

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITÉ ABDERRAHMANE MIRA DE BEJAIA
FACULTÉ DES SCIENCES EXACTES
DÉPARTEMENT DE RECHERCHE OPÉRATIONNELLE



Mémoire de fin d'études

En vue de l'obtention du diplôme de Master
en Mathématiques Appliquées

Option : Mathématiques Financières

THÈME :

*Evaluation des risques de crédit bancaire par la
méthode de Scoring et les SVMs*

Présenté par :

M. HADJ MAHAMMED Messaoud

Mlle HAMICHI Chahrazed

Devant le jury composé de :

Président	Dr. Nassim TOUCHE	Université de Béjaïa
Encadreur	Dr. Belkacem BRAHMI	Université de Béjaïa
Examineur	Dr. Rabah DJABRI	Université de Béjaïa
Examineur	Dr. Larbi ASLI	Université de Béjaïa

Année Universitaires : 2019-2020

Remerciements

Nous remercions Dieu de nous avoir donné le courage et la patience afin de terminer ce travail.

*Nous tenons à exprimer toute notre gratitude envers notre, encadreur **Dr. Belkacem BRAHMI** pour son soutien et son aide. En nous écoutant patiemment, et en discutant maintes fois de la nature et l'avancement de notre travail, il nous a permis de synthétiser, comprendre et expliquer un grand nombre de questions. Ces conseils et sa gentillesse nous a apporté un précieux soutien, qu'il soit chaleureusement remercié ici.*

Nos remerciements sont aussi adressés aux membres du jury qui nous ont fait l'honneur d'accepter de juger notre travail.

*Notre gratitude va également à **M. N. TOUCHE** pour avoir accepté de présider le jury de la soutenance.*

*Nos remerciements s'adressent également à **M. R. DJABRI** et **M. L. ASLI** pour l'honneur qu'ils nous ont fait en acceptant d'examiner ce mémoire.*

Nous remercions aussi tous les enseignants du département de Recherche Opérationnelle qui nous ont permis d'améliorer notre formation.

Merci à nos parents et à nos frères et sœurs pour nous avoir inculqué le goût d'apprendre, de nous avoir enseigné à penser et de nous avoir encouragé sans cesse pour aller plus loin.

Enfin, nous n'oublions pas de remercier ceux qui nous ont aidé d'une manière ou d'une autre d'élaborer ce travail.

Dédicaces

Je dédie ce modeste travail

À Mes très chers parents pour tous leurs sacrifices.

À nos frères et sœurs et la grande famille en leurs souhaitant un avenir radieux et plein de réussite.

À tous nos amis avec qui nous avons partagé les meilleurs moments de notre vie.

À notre encadreur et à tous les enseignants et le personnel de l'université de Bejaia.

HADJ MAHAMMED MESSAOUD

Dédicaces

A mes chers parents Que nulle dédicace ne puisse exprimer ce que je leurs dois pour leur bienveillance, leur affection et leur soutien

A mes chers frères En témoignage de mes sincères reconnaissances pour les efforts qu'ils ont consenti pour l'accomplissement de mes études.

Notre encadreur et à tous les enseignants et le personnel de l'université de Bejaia.

A ma famille et mes amis Pour leur aide et leur soutien moral durant l'élaboration du travail de fin d'études.

HAMICHI CHAHRAZED

TABLE DES MATIÈRES

Table des matières	I
Listes des figures	IV
Introduction Générale	1
I Notions de base sur les banques et les crédits	3
I.1 Notions de base	3
I.1.1 Notion d'entreprise	3
I.1.2 Notion de banque	4
I.1.3 Notion de profil ou de rentabilité	4
I.1.4 Intérêts nets	4
I.2 Comprendre les banques	5
I.2.1 La typologie des banques	5
I.2.2 Financement des banques	6
I.2.3 Les métiers de la banque	8
I.2.4 Les moyens de paiement	8
I.2.5 Rôle de la banque	9
I.2.6 La classification des banques	9
I.2.7 Les ressources de la banque	10
I.2.8 Les services bancaires	10
I.3 Le crédit bancaire	11
I.3.1 Notion de crédit	11
I.3.2 Objectif du crédit	12
I.3.3 Classification du crédit	12
I.3.4 Types d'amortissement d'un crédit bancaire	13

I.3.5	Types de crédits bancaires	13
I.3.6	Financement d'exploitation	13
I.4	Le risque de crédit	15
I.4.1	La nature du risque de crédit	15
I.4.2	Typologie des risques de crédit	16
I.4.3	Évolution de la perception du risque de crédit	16
I.4.4	Formalisation du risque de crédit	17
I.4.5	La mesure du risque de crédit	18
I.4.6	Les risques bancaires	19
II	Méthodes d'évaluation du risque de crédit bancaire	22
II.1	Méthode de scoring	22
II.1.1	Crédits scoring	22
II.1.2	Principe de la méthode	23
II.1.3	Les étapes de la méthode de scoring	23
II.1.4	Panorama des techniques de scoring appliquées au crédits aux particuliers	24
II.1.5	Application du crédit scoring aux particuliers	28
II.1.6	Les différentes techniques de discrimination et de prédiction appliquées au scoring	30
II.1.7	Condition d'utilisation efficace de la méthode :	33
II.1.8	Les avantages et les inconvénient de scoring	33
II.2	Méthode des séparateurs à vaste marge	35
II.2.1	Définition des SVMs	35
II.2.2	Objectif d'un SVM	36
II.2.3	Notion de base	37
II.2.4	classifieurs linéaires	38
II.2.5	La non linéarité (cas non séparable/ marge molle)	42
II.2.6	Les domaines d'applications	45
II.2.7	Les avantages et les inconvénients du SVM	45
III	Application des méthodes	47
III.1	La méthode des scores	47
III.1.1	Modèle Altman	47
III.1.2	Méthode statistique de construction du score	49
III.2	Application méthode SVM	53
III.2.1	Ensemble des données	53

III.2.2 Modèle mathématique	53
Conclusion Générale	57

TABLE DES FIGURES

II.1	Règles de décision pour le modèle d'Altman [2]	25
II.2	Règle de décision pour le modèle Conan et Holder [2]	27
II.3	Exemple support vecteurs machines SVM.	36
II.4	Principe d'hyperplan séparateur, il existe plusieurs celui qui correspond au minimum d'erreur est hyperplan optimal.	37
II.5	Classification d'un nouvel exemple inconnu.	38
II.6	Plus en exemple est éloigné du mauvais coté du séparateur (point bleu), plus la variable de relâchement ξ_i a une valeur importante.	41
II.7	Un seul point est mal classé (point bleu). L'écart mesure la distance du point à la marge numérique de l'hyperplan séparateur.	41
II.8	Illustration de passage de \mathbb{R}^3	43
III.1	Régression de Z en R_2	51
III.2	Application de la règle de décision	52
III.3	Le classement des entreprises d'après la valeur du score	53
III.4	Représentation graphique de la surface séparatrice entre les deux classes par la base German.	55
III.5	Classification des exemples de la base German.	56

INTRODUCTION GÉNÉRALE

La crise financière qui a secoué le monde au cours des dernières années, s'est exprimée par la faillite de grandes banques, elle a provoqué une remise en cause des modèles de gestion des risques bancaires, notamment le risque crédit. Ce risque doit être géré dorénavant par des méthodes fiables en mesure de réduire son impact négatif sur la rentabilité des banques.

Parmi ces méthodes, nous citons la méthode de scