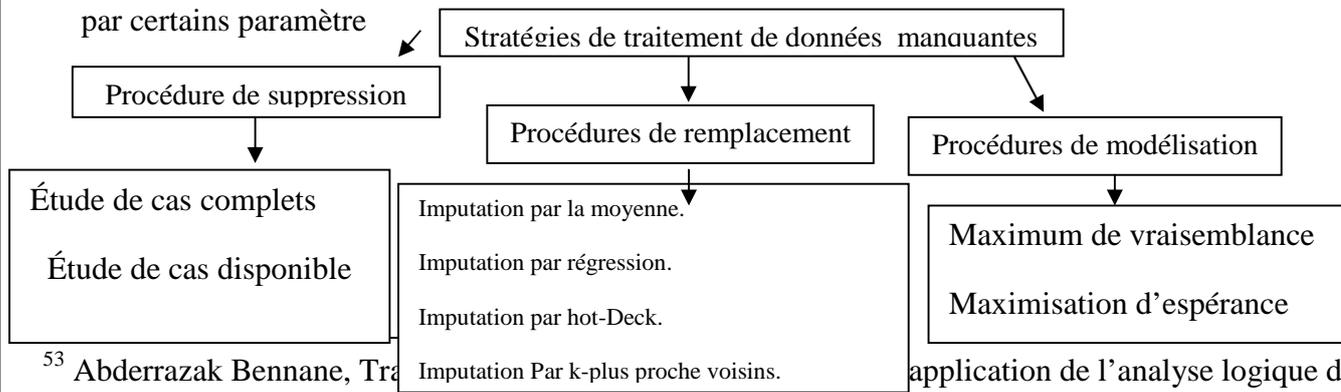
Une donnée incomplète est une donnée pour laquelle la valeur de certains attributs est inconnue. Ces valeurs sont dites manquantes. Soit l'observation est un vecteur des valeurs de certains indicateurs ou attributs, les valeurs manquantes peuvent être de deux natures :

- Valeur manquante totale, c'est-à-dire que toute l'observation manque.
- Valeur manquante partielle, c'est-à-dire que l'observation est présente mais il manque certaines valeurs de cette observation.

Traitement des valeurs manquantes :

Selon Kline (1998), Song et Shepherd(2007), il y a trois stratégies possibles pour traiter les données manquantes :

- L'utilisation des procédures de suppression ;
- L'utilisation des procédures de remplacement, (substitution) des données manquantes par les valeurs présentes ;
- L'utilisation des procédures de modélisation de la distribution des données manquantes et les estimés par certains paramètre



⁵³ Abderrazak Bennane, Tra application de l'analyse logique des données à la maintenance conditionnelle, Département de mathématique et de génie industriel, Ecole Polytechnique de Montréal.

Procédures de suppression

Étude des cas complets (Listwise deletion) : Cette méthode permet de se ramener à une base de données complète par réduction de la dimension du problème. Pour cela, tous les exemples de la base contenant des valeurs manquantes sont supprimés (On peut également choisir de supprimer toutes les variables dont certaines observations manquent, mais il faut être prudent car certaines peuvent être essentielles pour l'analyse).

Étude des cas disponibles (Pairwise deletion) : Dans cette méthode, on ne considère que les cas où ces variables sont complètement observées. Par exemple, si la valeur de l'attribut A est absente pour une observation, les autres valeurs pour le reste des attributs de la même observation pourraient encore être employées pour calculer des corrélations, telles que celle entre les attributs B et C. Comparée à la première méthode (Étude des cas complets), selon Roth, (1994), cette méthode conserve beaucoup plus de données qui auraient été perdues si on employait la méthode d'étude des cas complets.

Procédures de remplacement

Ces procédures visent à se ramener à une base complète en trouvant un moyen adéquat de remplacer les valeurs manquantes. On nomme ce procédé imputation, complétion ou substitution.

Imputation par la moyenne : Les valeurs manquantes de chaque attribut sont remplacées par la moyenne de l'attribut considéré. Il y a deux variantes de l'imputation par la moyenne : Imputation par la moyenne totale, imputation par la moyenne de sous-groupe. Pour l'imputation par la moyenne totale, la valeur absente d'un attribut est remplacée par la moyenne des valeurs de cet attribut de toutes les observations. Pour l'imputation par la moyenne de sous-groupe (classe), la valeur manquante est remplacée par la moyenne du sous-groupe (classe) de l'attribut en question. L'inconvénient de cette méthode est la sous-estimation de la variance et de biaiser la corrélation entre les attributs, cela veut dire que la distribution des données est loin d'être préservée.

Imputation par régression : C'est une approche en deux étapes : d'abord, on estime les rapports entre les attributs, et puis on emploie les coefficients de régression pour estimer la valeur manquante (Frane, 1976). La condition fondamentale de l'utilisation de l'imputation par régression est l'existence d'une corrélation linéaire entre les attributs. La technique suppose également que les valeurs sont manquantes au hasard. Dans le contexte des valeurs manquantes, deux modèles de régression sont en général

employés: la régression linéaire et la régression logistique. Cette dernière est plutôt utilisée pour traiter les variables discrètes, alors que la régression linéaire est appliquée sur des variables continues (Little et Rubin, 2002).

Imputation par hot-Deck : L'imputation hot-Deck est une procédure qui consiste à remplacer les valeurs manquantes d'une observation par des valeurs empruntées à d'autres observations similaires, définies comme étant celles pour lesquelles les valeurs sont les plus identiques à celles de l'observation présentant une donnée manquante, l'hypothèse sur laquelle elle s'appuie est que les probabilités de présence des valeurs sont égales dans les cas d'imputation.

Imputation par k-plus proches voisins : La technique k-plus proche voisin est une technique utilisée pour la substitution des valeurs manquantes avec la valeur du plus proche voisin dans l'ensemble de données. Pour chaque observation contenant des valeurs manquantes, on recherche ses k plus proches voisins. Dans le cas de variables continues, la valeur de remplacement correspond simplement à une moyenne pondérée des valeurs prises par ces k voisins pour la variable en question.

La difficulté réside dans le choix du paramètre k et de la métrique utilisée, les distances les plus utilisées étant l'eulidienne, celle de Mahalanobis ou encore celle de Pearson.

Imputation multiple : Cette méthode est suggérée par Rubin (1978), cela fait plus de 30 ans, et décrite en détail par Rubin (1987) et Schafer (1997). Dans le but de prédire une valeur pour toute donnée manquante, l'imputation multiple, au lieu de procurer une seule matrice à analyser, va à la place produire m matrices de données plausibles. Ces m matrices (souvent cinq suffisent) contiennent les mêmes données observées, mais les valeurs pour les données prédites peuvent être différentes. Cette variabilité entre les valeurs prédites des m matrices reflète l'incertitude face à l'imputation (Fichman, et Cummings, 2003) [48]. Ces matrices de données sont ensuite analysées comme si elles étaient des bases de données complètes, et elles sont combinées dans une unique base de données récapitulative.

Procédures basées sur un modèle

Maximum de vraisemblance : Sous sa forme plus simple, l'approche de maximum de vraisemblance pour analyser des données manquantes, suppose que les données observées sont tirées d'une distribution normale multi variée. Au lieu d'imputer des valeurs aux données manquantes, ces méthodes définissent un modèle à partir des données disponibles et basent les inférences de ce modèle sur la vraisemblance de la distribution des données sous ce modèle.

Maximisation d'espérance : Une dernière approche assez fréquente consiste à utiliser l'algorithme de maximisation d'espérance EM (Expectation-Maximization) pour estimer les valeurs manquantes qui est un processus itératif. Il est généralement utilisé pour estimer les paramètres d'une densité de probabilité. Il

peut être appliqué sur des bases de données incomplètes, et présente l'avantage de procéder à l'estimation des valeurs manquantes en parallèle de l'estimation des paramètres.

2.3.2 Les valeurs aberrantes⁵⁴

On appelle valeur extrême (ou aberrante) une donnée observée pour une variable qui semble anormale au regard des valeurs dont on dispose pour les autres observations de l'échantillon. On distingue deux types de situation dans lesquelles on rencontre des valeurs extrêmes :

- Une valeur extrême peut indiquer une erreur de lecture, une erreur de saisie ou un événement particulier qui a perturbé le phénomène observé au point de le rendre incomparable aux autres. Dans de tels cas, il faut soit corriger la valeur extrême si c'est possible, ou sinon supprimer l'observation.
- Une valeur extrême peut également être liée à un événement atypique, mais néanmoins connu ou intéressant à étudier.

Selon (Barnett et Lewis, 1994), une valeur aberrante est « une observation (ou sous-ensemble d'observations) qui semble être incompatible avec le reste de cet ensemble de données ».

2.4 Le choix des variables explicatives

Les variables explicatives de défaut correspondent aux facteurs de risque dont la mise en jeu a conduit à une situation de défaut. Elles doivent traduire des dimensions variées du risque de sorte qu'elles ne soient pas corrélées.⁵⁵

Divers types de données peuvent être utilisés :

- Comptables et financiers : comprennent les ratios comptables et financiers, tels que : ratios de liquidité, ratios de structure financière, ratios de trésorerie, ratios de rentabilité
- Bancaires : concerne le fonctionnement de compte bancaire de l'entreprise.
- Qualitatives : c'est toutes les données qui ne peuvent pas être quantifiées (l'âge, la profession, la localisation, le statut juridique ...).Elle sont rendues dichotomiques grâce à un codage.

2.5 Choix de la méthode statistique

Il existe plusieurs techniques pour la construction des modèles de score, on peut les diviser en deux groupes : l'approche paramétrique et l'approche non paramétrique.

⁵⁴<https://www.xlstat.com>

⁵⁵<https://wikimemoires.net/modèles>, Modèles de scoring: construction du système de scoring.

2.5.1 L'approche paramétrique

Ce sont des approches qui se basent sur l'hypothèse, la distribution des indicateurs et sur les échantillons des entreprises saines et celles défaillantes. Le plus souvent ces hypothèses ne sont pas vérifiées ce qui peut mettre en cause les résultats de ce genre de modèle. Cependant, ces modèles sont toujours utilisés à cause de leur capacité prédictive prouvée dans la pratique

On peut citer plusieurs exemples :

- L'analyse discriminante : c'est une technique statistique qui vise à décrire, expliquer et prédire l'appartenance à des groupes prédéfinies d'un ensemble d'observations à partir d'une série de variables prédictives (descripteurs, variables exogènes ...)
- La régression logistique : c'est une technique statistique dont l'objet est de produire un modèle permettant de prédire les valeurs prises par une variable catégorielle, le plus souvent binaire, en se basant sur une série de variables explicatives continues et /ou binaires.

Contrairement à l'analyse discriminante, la régression logistique n'exige pas une distribution normale des prédicteurs ni l'homogénéité des variances.

2.5.2 L'approche non paramétrique

Contrairement aux approches paramétriques, ces approches ne se basent pas sur des hypothèses quant aux distributions statistiques des variables. Plusieurs études ont montré leurs performances tout en présentant quelques inconvénients tels que l'absence d'une fonction score et un temps de calcul relativement long.

On peut citer :

- Techniques d'intelligence artificielle : l'intelligence artificielle consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle comme l'analyse neurale.⁵⁶
- Techniques d'enveloppement des données (DEA) : la méthode DEA est une méthode non paramétrique qui n'impose aucune forme fonctionnelle spécifique. Elle est fondée sur la programmation linéaire pour identifier des fonctions de production empiriques. Son intérêt majeur réside dans le calcul d'une mesure synthétique et compréhensive de la performance d'organisation appelé DMUS⁵⁷.

⁵⁶ L'analyse neuronale regroupe des modèles dont l'intention est d'imiter certaines fonctions du cerveau humain en reproduisant certaines de ces structures de base. Elle est inspirée d'un modèle neurophysiologique d'apprentissage.

⁵⁷ Decision Making Unite : Unité décisionnelle qui peuvent être des firmes, des agences, des établissements publique,... etc.

2.6 Modélisation et tests

C'est l'étape de la construction de la fonction score en combinant les variables explicatifs significatifs et la variable dépendante. Après l'estimation du modèle, on procède aux tests statistiques pour s'assurer de la robustesse du modèle.

Cette fonction nous permettra de distinguer les bons emprunteurs des mauvais.

2.7 La validation

La dernière étape de la construction d'un modèle de Crédit Scoring est celle de la validation. Il s'agit de mesurer la performance du modèle obtenu ainsi que sa capacité prédictive.

La validation du score se fait donc à l'aide d'observations supplémentaires, mises de côté. L'évaluation de la performance du modèle est effectuée pour savoir si le modèle va dans la direction de la réalité et s'il correspond à la logique financière. Toute déviation de ses deux points doit être expliquée sinon rejetée.

Section 3 : Techniques de classification et de validation des modèles de crédit scoring.

L'élaboration d'un système de scoring peut se faire par divers méthodes comme nous l'avons vu dans la section précédente, certaines sont paramétriques, elles se basent sur le raisonnement probabiliste, et d'autres non paramétrique, elles se basent sur le raisonnement humain.

Dans notre cas c'est les méthodes paramétriques les plus utilisées en crédit scoring (l'analyse discriminante, la régression linéaire et la régression logit et probit) que nous aborderons.

1 L'analyse discriminante

1.1 Définition

L'analyse discriminante est une technique statistique qui permet de discriminer entre les observations compte tenu de leurs caractéristiques individuelles. Elle est utilisée afin de classer et/ou prévoir une variable dépendante de type qualitatif. Son application empirique a commencé depuis les années 1930 avec les travaux de Fisher et Mahalanobis (1936).

L'Analyse discriminante de Fisher consiste à chercher des combinaisons linéaires de descripteurs quantitatifs, indicateurs synthétiques qui permettent de classer les individus correctement dans chacun des groupes. Cette méthode, essentiellement analytique, est basée sur des concepts géométriques, cependant

pour que ces combinaisons linéaires puissent être optimales (au sens où le risque d'erreur de classement serait alors minimal), les données doivent vérifier certaines hypothèses de nature probabiliste, tel que la normalité des données et d'homoscédasticité.⁵⁸

1.2 Construction de modèle

Pour effectuer une analyse discriminante deux approches peuvent être utilisés :

1.2.1 L'approche géométrique

Le principe de cette méthode est d'avoir une droite (un hyperplan / une séparation linéaire) entre deux groupes d'observations. Cette droite est une combinaison linéaire des variables explicatives, toutes continues, qui décrivent les deux groupes d'observations. Chacun des deux groupes d'observations est « remplacé » (synthétisé) par son barycentre (point de coordonnées moyennes de toutes les variables explicatives). La séparation entre les deux populations est faite selon une droite perpendiculaire au segment reliant les deux barycentres.⁵⁹ Cette approche ne repose sur aucune hypothèse probabiliste.

1.2.2 L'approche probabiliste

Elle se base sur des hypothèses de probabilités en prenant en compte les probabilités à priori d'appartenance des individus à l'une des deux sous populations et l'équivalence des matrices de variance-covariance.

1.3 Sélection des variables⁶⁰

Plusieurs méthodes peuvent être utilisées dans le choix des variables à inclure dans l'édification des fonctions discriminantes.

On peut citer l'approche hiérarchique (stepwise) où les variables sont introduites une à une selon leur capacité décroissante à mettre en évidence la différence entre les groupes. Au cours des sélections successives, il est possible que des variables déjà entrées perdent leur pouvoir de discrimination: la raison en est une redondance d'information, c'est-à-dire que le pouvoir de discrimination de cette variable est désormais inclus dans quelque combinaison de nouvelles variables retenues. Donc à chaque étape de l'analyse, on procède à l'élimination des variables devenues inutiles. La variable qui fait diminuer le lambda de Wilks est introduit.

⁵⁸ Dominique Desbois, Une introduction à l'analyse discriminante avec spss pour Windows.

⁵⁹ <https://od-datamining.com>

⁶⁰ Michel Calciu (EUDIL-IAE de Lille) et Christophe Benavente (IAE des pays de l'Adour), L'analyse discriminante.

Lambda de Wilks : Le lambda de Wilks est une statistique qui compare simultanément plusieurs moyennes en mesurant la part d'inertie intercalasse dans l'inertie totale. Plus précisément, c'est le déterminant de la matrice des variances-covariances intercalasse rapporté au déterminant de la matrice des variances-covariances totale. Il est donc compris entre 0 (bonne discrimination) et 1.⁶¹

Le Lambda Λ de Wilks est donné par la formule : $\Lambda = \frac{|M|}{|N|}$

Où : M : Matrice des variance-covariance intra classe ;

N : Matrice de variance-covariance totale.

On suppose que les matrices de covariance des différentes sous-populations sont identiques (ce qui se vérifie avec un test de Box) et que les variables présentent des distributions normales (on peut le vérifier par le test de Kolmogorov-Smirnov).

- Il n'existe pas de niveau unique à partir duquel notre lambda indique que les sous-populations sont suffisamment différentes. Tout dépend du nombre de classes, de l'effectif, du nombre de variables et bien sûr du niveau de confiance que l'on se donne.
- La signification de la statistique de lambda de Wilks entre deux pas est vérifiée par la statistique de variation de Fisher.⁶²

$$(n-k-p/k-1) ((\Lambda_p/\Lambda_{p+1}) - 1) \rightsquigarrow (k-1, n-k-p)$$

Où : n : Nombre d'observation.

K : Nombre de groupe.

Λ_p : Lambda de Wilks pour la p^{ème} variable.

On rejette l'hypothèse de l'apport de la variable à la discrimination, si :

$$(n-k-p/k-1) ((\Lambda_p/\Lambda_{p-1}) - 1) > F_{1-\alpha}(k-1, n-k-p)$$

- Plus Λ de Wilks est petit, plus la discrimination entre les groupes est meilleure.

⁶¹ <http://www.jybaudot.fr>, Techniques et concepts de l'entreprise, de la finance et de l'économie (et fondements mathématiques).

⁶² G. Saporta, Probabilités, analyse de données et statistique, édition TECHNIP, Paris.

1.4 Test de validation⁶³

Elle consiste à estimer le taux théorique d'erreur de classement (appelé aussi risque ou cout) à partir d'un ensemble d'individus répartis en K groupes soumis à une analyse discriminante. Cette estimation s'obtient de différentes manières, que l'on peut mettre en œuvre dans les analyses.

1.4.1 Méthode de resubstitution

De manière classique en analyse discriminante, pour évaluer les performances d'une fonction de classement, nous confrontons ses prédictions avec les vraies valeurs de la variable à prédire sur un fichier de données. Le tableau croisé qui en résulte s'appelle une matrice de confusion de terme général s avec : en lignes les vrais groupes d'appartenance, en colonnes les groupes d'appartenance prédits. Le taux d'erreur ou taux de mauvais classement est tout simplement, lorsque la prédiction ne coïncide pas avec la vraie valeur, rapporté à l'effectif du fichier de données.

L'estimation du taux théorique d'erreur de classement est égale au pourcentage d'individus de l'échantillon mal classés par la règle d'affectation, on parle de taux d'erreur en resubstitution :

$$\widehat{C}^{res} = 1 - \frac{1}{n} \sum_{k=1}^K n_{kk}$$

Cette estimation, obtenue en appliquant la règle d'affectation à tous les individus ayant contribué à la construire, est une estimation dite échantillon dépendante. C'est une estimation biaisée.

Remarque : Cette approche n'est pas recommandée en pratique puisqu'on a tendance à sous-estimer le taux d'erreur. Pour obtenir une estimation plus précise du taux d'erreur théorique de classement, on a recours à la méthode suivante.

1.4.2 Méthode de l'échantillon-test

Cette méthode est valable dans le cas des grands échantillons. La procédure est de construire la fonction de classement sur une fraction de données, dites d'apprentissage, puis de l'évaluer sur une autre fraction de données, dites de test. On effectue au hasard un tirage de 20 à 30% des individus dans chaque groupe, ce qui constitue l'échantillon test, répartis en K sous-ensembles (T_1, \dots, T_K), le reste (de 70 à 80%) est l'échantillon d'apprentissage (échantillon de base) constitué de K sous-ensembles (B_1, \dots, B_K). L'analyse

⁶³ Jean-Pierre Nakache, Josiane Confais, Statistique explicative appliquée: Analyse discriminante, modèle logistique, segmentation par arbre, édition TECHNIP.

discriminante est alors effectuée sur les individus de l'échantillon de base. Il en résulte une règle d'affectation qui, appliquée à l'échantillon-test, fournit une estimation du taux théorique d'erreur de classement qui est le pourcentage d'individus de l'échantillon-test mal classés par la règle d'affectation issu du tableau de classement.

$$\hat{C}^{ct} = 1 - \frac{1}{n} \sum_{k=1}^K n_{kk}$$

\hat{C}^{ct} Est une estimation meilleure (estimation sans biais) que celle obtenue par resubstitution, puisqu'elle est obtenue à partir d'un échantillon indépendant de l'échantillon de base.

Remarque : On peut sélectionner une règle de classement permettant de minimiser l'erreur de classement afin de sélectionner les variables discriminantes en utilisant la méthode dite séquentielle:

1. On sélectionne la variable j_1 qui permet d'obtenir le plus faible taux d'erreur dans un modèle à 1 variable.
2. On sélectionne la variable j_2 qui permet d'obtenir le plus faible taux d'erreur, lorsqu'elle est utilisée conjointement à j_1 dans un modèle à 2 variables.
3. On sélectionne la variable j_3 qui permet d'obtenir le plus faible taux d'erreur, lorsqu'elle est utilisée conjointement à j_1 et j_2 dans un modèle à 3 variables etc.

2 La régression logistique :

2.1 Définition :

La régression logistique est l'analyse appropriée quand la variable dépendante est dichotomique⁶⁴ (binaire). Comme toutes les analyses de régression, la régression logistique est une analyse prédictive. Elle sert à décrire les données et à expliquer la relation entre une variable dépendante et une ou plusieurs variables indépendantes nominales ou ordinales.

Elle est similaire à la régression linéaire mais elle convient aux modèles dans lesquelles les variables sont dichotomiques. Les coefficients de la régression logistique peuvent servir à estimer des taux de probabilité pour chacune des variables indépendantes d'un modèle. La régression logistique s'applique à une plus large gamme de situations de recherche que l'analyse discriminante.

⁶⁴ Les variables dichotomiques ce sont des variables qualitatives qui ne prennent que 2 modalités (0/1, oui/non, vrai/faux).

2.2 Construction du modèle régression logistique binaire :

On considère un échantillon de N individus indicés $i=1, \dots, N$. Pour chaque individu, on observe si un certain événement s'est réalisé et l'on note y_i la variable codée associée à l'événement. On pose, $\forall i \in [1, N]$

$$Y_i = \begin{cases} 1 & \text{si l'événement s'est réalisé pour l'individu } i \\ 0 & \text{si l'événement ne s'est pas réalisé pour l'individu } i \end{cases}$$

On considère le modèle dichotomique suivant :

$$p = p(x) = F(x\beta), \forall i = 1 \dots N$$

) La fonction de répartition $F(\cdot)$ correspond à la fonction logistique : $\forall w \in \mathbb{R}$:

$$F(w) = \frac{1}{1 + \exp(-w)} = \frac{\exp(w)}{1 + \exp(w)}$$

➤ Le modèle logit définit la probabilité associée à l'événement $y_i=1$, comme la valeur de la fonction de répartition de la loi logistique considérée au point x_i :

$$\text{Modèle logit: } p_i = p(x_i) = \frac{1}{1 + e^{-x_i \beta}}, \forall i=1, \dots, N$$

L'estimation des coefficients β de modèle logistique s'effectue par l'estimation du maximum de vraisemblance.

La fonction de maximum de vraisemblance est écrite comme suit:

$$L(\beta; Y) = \prod_{i=1}^n (p)^{y_i} (1 - p)^{1 - y_i}$$

On travaillera avec le log de la fonction de vraisemblance qui nous facilitera les calculs :

$$l(\beta; Y) = \ln L(\beta; Y) = \sum_{i=1}^n (Y_i \log(p) + (1 - Y_i) \log(1 - p))$$

Nous effectuerons une différenciation de la fonction $l(\beta; Y)$ par rapport au vecteur β pour pouvoir estimer les coefficients qui maximise le log de la fonction de vraisemblance.

2.3 La sélection des variables :

- Nous voulons expliquer les données de la façon la plus simple possible les prédicateurs redondants devraient être supprimés car ils sont inutiles et ils ajouteront du bruit à l'estimation d'autres quantités qui nous intéressent. La colinéarité est causée par un trop grand nombre de variables qui essaient de faire le même travail.
- A travers certaines procédures de sélection on pourra éviter ces problèmes.
- La sélection des variables vise à sélectionner le meilleur sous-ensemble de variables prédicateurs. Elle peut s'effectuer par trois étapes comme le montre le schéma ci-dessous : la méthode ascendante (Forward selection), la méthode descendante (backward selection) et la sélection pas à pas (stepwise selection).

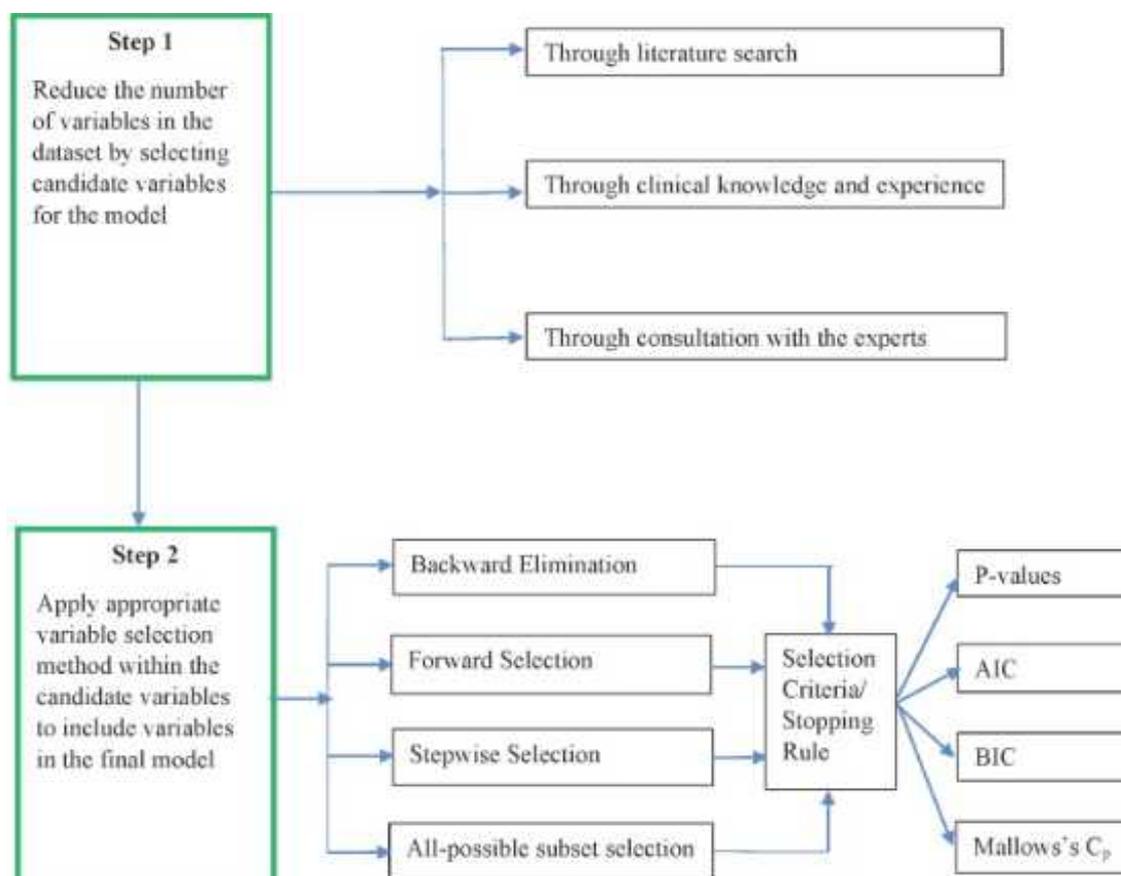


Figure 1: Description des étapes à suivre pour la sélection des variables lors de la création du modèle.

2.3.1 La sélection ascendante : (Forward selection)

Cette procédure part d'un modèle vide sans aucune variable, il contient juste la constante β .

- Les variables sont ajoutées une à une ;

- A chaque itération, on effectue un test de significativité, où on retient la variable dont le p-value est petit par rapport au seuil fixé ;
- Le processus s'arrête soit quand il n'y a plus de variable à ajouter, soit quand un certain critère est satisfait ;
- Une fois qu'une variable a été ajoutée, la FS ne peut la retirer.

2.3.2 La sélection descendante : (backward selection)

Cette procédure à l'inverse de l'ascendante, commence par un modèle complet qui contient toutes les variables.

- On supprime une à une les variables x dans le modèle ;
- A chaque itération, on supprime la variable dont le p-value la plus élevée est supérieure à au seuil fixé (variable non significatif).

2.3.3 La sélection pas à pas : (stepwise)

- Cette méthode est une combinaison de procédures de sélection avant et arrière qui permet de se déplacer dans les deux sens, en ajoutant et en supprimant des variables à différentes étapes.
- Le processus peut commencer par une approche d'élimination en amont et de sélection en aval. Par exemple, si la sélection par étapes commence par la sélection avant, les variables sont ajoutées au modèle une par une en fonction de la signification statistique. À chaque étape, après l'ajout d'une variable, la procédure vérifie toutes les variables déjà ajoutées au modèle pour supprimer toute variable non significative dans le modèle. Le processus se poursuit jusqu'à ce que chaque variable du modèle soit significative et que chaque variable exclue soit non significative.
- La sélection par étapes nécessite deux niveaux de signification distincts (seuils) pour ajouter et supprimer des variables du modèle. Les niveaux de signification pour l'ajout de variables doivent être inférieurs aux niveaux de signification pour la suppression des variables afin que la procédure n'entre pas dans une boucle infinie.
- Dans le cadre de la sélection par étapes, l'élimination en amont est souvent privilégiée, car dans l'élimination en amont, le modèle complet est pris en compte et l'effet de toutes les variables candidates est évalué.

2.4 Interprétation des paramètres :

Le modèle logit nous permet d'interpréter les coefficients des variables explicatives grâce à « l'Odds Ratio ».

- L'Odds Ratio est une mesure statistique, exprimant le degré de dépendance entre des variables aléatoires. Il représente les chances qu'un résultat se produise (Y=1) compte tenu d'une exposition particulière, par rapport aux chances que le résultat se produise en l'absence de cette exposition. Il est défini par la formule suivante :

$$O = \frac{G(x+1)/(1-G(x+1))}{F(x)/(1-F(x))} = e^{\beta}$$

$$\text{Avec: } G(x) = G(x) = \frac{1}{1+e^{-x}}$$

- OR = 1 : L'exposition n'affecte pas les chances de résultat
- OR > 1 : Exposition associée à des chances de résultat plus élevées
- OR < 1 : Exposition associée à une moindre probabilité de résultat

Intervalle de confiance :

L'intervalle de confiance (IC) à 95% est utilisé pour estimer la précision de l'OR. En pratique, l'IC à 95% est souvent utilisé comme indicateur de la présence d'une signification statistique s'il ne chevauche pas la valeur nulle (par exemple OR = 1).

Un grand CI indique un faible niveau de précision de l'OR, tandis qu'un petit CI indique une plus grande précision de l'OR.

2.5 Validation du modèle :

Après l'estimation des coefficients, plusieurs étapes sont nécessaires pour évaluer la pertinence, l'adéquation et l'utilité du modèle. Tout d'abord, l'importance de chacune des variables explicatives est appréciée en effectuant des tests statistiques de la signification des coefficients. La qualité globale de l'ajustement du modèle est ensuite testée. De plus, la capacité du modèle à faire la distinction entre les deux groupes définis par la variable de réponse est évaluée.

2.5.1 Le test de Wald :⁶⁵

Les statistiques Wald sont utilisées pour tester la significativité des coefficients individuels dans le modèle et sont calculées comme suit:

$$\left(\frac{C}{L'éc} \frac{t_j}{d} \frac{it}{c} \right)^2$$

⁶⁵Liz Cheek et Jonathan Ball' Statistics review 14: Logistic regression.

- Chaque statistique de Wald est comparée à une distribution khi-deux (2) avec 1 degré de liberté.
- Les statistiques de Wald sont faciles à calculer mais leur fiabilité est discutable, en particulier pour les petits échantillons.
- Pour les données qui produisent de grandes estimations du coefficient, l'erreur-type est souvent gonflée, ce qui entraîne une statistique de Wald plus faible, et par conséquent, la variable explicative peut être supposée à tort comme sans importance dans le modèle.
- La constante n'a pas d'interprétation pratique simple mais est généralement conservée dans le modèle indépendamment de sa signification.

2.5.2 Test de significativité globale :⁶⁶

- C'est un test de spécification du modèle qui compare généralement deux modèles.
- La statistique test qui est associée correspond à la différence entre les logs de vraisemblance respectifs du modèle dit contraint et du modèle sans restriction.
- La technique d'estimation par maximum de vraisemblance utilisée pour dériver les paramètres estimés de nos modèles maximise la fonction de vraisemblance.
- L'idée sous-jacente à ce test est que l'omission de toutes variables explicatives pertinentes (modèle contraint) impliquera une diminution importante du log vraisemblance. Ainsi, ce ratio devrait traduire dans quelle mesure l'omission de certaines variables viendra affecter notre inférence indiquant à cet effet, la pertinence d'une spécification par rapport à une autre. La statistique test correspondante, du nom de ratio de vraisemblance est la suivante:

$$L = 2(\log li_{ho_{s_r}} - \log li_{ho_{a_r}})$$

- Cette statistique est comparée à une khi-deux. La règle de décision est que si cette statistique est plus grande que le quantile associé au khi-deux, on rejette l'hypothèse nulle qui correspond au modèle restreint. De façon équivalente, plus le log vraisemblance est petit, meilleure est l'estimation et meilleur est le modèle estimé. L'avantage de ce test est qu'il assure la plus grande puissance selon le théorème de Neyman-Pearson.

2.5.3 Test de Hosmer et Lemeshow :

- Hosmer et Lemeshow ont proposés un test du khi deux sur un tableau où les modalités sont définies à partir des probabilités prédites.
- Le test Hosmer-Lemeshow (test HL) est un test de qualité d'ajustement pour la régression logistique, en particulier pour les modèles de prédiction des risques. Il indique dans quelle mesure vos données

⁶⁶Woybi Touré, Modèle de probabilité de défaut des prêts d'une banque canadienne : science de gestion (ingénierie financière).

correspondent au modèle. Plus précisément, le test HL calcule si les taux d'événements observés correspondent aux taux d'événements attendus dans les sous-groupes de population.

- Un test significatif indique que le modèle n'est pas un bon ajustement et un test non significatif indique un bon ajustement.

2.5.4 Le coefficient de détermination généralisé :

- La statistique R n'est pas fournie par SPSS, mais peut être calculée à la main. Elle représente la corrélation partielle entre chaque variable indépendante et la variable dépendante et se situe toujours entre - 1 et 1. Lorsque la valence est positive, la valeur du prédicteur augmente en même temps que la probabilité Y, alors que lorsque la valence est négative, la probabilité Y diminue quand la valeur du prédicteur augmente. Plus petite est la valeur calculée, plus faiblement le prédicteur contribue au modèle.⁶⁷

$$R = +/- \sqrt{\frac{\text{Wald} - (2 \times \text{ddl})}{-2LL(\text{original})}}$$

3 La régression Probit :

Les modèles probit ont été introduits par Chester Bliss en 1934. C'est un type de régression où la variable dépendante ne peut prendre que deux valeurs (oui/non).

En raison des caractéristiques non linéaires de ces modèles, il est nécessaire d'utiliser la méthode du maximum de vraisemblance pour l'estimation des paramètres.

Dans le cas du modèle probit, nous utilisons la fonction de distribution cumulative de la loi normale :

$$P_i = \int_{-\infty}^{\alpha + \beta x_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt.$$

Étant donné P_i et en supposant que les valeurs par défaut sont indépendantes, nous pouvons former le logarithme de la fonction de vraisemblance comme suit:

$$\ln L = \sum_{i=1}^n y_i \ln P_i + \sum_{i=1}^n (1 - y_i) \ln(1 - P_i)$$

Pour la sélection des variables, nous pouvons utiliser la méthode par stepwise et pour tester la signification des modèles, le test du rapport log-vraisemblance et le test de Wald sont généralement utilisés (mêmes procédures que la régression Logit).

⁶⁷<http://spss.espaceweb.usherbrooke.ca/>

Section 4 : Validation des modèles de crédit scoring

Il est important une fois que le modèle scoring est fait de s'assurer de sa performance et ce qui le privilège par rapport aux modèles existant déjà. Pour cela on présentera dans cette section les différentes méthodes qui nous permettront de le faire.

1 Le taux de bon classement :

- L'évaluation de la règle de décision se fait par l'estimation du taux d'erreur.
- Le taux de bons classements et le taux d'erreur sont estimés à partir des données qui concernent l'affectation possibles d'une observation à une classe, fournis par le tableau d'affectation (matrice de confusion).

Tableau 4: Table d'affectation

	Affectation	
Situation réelle	Mauvais payeur	Bon payeur
Mauvais payeur	Taux de vrai positif	Taux de faux positif
Bon payeur	Taux de faux négatif	Taux de vrai positif

- Le taux de bon classement réel estimé pour les Mauvais payeurs est donné par la relation : $\frac{T}{T + T}$
- Le taux de bon classement réel estimé pour les Bons payeurs est donné par la relation : $\frac{T}{T + T}$
- Le taux de bon classement réel général estimé sur l'échantillon est donné par la relation : $\frac{T + T}{T + T + T + T}$

On note que :

L'erreur de type 1: représente les dossiers de crédit réellement mauvais, déclarés par le modèle comme Bons (risque de non remboursement du crédit)

L'erreur de type 2: représente les dossiers de crédit réellement bons, qualifiés par le modèle de mauvais payeurs (manque à gagner). Chaque erreur de classement engendre un coût supplémentaire relatif à la mauvaise décision que l'on aurait prise. Il s'agit alors de minimiser ce coût causé par le classement incorrect des éléments.

2 La courbe : ROC

Une courbe ROC (**Receiver Operating Characteristic**) est un graphique représentant les performances d'un modèle de classification pour tous les seuils de classification. Cette courbe trace le taux de vrais positifs (VP) en fonction du taux de faux positifs (FP): Taux de vrais positifs, Taux de faux positifs. Une courbe ROC qui trace les valeurs TVP et TFP pour différents seuils de classification.

Choix d'un seuil :

On utilise parfois la courbe ROC pour choisir un seuil. En pratique, on peut prendre le seuil correspondant au point de la courbe la plus éloigné de la première bissectrice et le plus prêt du point supérieur gauche (0,1). Ou encore le seuil correspondant au point où la pente de la courbe est la plus proche de 0. Mais on peut également choisir le seuil qui minimise le taux d'erreur (ou un autre indicateur) sur l'ensemble d'apprentissage

Diminuer la valeur du seuil de classification permet de classer plus d'éléments comme positifs, ce qui augmente le nombre de faux positifs et de vrais positifs. La figure ci-dessous représente une courbe ROC classique.

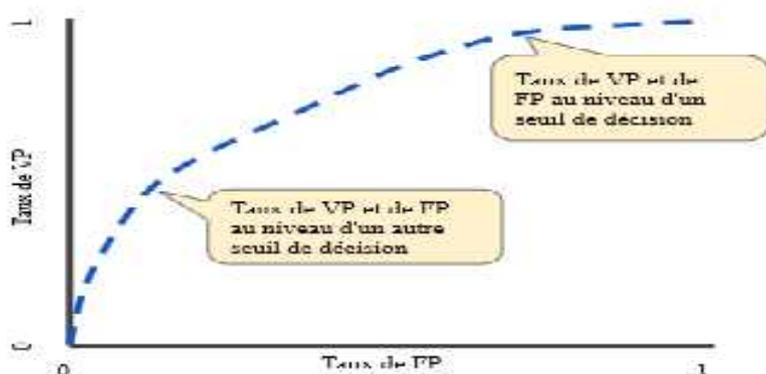


Figure 2 : Taux de VP et de FP pour différents seuils de classification

❖ Pour calculer les points d'une courbe ROC, nous pourrions effectuer plusieurs évaluations d'un modèle de régression logistique en variant les seuils de classification, mais ce serait inefficace. Nous pouvons en revanche calculer efficacement l'aire sous cette courbe, ou AUC, grâce à un algorithme de tri. .

❖ L'indice AUC signifie "aire sous la courbe ROC". Cette valeur mesure l'intégralité de l'aire à deux dimensions située sous l'ensemble de la courbe ROC (par calculs d'intégrales) de (0,0) à (1,1).

3 La méthode de l'échantillon-test⁶⁸

Cette méthode consiste à partager l'échantillon en deux parties :

- ❖ une partie (de l'ordre de 80%) sert d'échantillon d'apprentissage de la règle de décision et l'autre partie sert d'échantillon-test et permet de tester la règle d'affectation et donc de calculer le taux d'erreurs.
- ❖ Cette méthode est plus fiable mais nécessite un échantillon plus important.
- ❖ En effectuant plusieurs tirages aléatoires d'échantillons d'apprentissage (et donc d'échantillons-test), on peut améliorer l'estimation du taux d'erreurs en calculant la moyenne (et l'écart-type) des valeurs des taux d'erreurs obtenues à chaque tirage

Conclusion

Comme nous l'avons vu dans ce chapitre la construction d'un modèle de crédit Scoring nécessite une méthodologie bien définie qui englobe plusieurs étapes (construction du modèle, sélection des variables et la validation) et divers tests statistiques qui nous permettent d'estimer la significativité de nos variables. En dernier lieu nous avons abordés quelques méthodes qui nous permettent d'estimer la performance de notre modèle.

Ceci nous conduit à envisager la construction d'un modèle de crédit scoring dans le chapitre suivant en appliquant tout ce que nous avons acquis sur le plan théorique.

⁶⁸ Marie Chavent, « Le Scoring » cours- Master MIMSE – 2^{ème} année, université de Bordeaux.

Chapitre III : la mise en œuvre d'un crédit scoring au niveau du CPA

introduction :

Afin de mettre en application les divers éléments théoriques présentés dans les chapitres précédents, nous avons entrepris un stage pratique au sein de la Direction de Surveillance de Risque Crédit –DSRC- duCrédit Populaire d’Algérie (CPA), où nous avons pu collecter des informations sur 144 entreprises privées de statuts juridiques et secteurs différents afin d’élaborer un modèle de notation « scoring ».Le présent chapitre est subdivisé en deux parties. En premier lieu nous faisons une présentation générale de l’établissement d’accueil puis nous présenterons les données collectées et le modèle scoring élaboré.

Chapitre III : la mise en œuvre d'un crédit scoring au niveau du CPA

Section 1 : Présentation du Crédit Populaire d'Algérie

1 Présentation générale

Le Crédit Populaire d'Algérie (la Banque) est une banque publique algérienne qui a son importance dans le financement de l'économie algérienne avec une part de marché estimée à 13%, elle est issue du démembrement des banques populaires françaises après l'indépendance de l'Algérie en 1962. Les informations suivantes permettent de mieux présenter le CPA:

- Nom de la Banque : Crédit Populaire d'Algérie ;
- Année de création : en 1966 ;
- Forme juridique : Entreprise Publique Économique, Société par actions ;
- Capital : 48 000 000 000 Dinars Algériens ;
- Nombre d'agences bancaires : 141 agences regroupées en 15 groupes d'exploitation (directions régionales) ;
- Nombre d'employés : 3738 pendant l'année 2015 ;
- Identité visuel :



- Signature : « Une banque à votre écoute » ;

Dans ce paragraphe nous présentons quelques événements importants qui ont façonnés la banque:

- 1966 : Création de la banque en reprenant l'activité de cinq (05) banques étrangères qui sont les banques populaires commerciales et industrielles (Alger, Oran, Annaba et Constantine) et la Banque Populaire du Crédit d'Algérie ;
- 1967 : Le CPA reprend l'activité de quatre (04) autres banques étrangères ;
- 1985 : Le CPA donne naissance à la BDL (Banque de Développement Locale) ;
- 1988 : La Banque devient une entreprise publique économique, société par actions ;
- 1997 : Le CPA obtient son agrément auprès du Conseil de la Monnaie et de Crédit ;
- 2002 : Première tentative de privatisation, non concluante ;
- 2007 : Deuxième tentative de privatisation, également non concluante.

2 Activités de la Banque :

Le CPA a pour mission de promouvoir le développement des secteurs : BTPH, santé et médicaments, commerce et distribution, hôtellerie et tourisme, les médias, l'artisanat. Parmi ses activités nous pouvons citer :

- **Activité commerciale :** le CPA intervient dans le financement de l'économie via l'octroi du crédit aux entreprises ainsi qu'aux particuliers et professionnels, comme il accomplit une fonction de bancarisation, collecte des ressources et diversification des produits et services ;
- **Activité internationale :** cette activité est inscrite parmi les priorités de la Banque. Le CPA poursuit l'effort de développement de son offre en ce qui concerne ce volet ;
- **Activité de trésorerie :** sur la place interbancaire, la Banque maintient généralement une position de prêteuse, elle répond favorablement aux exigences de la Banque d'Algérie en matières des réserves obligatoires ;
- **Activité de marché :** le CPA gère un portefeuille dominé par les titres de l'État et les obligations des sociétés et des actions.

3 Axes stratégiques

Le CPA définit son plan stratégique d'une façon quinquennale en se fixant des objectifs réalisables sur le moyen terme. La banque cible les axes stratégiques concernés puis elle définit les actions stratégiques ad hoc à performer durant les cinq années.

La finalité de la Banque via sa stratégie est de garantir sa pérennité et viser l'innovation et l'amélioration de la gamme des produits d'une façon continue pour mieux répondre aux attentes de la clientèle.

4 Organisation de la Banque

L'organisation de la banque comprend :

- L'Assemblée Générale des Actionnaires,
- Le Conseil d'Administration,
- La Présidence Direction Générale,
- Le Cabinet du PDG, le Secrétariat Général,
- Onze (11) Divisions et la Direction de l'Organisation et de la Réglementation placée sous l'autorité directe du Président Directeur Général.
- Elle comprend, également, des organes d'aide à la prise de décision, d'assistance à la gestion et de coordination des activités.

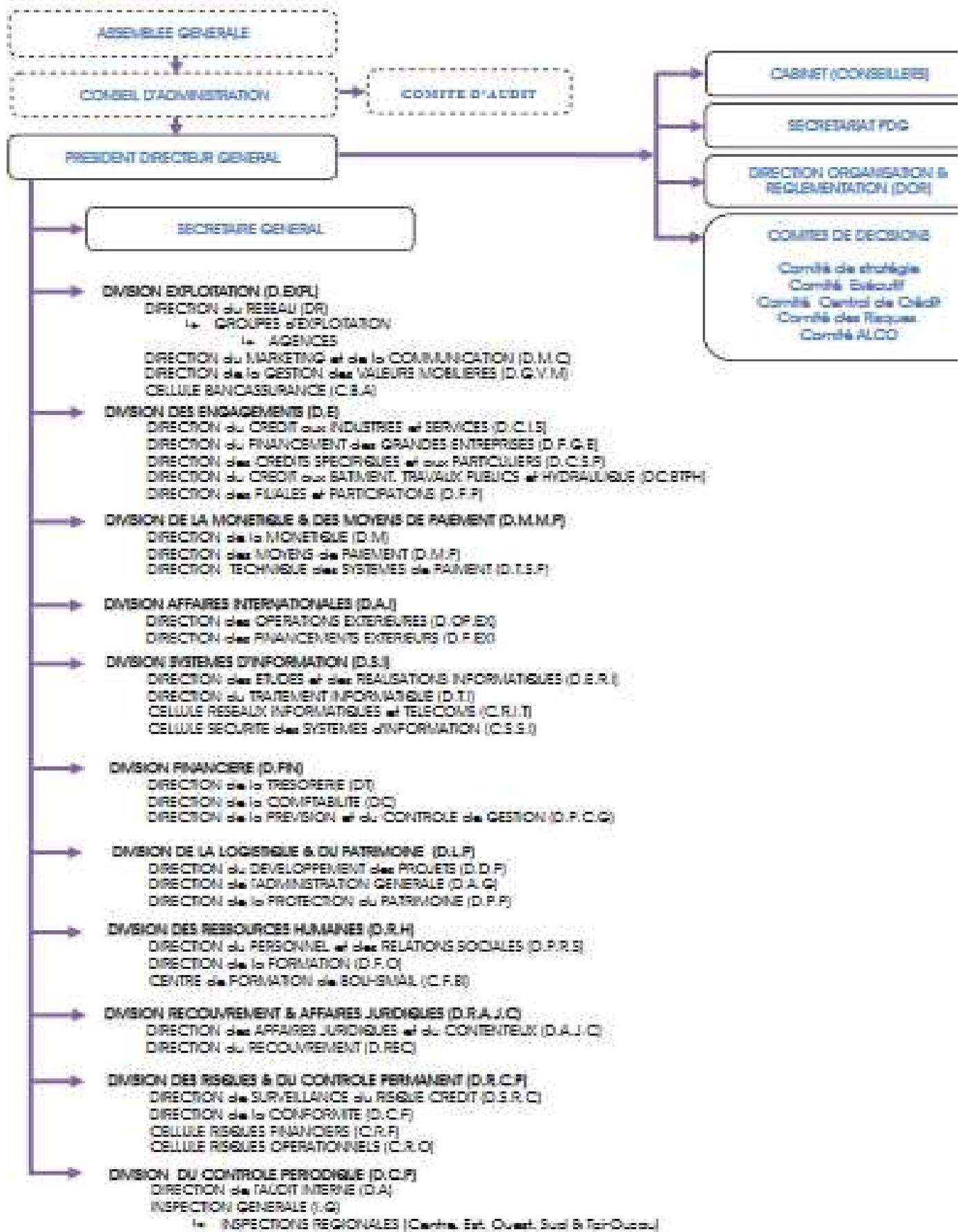


Figure 3: Organisation du CPA

5 La Direction de Surveillance de Risque Crédit –DSRC

Après qu'elle a été rattachée à la DGA des engagements sous l'appellation : « DESE » (Direction des Études et Surveillance des Engagements) ; la DSRC devient supervisée par la Division des Risques et de Contrôle Permanent (DRCP).

L'organisation de la DSRC peut être schématisée selon la figure suivante :



Figure 4: Organisation de la DSRC

➤ Département de la centrale des risques

La centrale des risques est un fichier national établi par la Banque d'Algérie, il contient la situation des engagements de tous les clients envers toutes les banques, la centrale des risques permet d'avoir une information sur tout client concernant sa situation d'impayés vis-à-vis des banques consœurs.

Le département s'occupe de la mise à jours des engagements qui concernent le CPA, il enregistre dans la centrale des risques les entrées et les sorties d'impayés, comme il est habilité à consulter la centrale des risque en ce qui concerne les clients potentiels, et l'enregistrement de nouveaux contractants de crédit.

➤ Département risques et règles prudentielles

Ce département travaille principalement en relation avec la structure chargée de la comptabilité de la Banque ainsi que la direction de traitement informatique. Le département établit le provisionnement des créances en organisant d'une façon trimestrielle un comité de provisionnement sous la direction du PDG.

Section 2 : Elaboration de modèle de scoring

Dans cette section nous allons présenter les étapes qu'on a suivies pour construire notre modèle de crédit scoring.

1 Construction de la base de données :

Nous avons collecter des informations sur 144 entreprises de type PME détenus par des investisseurs privés, qui ont bénéficiés d'un crédit au sein de la banque CPA durant une période qui s'étale sur quatre ans, soit 2013, 2014,2015 et 2016.Ces entreprises sont de différents secteurs d'activités comme le montre l'histogramme suivant:

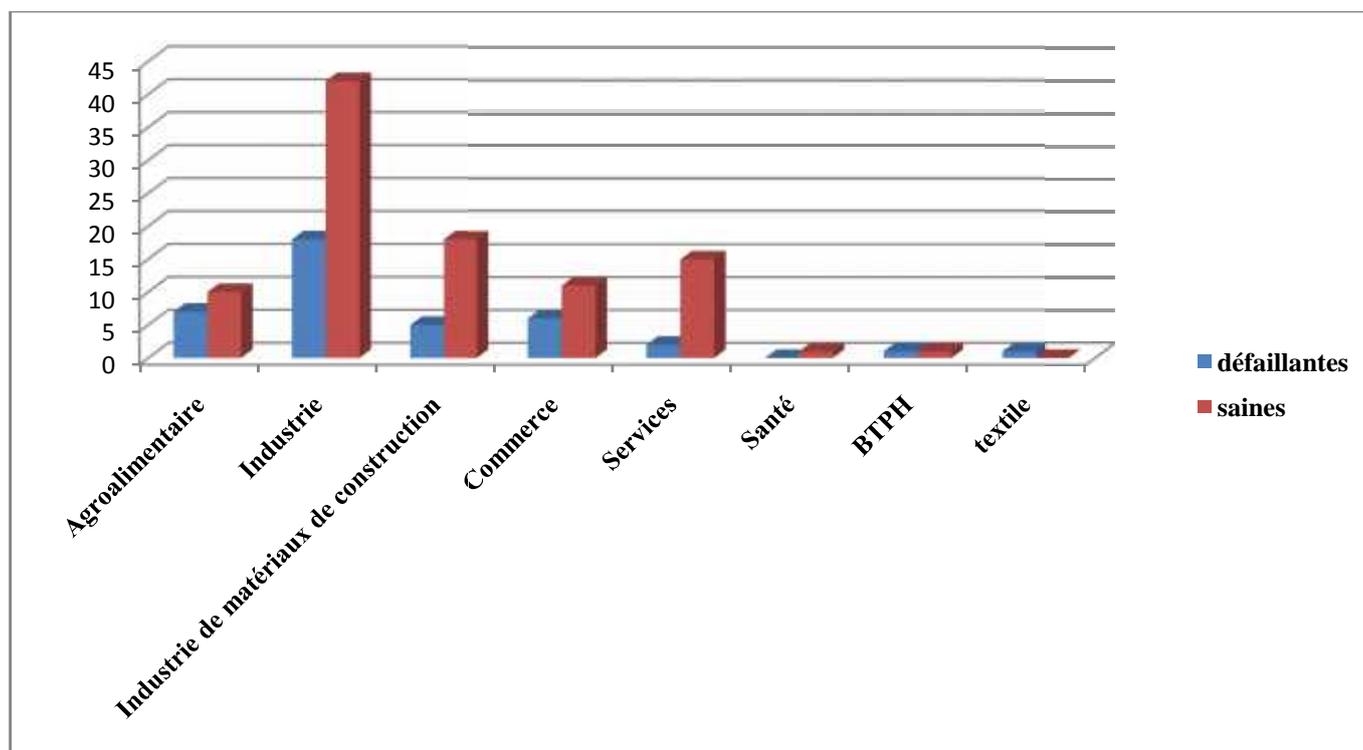


Figure5 : Secteurs d'activité des entreprises de notre échantillon

- D'après cet histogramme on remarque que le CPA octroie plus de crédits aux entreprises qui activent dans le secteur de l'industrie par rapport aux restes des secteurs.
- Les secteurs traditionnels de la banque comme le BTPH et la santé deviennent minoritaires.

Notre échantillon se compose de différents secteurs juridiques. D'après le tableau ci-dessous la principale forme juridique des entreprises auxquelles la banque a octroyé des crédits est du type SARL (62%).

Tableau 5: Les formes juridiques

Forme juridique	Pourcentage
EURL	28%
SARL	62%
SNC	6%
SPA	4%
Affaire	1%
TOTAL	100%

Source : élaboré par nous même.

Avant de présenter la répartition de notre échantillon, il convient de signaler quelques anomalies qui pourront affecter la qualité du modèle :

- La base de données utilisée pour construire notre modèle ne contient que les dossiers acceptés, les dossiers rejetés n'étant pas archivés par la banque et donc non pris en considération ;
- La taille de l'échantillon est relativement limitée ;

Nous avons ainsi une base de données qui contient divers ratios financiers sur 144 entreprises dont 41 entreprises défailtantes et 103 entreprises saines.

2 Critère de défaut :

Nous avons choisi comme critère de défaut l'incapacité de client à rembourser la banque. L'apparition d'impayés est le critère servant à classer une entreprise comme étant défailtante.

La codification des deux groupes est de 0 pour les entreprises en défaut et 1 pour celles qui sont saines, c'est-à-dire celles qui honorent leurs engagements

3 Construction de modèle :

Afin de construire notre modèle, nous allons utiliser des variables comptables qui sont sous formes de ratios financiers les plus utilisés par les analystes : ratios de liquidité, ratios de trésorerie, ratios de rentabilité et ratios de structure financière.

Le traitement de nos données s'effectuera par le logiciel d'analyse statistique IBM SPSS 21⁶⁹ (Statistical package for the social sciences), qui est l'un des logiciels les plus utilisés dans le monde professionnel.

Nous avons utilisé le modèle de la régression logistique pour construire notre modèle .L'utilisation du modèle logistique est justifié par le fait que la variable à expliquer est une variable qualitative dichotomique : défailante ou saine.

Pour la sélection des variables nous avons utilisés la procédure descendante (backward) de la méthode pas à pas (stepwise).Dans cette méthode, le modèle initial comprend toutes les variables, SPSS va retirer la variable ayant la plus faible contribution au modèle. Si la variation du R² n'est pas significative on élimine la variable. La procédure va être répétée jusqu'à ce que toutes les variables conservées contribuent significativement à l'amélioration du R².

Dans notre cas, cette procédure nos a permet de sélectionner sept variables comme le montre le tableau ci dessous :

*Deux ratios de trésorerie: $\frac{F d r t}{C}$ et $\frac{e j t d C (F x 3)}{C}$ et $\frac{V a é n}{d' a i}$

*Un seul ratio de liquidité : $\frac{d é}{a}$

*Trois sont des ratios de structure financière : $\frac{F p}{P n c}$; $\frac{T d}{T a}$ et $\frac{D à l i t i}{C é d' a f i è r e}$

*Un le ratio de chiffres d'affaires confiées : $\frac{M c é n}{C}$

Tableau 6 : Les variables sélectionné⁷⁰

Itération		-2log- vraisemblance
Etape 1	1	124,754
	2	116,398
	3	115,050
	4	114,547
	5	114,031
	6	113,931
	7	113,930
	8	113,930
Etape 17	1	135,126
	2	127,667
	3	125,794
	4	125,383
	5	125,348
	6	125,347
	7	125,347

Source : Traitement SPSS

En se basant sur le tableau de sélection des variables et celui de test de spécification du modèle on peut conclure que :

- ✓ L'indice de significativité globale à l'étape 1 lorsqu'on a intégré toutes les variables était de 58,109, puis a diminué à 46,692 une fois qu'on a exclu les variables inappropriées. Avec 7 degré de liberté, la p-value est inférieure à 5%. Donc on peut conclure que globalement ces sept variables sont significativement associés au défaut.

Tableau 7 : Test de spécification du modèle

		Khi-Chi-deux	ddl	Sig.
Etape 1	Etape	58,109	23	,000
	Bloc	58,109	23	,000
	Modèle	58,109	23	,000
Etape 17 ^a	Etape	-1,374	1	,241
	Bloc	46,692	7	,000
	Modèle	46,692	7	,000

Source : Traitement SPSS

Pour s'assurer de la contribution des variables que le modèle a retenu à la prédiction de défaut, un test sur les coefficients qui estime ces variables s'avère nécessaire. Le tableau suivant montre le résultat de test de Wald sur la significativité des coefficients des variables et l'indicateur statistique l'odds ratio qui nous permet d'avoir une idée claire sur les variables.

Tableau 8 : Résultat de la régression logistique

Variables dans l'équation

	A	E.S.	Wald	Ddl	Sig.	Exp(B)	IC pour Exp(B) 95%	
							Inférieur	Supérieur
Valeur Ajoutée / CA	4,359	1,763	6,117	1	,013	78,211	2,471	2475,531
FondderoulementenjourdeCAFRx360 CA	,000	,000	2,332	1	,127	1,000	1,000	1,001
Disponibilité net /Actif	3,766	1,299	8,400	1	,004	43,213	3,385	551,669
Etape 17 Total dettes / Total Actif	2,032	1,170	3,013	1	,083	7,627	,769	75,610
Fonds propres / Passif non courant	3,914	1,320	8,793	1	,003	50,097	3,770	665,788
Dettes à long terme / CAF	1,353	,436	9,624	1	,002	3,869	1,646	9,095
MVC	2,733	,560	23,824	1	,000	15,379	5,132	46,084
Constante	-6,430	1,950	10,878	1	,001	,002		

Source : Traitement SPSS

D'après ces résultats, toutes les variables sont significatives dans ce modèle à l'exception de deux variables : le ratio de fonds de roulement en jour et le ratio (Total dettes / Total Actif), dont leurs p-value est supérieure à 5%.

On remarque aussi que le ratio : Dettes à long terme/CAF et MVC ont l'erreur standard le plus faible ce qui implique que leurs coefficients sont très prédictifs dans leurs estimation.

On constate aussi que toutes les variables significatives ont une relation positive avec la probabilité de défaut ($A > 0$), sauf la constante qui est négative.

Test de corrélation : D'après le tableau⁷¹ de corrélation, les variables indépendantes sont faiblement corrélées entre eux et fortement corrélés avec la variable dépendante.

- Suivants les résultats obtenus précédemment on peut écrire la fonction score généré par le modèle logistique comme suit :

$$\text{Score} = -6.914 + 2.733\text{MVC} + 1.353 \frac{\text{DLT}}{\text{CAF}} + 3.914 \frac{\text{FP}}{\text{PNC}} + 3.766 \frac{\text{Dispo net}}{\text{Actif}} + 4.359 \frac{\text{VA}}{\text{CA}}$$

- Et la fonction de la probabilité de non défaillance est donnée par cette formule :

$P_i = V(\text{Fonction Score})$

$$= \frac{e^{-6.9 + 2.7 M + 1.3 \frac{D}{C} + 3.9 \frac{F}{P} + 3.7 \frac{D}{A} + 4.3 \frac{V}{C}}}{1 + e^{-6.9 + 2.7 M + 1.3 \frac{D}{C} + 3.9 \frac{F}{P} + 3.7 \frac{D}{A} + 4.3 \frac{V}{C}}}$$

Analyse de la cohérence des signes des coefficients trouvés et vérification de leur conformité avec la logique financière :

1- Valeur ajoutée / Chiffres d'affaires : La valeur ajoutée est utilisée comme un indicateur de la rentabilité économique de l'entreprise.

Plus ce taux est élevé, plus l'entreprise contribue à créer de la valeur et donc à être rentable et à s'éloigner de risque de défaillance puisque elle pourra s'autofinancer et s'éloigner de l'endettement. C'est ce qui explique la relation positive avec le score.

⁷¹ Présenté dans l'annexe 2

2- DLMT/CAF: La relation de score et le taux d'endettement (DLMT/CAF) est positive, contrairement à l'hypothèse émise. Mais ce résultat est conforme avec les données de la banque car la moyenne de ce ratio pour les entreprises saines est de 41,42% supérieure à celle des entreprises défaillantes qui est de 28,54%. Une exploration plus poussée des données pourrait faire ressortir l'explication de cette relation contre-intuitive.

3- Disponibilité / Actif : représente un ratio de liquidité qui étudie la solvabilité de l'entreprise à court terme. L'entreprise doit avoir une liquidité suffisante pour honorer ces échéances et aussi optimiser la rentabilité des fonds. Dans le cas contraire elle risque de faire face au risque d'illiquidité et donc elle sera insolvable et incapable de faire face à ces engagements financiers.

4- Fonds propres / Passif non courant : Les fonds propres représentent les ressources financières que possède une entreprise. Ils reflètent sa santé financière c'est pour cela qu'il est très important de savoir comment les gérer. Ce ratio indique la solvabilité de l'entreprise et sa capacité à payer ses dettes à long terme ou non. Lorsque les fonds propres sont faibles, l'entreprise aura des difficultés à rembourser ces dettes financières, et dans le cas contraire si elle possède des fonds propres stables ou élevés elle n'aura pas des difficultés à faire face à ses engagements et sa situation financière sera bonne et aura un risque de défaillance faible. C'est ce qui explique la relation positive de ce ratio avec le score.

5- Mouvement de chiffres d'affaires confiés / Chiffres d'affaires : ce ratio représente la part de chiffres d'affaires que le client a confié à sa banque par rapport au total de son chiffre d'affaires. Il indique le niveau de fidélité que l'entreprise a envers cette banque. La relation positive entre le score et ce ratio reflète la bonne relation qui relie l'entreprise et la banque et est un indicateur de faibles risque non remboursement.

4 Validation :

a. Test de Significativité globale :

Tableau 9 : Tableau récapitulatif

Etape	-2log-vraisemblance	R-deux de Cox & Snell	R-deux de Nagelkerke
1	113,930 ^a	,332	,476
17	125,347 ^b	,277	,397

Source : Traitement SPSS

- ✓ on peut mieux apprécier la qualité de cette régression par un indicateur statistique qui est le coefficient de détermination R^2 .
- ✓ SPSS nous donne deux versions de ce coefficient qui sont produites par deux auteurs différents mais qui sont généralement proches l'une de l'autre.

✓ D'après le tableau récapitulatif on remarque que l'indicateur de Cox et Snell et celui de Nagelkerke sont de 27.7% et 39.7 respectivement, ce qui veut dire que 27.7% et 39.7% de la variable de probabilité de défaut est attribuable aux cinq variables que le modèle a retenu.

b. Test de Hosmer-Lemeshow :

Tableau 10 : Défaillances observées et prédites

Tableau de contingence pour le test de Hosmer-Lemeshow

	Défaut = défaut		Défaut = saine		Total
	Observations	Attendu	Observations	Attendu	
Etape 1	1	14	13,908	0,092	14
	2	14	13,584	0,416	14
	3	13	13,065	1,935	14
	4	12	12,499	2,501	14
	5	13	12,174	1,826	14
	6	11	11,544	3,456	14
	7	12	9,984	2,016	14
	8	6	7,807	8,193	14
	9	3	5,329	11,671	14
	10	5	3,107	13,893	18
Etape 17	1	14	13,695	0,305	14
	2	14	13,252	0,748	14
	3	11	12,781	3,219	14
	4	13	12,369	1,631	14
	5	11	11,853	3,147	14
	6	14	11,169	0,831	14
	7	11	10,016	3,984	14
	8	5	8,252	9,748	14
	9	4	5,550	10,450	14
	10	6	4,063	12,937	18

Source : Traitement de SPSS

Tableau 11 : Test de Hosmer-Lemeshow

Test de Hosmer-Lemeshow

Etape	Khi-Chi-deux	Ddl	Sig.
1	6,690	8	,570
17	13,547	8	,094

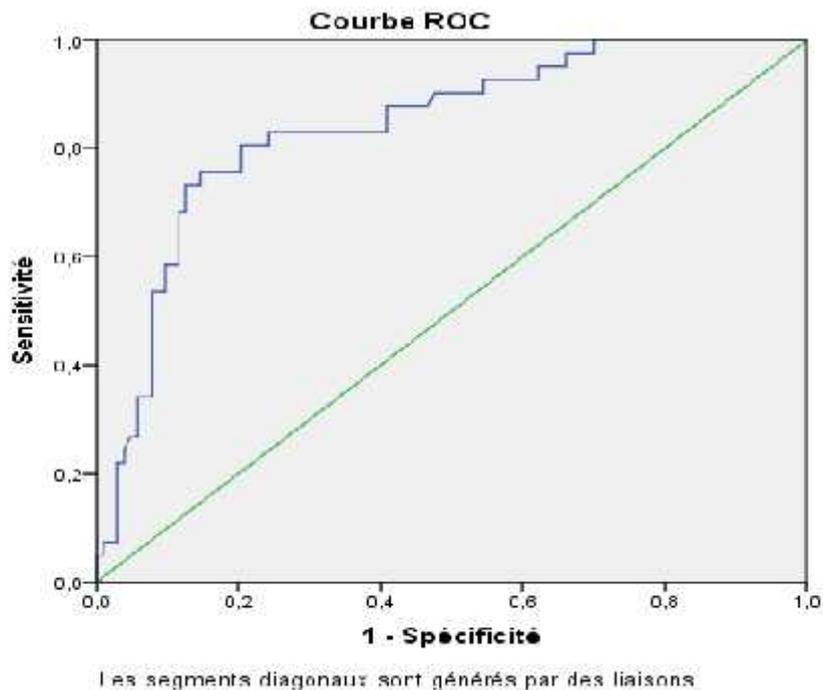
Source : Traitement SPSS

- ✓ Ce test permet de juger la qualité de l'ajustement du modèle. Il évalue s'il y a une différence significative entre les valeurs observées et les valeurs prédites. Évidemment, on préfère que ce test soit non significatif.
- ✓ Au seuil de 5%, l'ajustement du modèle est bon car la probabilité (ou significativité) du Khi carré à 8 degrés de liberté (ddl) vaut 0,094, soit 9.4%(>5%), aussi la distance entre les fréquences observées et prédites au moyen de la statistique du Khi 2 est petite (13,547). Le modèle est donc bien calibré.

5 Performance du modèle:

a. La courbe ROC: (Receiver Operating Characteristic):

Figure 6 : La courbe ROC de la régression logistique



Source : Traitement SPSS

Tableau 12 : Surface sous la courbe ROC

Zone sous la courbe

Variable(s) de résultats tests: Probabilité de prédiction

Zone	Erreur Std. ^a	Signif. asymptotique	Intervalle de confiance 95% asymptotique	
			Borne inférieure	Borne supérieure

,839	,036	,000	,768	,911
------	------	------	------	------

La ou les variables de résultats tests : Probabilité de prédiction comportent au moins une liaison entre le groupe d'état réel positif et le groupe d'état réel négatif. Les statistiques peuvent être déformées.

- a. Dans l'hypothèse non-paramétrique
- b. Hypothèse nulle /: zone vraie = 0.5

Source : Traitement SPSS

- Sensibilité : taux de vrai positif, c'est la probabilité pour laquelle on décide que l'entreprise est positive (saine) sachant qu'en réalité est positif.
- Spécifié : taux de vrai négatif, c'est-à-dire le taux pour qu'une entreprise soit négative sachant qu'elle est négative (défaillante).
- La diagonale : droite d'un modèle de classification aléatoire.

L'estimation de la surface sous la courbe ROC de la probabilité prédite donnée par l'indice de performance est égale à 0.839 qui est proche de 1 ce qui signifie que c'est un bon modèle, et puisque on travaille sur un échantillon, la valeur de l'indice se trouve à 95% dans cet intervalle de confiance : [0,768 ; 0,911].

Aussi le p-value est inférieure à 5%. Donc on rejette l'hypothèse de nullité qui dit que le modèle de la régression logistique est semblable à celui de modèle aléatoire.

La valeur de L'erreur est très faible (0.036).

- ✓ Donc d'après ces résultats on peut dire que notre modèle est performant.

b. Le taux de bon classement :

Tableau 13: Tableau de classement :

Observations		Prévisions		
		Défaut		Pourcentage correct
		défaut	saine	
Etape 1	Défaut	26	15	63,4
	saine	9	94	91,3
	Pourcentage global			83,3
Etape 17	Défaut	23	13	56,1
	saine	10	93	90,3
	Pourcentage global			80,6

*La valeur de césure est de 0.5

Source : Traitement de SPSS

- ✓ Le modèle nul a eu un taux de précision de 71,5 %, ce qui donne un taux d'erreur de 28,5%. Une fois que la régression est faite le taux atteint 80,6 % (taux d'erreur est 19,4%). Donc le taux de bon classement s'est amélioré.
- ✓ Le taux prédit égal (bien classées) à : $(23+93)/144 = 116/144 = 80.6 \%$
- ✓ Le taux d'erreur est égal à : $(18+10)/144 = 19,4\%$
- ✓ Le taux de bonne capacité prédictive des entreprises défaillantes = $23 / (23+18) = 56.1\%$
- ✓ Le taux de bonne capacité prédictive des entreprises saines = $93 / (93+10) = 90.3\%$
- ✓ Le taux de bonne classification des entreprises saines = $93 / (93+18) = 83.8\%$
- ✓ Le taux de bonne classification des entreprises défaillantes = $23 / (23+10) = 69.7\%$

Le modèle construit classe 116 entreprises parmi 144 dans leurs classes d'origine avec seulement un taux d'erreur de 19.4 %, donc le modèle est performant.

Conclusion

Afin de déterminer les indicateurs de défaillance des entreprises ayant demandées un crédit au CPA en utilisant la méthode de scoring, nous avons élaboré un modèle de prévision et de prédiction de la défaillance en utilisant la régression logistique. Les résultats que nous avons obtenus sont relativement satisfaisants puisque le modèle nous a permis de déterminer cinq variables pertinentes qui distinguent entre les deux groupes d'entreprises à savoir les saines et les défaillantes. Le signe des coefficients des variables sont positifs ce qui est logique pour les quatre variables : (VA/CA), (DISPO/Actif), (FP/PNC), (MVC/CA). Mais ce n'est pas le cas pour la variable (DMLT/CAF) car ce signe positif avec le score est contre-intuitif. Une réflexion plus poussée est nécessaire pour clarifier cette relation.

L'élaboration d'un modèle de prévision par la méthode de scoring permet aux établissements de crédit d'avoir un outil d'aide aux traitements des demandes de crédit, ainsi qu'à la prise de décision d'octroi ou de refus de la demande.

Conclusion générale

La banque est souvent confrontée à une multitude de risques du fait qu'elle octroie des crédits, notamment le risque de non-remboursement qui est au cœur des préoccupations de tout prêteur.

La mise en place d'un système de prévention performant est indispensable aux banques afin de minimiser et prévoir ces risques, qui peuvent avoir des conséquences néfastes sur sa situation financière.

La gestion des risques que subissent les banques se résumait auparavant aux méthodes traditionnelles (diagnostic financier). Après la crise des Subprimes qui a mis en péril toute l'économie mondiale, les établissements de crédits ont essayé de renforcer leurs systèmes de prévention des risques en adoptant des méthodes plus performantes.

Il est primordial pour les établissements de crédit d'effectuer une étude objective avant de prendre la décision d'octroi ou de refus des demandes de crédit pour se prévenir des demandeurs défaillants. C'est pour cela qu'ils doivent se doter des outils d'aide à la décision qui permettent d'identifier les indicateurs de défaillance.

Le crédit scoring qui est une méthode d'évaluation de risque de crédit consiste en l'utilisation des données historiques et de techniques statistiques, dans le but de faire apparaître la contribution de certaines variables dans le critère de «défaillance».

L'objectif de notre travail de recherche est de développer un modèle de Crédit Scoring permettant de prédire la probabilité de défaut des nouveaux demandeurs de crédit par l'identification des critères de défaillance des entreprises ayant bénéficiées déjà d'un crédit au sein de CPA. Pour ce faire, nous avons utilisés un outil statistique qui est la régression logistique et un échantillon de 144 entreprises ayant obtenu un crédit du CPA.

Dans un premier lieu, il ressort de la régression logistique que nous avons effectuée, que la valeur ajoutée de l'entreprise et le montant de ses disponibilités et de ses fonds propres, ainsi que le mouvement d'affaires confiées, et le montant des dettes à long terme sont des indicateurs pertinents pour détecter si une entreprise est défaillante ou bien saine.

Dans un second lieu, nous avons constaté qu'en utilisant cette méthode statistique, nous pouvons aboutir à des résultats excellents quant à la prévision des risques des emprunteurs caractérisés par un taux de bon classement de 86,5%. Ce modèle nous a permis de relever les indicateurs de défaillance suivants:

) Valeur ajoutée / Chiffres d'affaires : Plus ce taux est élevé, plus l'entreprise contribue à créer de la valeur et donc à être rentable et à s'éloigner de risque de défaillance.

-) Disponibilité / Actif :L'entreprise doit avoir une liquidité suffisante pour honorer ces échéances et aussi optimiser la rentabilité des fonds.
-) Fonds propres / Passif non courant :Ce ratio indique la solvabilité de l'entreprise et sa capacité à payer ses dettes à long terme ou non. Quand l'entreprise possède des fonds propres stables ou élevés, elle réduit la probabilité d'avoir des difficultés à faire face à ses engagements et sa situation financière sera bonne et aura un risque de défaillance faible.
 -) Mouvement de chiffres d'affaires confies / Chiffres d'affaires : Ce ratio indique le niveau de fidélité que l'entreprise a envers sa banque. La bonne relation qui relie l'entreprise et la banque est un indicateur de faible risque de non remboursement.
 -) DMLT/CAF :La relation entre ce ratio et la non défaillance d'une entreprise est positive contrairement à l'hypothèse que nous avons émise.

En dépit de sa performance, notre modèle peut être amélioré, ceci par l'obtention d'un échantillon plus large. De plus, il est impératif d'intégrer d'autres variables pour améliorer son pouvoir prédictif, pour qu'il puisse ainsi venir en appui aux décisions des chargés de crédit.

En effet, l'approche du crédit scoring est certes très profitable pour la banque dans le sens où elle lui permet de perfectionner son processus de gestion des risques, mais elle est d'autant plus profitable à l'économie nationale, parce qu'elle permet d'instaurer une transparence et une meilleure sélectivité en terme d'octroi de crédit; deux conditions nécessaires au développement de toute économie.

ANNEXES

ANNEXE 1

Base de données des variables significatifs :

entreprises	défaut	MVC	Valeur Ajoutée / CA	Fond de roulement en jour de CA (FR x 360)/CA	CAF / CA	Total dettes / Total Actif	Fonds propres / Passif non courant	Dettes long terme / CAF
1	0	1	0,4262	-146,7238	-0,0247	0,7192	0,3595	0,4134
2	0	1	0,2135	133,8605	0,0238	0,0818	1,0000	0,0000
3	1	0	0,1682	97,8017	0,0944	0,8927	0,1631	2,5202
4	1	0	0,1631	82,5798	-0,0259	0,5862	0,8790	0,2234
5	0	1	0,1827	72,6963	-0,0022	0,6466	0,6479	0,4911
6	0	1	0,6444	273,5414	-0,2196	0,3160	0,7470	0,2360
7	1	0	0,2708	35,2713	-0,0099	0,6052	0,8465	0,1135
8	1	1	0,2079	-5,1634	0,0201	0,7984	0,4131	0,8181
10	0	1	0,2468	104,6891	-0,0999	0,4680	1,0000	0,0000
11	0	1	0,4981	-102,8476	0,5743	0,8459	0,4299	-2,6851
12	1	1	0,3785	86,1760	0,0270	0,4903	0,8785	0,1407
13	0	1	0,7975	-158,5490	-0,0108	0,7374	0,3338	0,7513
14	0	1	0,3825	-30,8306	-0,0185	0,5837	0,6988	0,4940
15	0	0	0,2689	258,6042	-0,1605	0,7358	0,3930	1,6258
16	1	1	0,5325	146,6969	-0,1097	0,4278	0,6182	0,3838
17	0	0	0,2460	-14,5297	-0,0573	0,6164	0,5846	0,8217
18	0	0	0,0863	60,1901	0,0369	0,5451	0,4607	2,6058
19	1	1	0,2680	428,8066	-0,0897	0,4102	1,0000	0,0000
20	0	1	0,3568	152,2078	-0,0597	0,5338	1,0000	0,0000
21	0	1	0,5298	-150,1590	0,0208	0,4006	0,7467	0,3481
22	0	1	0,1986	-21,7187	-0,0402	0,9961	1,0000	0,0000
23	1	0	0,5971	-169,5324	0,0011	0,7887	0,3514	0,5528
24	0	1	0,1797	-260,6204	0,0797	0,9869	0,0236	4,7448
25	1	1	0,0323	145,4816	-0,0309	0,8389	0,3509	1,1853
26	1	0	0,4590	-545,0771	-0,1976	0,8037	0,4850	0,6347
28	0	1	0,6299	120,5815	0,1825	0,1687	0,8651	0,0984
29	1	0	0,1954	178,5870	0,3370	0,3464	1,0000	0,0000
30	0	1	0,0707	-2580,5539	-0,8409	0,4988	1,0000	0,0000
31	0	1	0,4940	35,4728	-0,0317	0,2763	1,0000	0,0000
32	0	1	0,3928	48,6620	-0,0496	0,4502	0,7230	0,2905
33	0	0	0,3938	199,6131	-0,0295	0,8472	0,5213	1,7716
34	1	0	0,1317	-58,5186	-0,0146	0,6944	0,8682	0,1036
35	0	1	0,0648	-7,7299	-0,1110	0,6950	1,0000	0,0000
36	0	0	0,4175	523,0422	-0,1789	0,3938	0,8969	1,5587

38	1	0	0,0945	122,4929	0,0342	0,2904	0,9040	0,3032
40	0	1	0,3997	-269,3219	0,0675	0,8730	1,0000	0,0000
41	0	0	0,3040	-196,4739	0,0055	0,8040	0,6342	0,2036
42	1	0	0,4983	-1021,5490	-0,0956	0,9521	0,3604	0,4602
43	0	1	0,2791	75,2285	0,0444	0,6187	0,7620	0,1535
44	0	0	0,5071	18,9875	-0,2731	0,3684	1,0000	0,0000
45	1	0	0,1480	-90,6801	-0,1450	0,8650	0,9136	0,0691
46	0	1	0,1482	50,1406	0,0042	0,2263	1,0000	0,0000
48	1	0	0,2200	-81497,7665	0,1467	1,2195	2,2580	1,0702
49	1	1	0,1625	-50,1947	-0,0855	1,0440	-0,0773	0,9560
50	0	1	0,6113	-769,8774	-0,0614	0,5169	0,9610	0,0311
51	0	1	0,0759	-111,0361	0,0329	0,5620	0,8067	0,7512
52	0	1	0,1387	-55,9631	-0,1745	0,5598	0,9740	0,0338
53	0	1	0,2688	-108,9015	0,0245	0,7496	0,4968	0,6850
55	1	1	0,5700	-14534,0995	4,9194	0,8505	0,6799	-0,1108
56	0	0	0,4697	202,1603	-0,0293	0,0301	1,0000	0,0000
57	1	1	0,1085	-60,5251	-0,1494	0,5011	0,9112	0,0607
58	0	1	0,1339	-66,2368	0,0316	0,4238	0,9467	0,0415
59	0	1	0,1808	77,1781	0,0246	0,5751	0,9097	0,5231
60	0	1	0,1417	31,0775	-0,0191	0,5358	0,9494	0,0309
61	0	1	0,0866	78,1929	-0,0393	0,3730	1,0000	0,0000
62	0	1	0,2179	-181,7843	0,0007	0,8977	0,4248	0,2703
63	0	1	0,4713	195,3724	-0,0120	0,0709	0,9723	0,0331
64	0	0	0,1664	74,0059	0,1097	0,8669	0,6495	0,0791
65	0	1	0,3000	-137,5517	-0,0011	0,6231	1,0000	0,0000
66	0	1	0,3318	202,6012	-0,0440	0,2622	0,8818	0,1806
71	1	0	0,1882	-32,0872	0,0988	0,6017	0,7294	0,1302
72	0	1	0,1917	53,2846	0,2212	0,3039	1,0000	0,0000
73	1	1	0,2090	13,4807	0,0557	0,5918	0,7947	0,3767
74	0	1	0,1302	77,2443	-0,0119	0,7667	0,3339	0,8093
75	1	0	0,3556	107,1096	0,3204	0,3920	1,0000	0,0000
77	0	1	0,2533	161,6355	0,0022	0,0608	0,9926	0,0104
78	0	1	0,1662	98,3211	0,0004	0,1887	0,9963	0,0064
79	0	1	0,3141	-44,5283	-0,0948	0,5623	0,8591	0,1063
80	0	1	0,4894	32,1624	-0,0761	0,9678	0,0538	1,9645
81	0	1	0,1580	-207,1884	-0,1171	0,8605	1,0000	0,0000
82	0	1	0,4457	18,8084	0,1327	0,4162	0,7951	0,0947
83	1	0	0,2707	-124,6137	-0,0910	0,6056	1,0000	0,0000
84	0	0	0,3909	827,3956	-0,0104	0,7914	0,2150	1,8826
85	0	0	0,4853	62,0901	0,0836	0,9017	0,3701	1,8541
86	0	0	0,1378	91,1131	-0,0281	0,6676	0,5432	0,3246
87	1	1	0,1644	-20,3000	-0,1654	0,2757	0,9247	0,4277
88	0	1	0,3725	43,1672	0,0556	0,4207	0,7859	0,1828
90	0	1	0,1844	144,6821	-0,0530	0,0954	1,0000	0,0000
91	1	0	0,3560	170,0552	0,0841	0,0949	1,0000	0,0000

93	0	1	0,2981	-55,2057	0,0002	0,6400	0,6699	0,2791
94	0	1	0,2542	-9,0506	-0,1000	0,8243	1,0000	0,0000
95	1	1	0,3637	164,2972	-0,0385	0,1766	0,2625	0,0026
97	1	0	0,2337	110,0496	-0,1007	0,5455	0,7135	0,4343
98	1	0	0,3149	620,6852	-0,0237	0,4008	0,7833	0,2447
99	0	1	0,1834	138,2734	-0,0447	0,3311	0,8122	0,2381
100	1	0	0,1082	376,1317	0,0630	0,4099	1,0000	0,0000
101	0	1	0,5144	-48,6478	-0,2344	0,4635	0,7926	0,0905
102	0	1	0,1128	38,4601	0,0325	0,6296	1,0000	0,0000
103	0	0	0,3404	54,3263	-0,2926	0,3333	0,6751	1,7956
104	0	0	0,1685	-52,4954	0,0424	0,6398	0,8374	0,1469
105	0	1	0,1020	-40,3562	-0,0636	0,8474	0,8590	0,1338
108	0	1	0,7975	-158,5490	-0,0108	0,7374	0,3338	0,7513
109	1	1	0,2688	-108,9015	0,0245	0,7496	0,4968	0,6850
112	0	1	0,0935	-56,2985	0,0608	0,5999	0,7381	1,2097
113	0	0	0,1468	10,2017	-0,0088	0,4945	0,8768	0,6166
114	1	1	0,1348	-17,1169	-0,0297	0,8570	0,9392	-0,0287
117	0	0	0,1999	-65,3870	-0,1708	0,6957	1,0000	0,0000
118	0	1	0,4872	-170,1648	-0,2965	0,5796	0,8090	0,2262
120	0	1	0,1664	74,0059	0,1097	0,8669	0,6495	0,0791
121	1	0	0,1664	74,0059	0,1097	0,8669	0,6495	0,0791
122	0	1	0,3125	-108,5504	-0,0077	0,7208	0,6014	0,1948
123	0	0	0,1056	20,9089	0,0596	0,8142	0,9919	0,0197
124	1	0	0,1586	172,9616	-0,1234	0,3627	0,9232	0,2338
125	1	1	0,0701	2,2115	-0,0038	0,9162	1,0000	0,0000
126	0	0	0,7540	-14,9195	0,0238	0,3905	0,8319	0,0702
127	0	1	0,2323	94,4445	0,1262	0,6054	1,0000	0,0000
128	0	0	0,0973	78,7350	0,0030	0,2131	1,0000	0,0000
129	0	1	0,2599	159,8192	-0,0104	0,3889	0,8717	0,9009
130	0	1	0,0786	26,8579	-0,0265	0,8760	0,9153	0,1942
131	0	0	0,2299	127,7677	0,0870	0,3402	0,9793	0,0249
133	1	0	0,2493	-666,5317	-1,1590	0,9064	0,8584	0,0438
134	0	1	0,2479	87,8555	0,1147	0,6898	0,9250	0,0755
136	0	0	0,0441	14,7023	0,0820	0,9402	0,6835	7,5283
137	0	1	0,5872	-287,0927	0,5208	0,9769	1,0000	0,0000
138	1	1	0,1151	26,7619	0,0312	0,9207	0,7481	0,1455
139	0	1	0,1450	118,4094	0,0234	0,5782	1,0000	0,0000
142	0	1	0,0365	0,5451	-0,0216	0,8066	1,0000	0,0000
143	0	1	0,5380	407,1433	0,1867	0,3899	1,0000	0,0000
147	0	1	0,1166	48,9830	-0,0898	0,5395	1,0000	0,0000
148	0	1	0,0630	-4,1573	0,1077	0,8161	1,0000	0,0000
149	0	1	0,4494	272,1007	-0,4862	0,5173	0,9259	0,1188
150	0	1	0,1620	-21,5835	0,0949	0,9250	1,0000	0,0000
151	0	1	0,0841	22,1238	-0,0750	0,9573	0,8147	0,1184
152	0	1	0,0964	28,4181	0,0646	0,8459	1,0000	0,0000

Annexe 2 :

Tableau de classement avant la régression logistique

Tableau de classement^{a,b}

	Observations	Prévisions		
		Défaut		Pourcentage correct
		défaut	saine	
Etape 0	Défaut	0	41	,0
	saine	0	103	100,0
Pourcentage global				71,5

a. La constante est incluse dans le modèle.

b. La valeur de césure est ,500

Annexe 3 : Tableau de corrélation des variables significatifs

Corrélations

	Défaut	Valeur Ajoutée / CA	Fondderoulement en jour de CAFRx360CA	Disponibilité net / Actif	Total dettes / Total Actif	Fonds propres / Passif en cours	Dettes à long terme / CAF	MVC	
Défaut	Corrélation de Pearson	1	,116	,154	,110	-,070	,034	,333*	
	Sig. (bilatérale)		,168	,066	,189	,401	,685	,000	
	N	144	144	144	144	144	144	144	
Valeur Ajoutée / CA	Corrélation de Pearson	,116	1	,001	,065	-,146	-,200*	-,064	,013
	Sig. (bilatérale)	,168		,991	,442	,080	,017	,448	,879
	N	144	144	144	144	144	144	144	144
Fondderoulement en jour de CAFRx360CA	Corrélation de Pearson	,154	,001	1	,011	-,229**	-,416**	-,051	,099
	Sig. (bilatérale)	,066	,991		,896	,006	,000	,540	,236
	N	144	144	144	144	144	144	144	144
Disponibilité net / Actif	Corrélation de Pearson	,110	,065	,011	1	-,243**	,106	-,072	-,335*
	Sig. (bilatérale)								

	Sig. (bilatérale)	,189	,442	,896		,003	,206	,388	,000
	N	144	144	144	144	144	144	144	144
	Corrélation de Pearson	-,070	-,146	-,229**	-,243**	1	-,328**	,255**	,011
TotaldettesTotal / Actif	Sig. (bilatérale)	,401	,080	,006	,003		,000	,002	,896
	N	144	144	144	144	144	144	144	144
	Corrélation de Pearson	,034	-,200*	-,416**	,106	-,328**	1	-,411**	,021
Fondspropres / Passifnoncourant	Sig. (bilatérale)	,685	,017	,000	,206	,000		,000	,800
	N	144	144	144	144	144	144	144	144
	Corrélation de Pearson	,063	-,064	-,051	-,072	,255**	-,411**	1	-,226*
Detteàlongterme / CAF	Sig. (bilatérale)	,455	,448	,540	,388	,002	,000		,007
	N	144	144	144	144	144	144	144	144
	Corrélation de Pearson	,333*	,013	,099	-,335**	,011	,021	-,226**	1
MVC	Sig. (bilatérale)	,000	,879	,236	,000	,896	,800	,007	
	N	144	144	144	144	144	144	144	144

** . La corrélation est significative au niveau 0.01 (bilatéral).

* . La corrélation est significative au niveau 0.05 (bilatéral).

Bibliographie

Ouvrages

Ceresoli Michèle et Guillaud Michel, « Titrisation : Gestion financière de la banque. », édition ESKA, paris, 1992.

Colasse B, l'analyse financière de l'entreprise, Ed. La Découverte, Paris, 2003.

Coussergues de Sylvie, Gestion de la banque: Normes et réglementation à jour, nouvelles stratégies bancaires 8^{ème} édition, Edition Dunod, 2017.

Hosmer David .W, Jr. Stanley Lemeshow, Rodney X. Sturdivant, Applied Logistic Regression, éditeur: John Wiley and Sons, cop. 2013

Naulleau Gérard et Rouach Michel, Le contrôle de gestion et financier, Revue bancaire.

Petit-Dutaillis Charles, Le risque de crédit bancaire, édition scientifique Ribière, Paris, 1967.

Scialom.L L, Economie bancaire, collection repère économique, Ed. La Découverte, Paris, 2013.

Traiss Pierre, Méthodologie statistique pour la discrimination et le classement, Application au ciblage des interventions nutritionnelles.

Articles et revus

Bensaida Salma et Lotfi Said, Measuring the impact of IFRS 9 on Bank credit risk, revue du contrôle de la comptabilité et de l'audit, Maroc, Juin 2018.

Berwick Viv, Cheek Liz et Ball Jonathan ,Statistics review14 : Logistic regression.

Cohen Benjamin et Gerald Edwards jr, la nouvelle ère des provisions pour pertes de crédit attendues, rapport trimestriel BRI, mars 2017.

Criteria: A review of the literature "Intelligent Systems in Accounting, Finance & Management".

Dionne George, Gestion des risques : Histoire, Définition et critiques, support de cours au centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, janvier 2013.

Dumontier Pet D.Dupre, Pilotage bancaire : les normes IAS et la réglementation Bale II, Ed. Revue banque, paris.

Gillet Adeline, Brostaux Yves & Palm Rodolphe, « Principaux modèles utilisés en régression logistique».

Gup.B. E & Kolari.J.W, Commercial banking: The management of risk.

Ouennoughi Sid-Ali, Modalités de détection des risques de défaillance au sein des PME, Essai d'analyse empirique.

Pointon & Abdou.H, "Credit scoring, statistical techniques and evaluation"

Revue de l'assurance N°24, janvier à mars 2019, éditée par le conseil national des assurances.

Szumilas.M, Explaining Odds Ratios.

Touati-Tlib Mohamed et Gliz Abdelkader, « Estimation du risque de crédit et qualité de l'information comptable en Algérie »

Ziaul Mohammad, Chowdhury Islam and Tanyir, "Variable selection strategies and its importance in clinical prediction modeling"

Mémoires et thèses de fin d'études

Adli Rahma et Habbi Hassina, « Application de la réglementation prudentielle dans le Système bancaire algérien », mémoire de master en sciences économique, FSESG, UMMTO.

Azzaoui Khaled, « Le financement des PME et le mécanisme de garantie en Algérie », Thèse de doctorat ,2016/2017, ESC.

Bennane Abderrazak, Traitement des valeurs manquantes pour l'application de l'analyse logistique des données à la maintenance conditionnelle, Département de mathématique et de génie industriel, Ecole Polytechnique de Montréal.

Boussoul Lamia « la notation interne : nouvel outil de gestion du risque de crédit », mémoire de fin d'étude, Ecole supérieure de banque.

Guizani Asma, Thèse de doctorat, Ecole doctorale Sousse en science de gestion, Laboratoire interdisciplinaire de recherche en sciences de l'action (paris),

Iscanoglu Aysegul, "Credit scoring methods and accuracy ratio", These de master, The institute of applied mathematics of the Middle East technical university.

Komorad Karel, "Credit Scoring Estimation", Mémoire de fin d'étude master, Institute for Statistics and Econometrics Humboldt University Spandauer Str.1D-10178 Berlin.

Rey .N, Les dérivés de crédit : instrument de couverture et facteurs d'instabilité. L'exemple des « crédits default swap. », Atelier dans le cadre de l'axe thématique « intermédiation et systèmes financiers. » du GDRE n°335, monnaie, banque et finance du CNRS, novembre 2007, université Paris 12.

Sites internet

<http://www.jybaudot.fr>,

<https://od-datamining.com>

<http://www.mvtechnologie.com>

<https://www.xlstat.com>

<http://spss.espaceweb.usherbrooke.ca/>

<http://theses.univ-lyon2.fr>

<https://detectionfaillite.fr>

<https://od-datamining.com>

<https://wikimemoires.net>

<https://www.researchgate.net>

<http://www.pandat.fr>

<https://www.academia.edu>

www.lafinancepourtous.com

<http://www.dge.gov.dz>

Résumé

Le risque de crédit est considéré comme le risque principal que subissent les établissements de crédit. Il représente la perte potentielle supportée par un établissement de crédit suite à une modification de la qualité de crédit d'une contrepartie ou d'un ensemble de contreparties.

Selon l'accord de Bale, il existe plusieurs méthodes de gestion et appréciation de risque de crédit : l'analyse financière, les systèmes experts, les réseaux neurones et le modèle de score qui est le thème de notre mémoire.

Ce dernier nécessite une méthodologie à respecter pour le mettre en place, telle que la conception de l'échantillon, l'élaboration du modèle et sa validation.

C'est dans ce cadre-là que notre travail a été conçu et il se résume dans la problématique suivante :

Comment les banques peuvent-elles évaluer d'une manière efficace le risque crédit ? Et quelle démarche à suivre pour mettre en place un système de score performant ?

Pour répondre à cette problématique, nous présentons dans un premier temps le risque de crédit, ses types et son impact sur l'activité bancaire ainsi que la réglementation internationale qui régit l'activité bancaire. Nous avons par la suite élaboré un modèle de credit scoring en se basant sur le modèle de régression logistique binaire. Pour son estimation, nous avons utilisé un échantillon de 144 entreprises ayant obtenu un crédit auprès du CPA. Nous avons ainsi fait ressortir cinq variables significatives pour expliquer la probabilité de défaut. Ce sont les ratios Valeur ajoutée / Chiffres d'affaires, Disponibilité / Actif, Fonds Propres / Passif non courant, Mouvement confiés / Chiffres d'affaires et DLMT / CAF. Mis à part ce dernier ratio, les autres ratios ont un signe conforme aux attentes.

Mots clés : Risque de crédit, Crédit Scoring, Modèle de Score, régression logistique.

Abstract

Credit risk is considered to be the main risk facing credit institutions. It represents the potential loss incurred by a credit institution following a change in the credit quality of a counterparty or a set of counterparties.

According to Bale's agreement, there are several methods of managing and assessing credit risk: financial analysis, expert systems, neural networks and the scoring model which is the subject of our thesis. The latter requires a methodology to be followed to set it up, such as sample design, model development and validation.

It is within this framework that our work was conceived and it is summed up in the following problem: How can banks effectively assess credit risk? And what approach to follow to set up an efficient scoring system?

To answer this problem, we first present the credit risk, these types and its impact on banking activity, as well as the international regulations that govern banking activity. We then developed a credit scoring model based on the binary logistic regression model. For its estimate, we used a sample of 144 companies that obtained a loan from the CPA. We have thus brought out five significant variables to explain the probability of default. These are the Value added / Turnover, Availability / Assets, Equity / Non-current Liabilities, Transactions entrusted / Turnover and DLMT / CAF ratios. Apart from this last ratio, the other ratios have a sign in line with expectations.

Keywords: Credit risk, Credit Scoring, Score model, logistic regression.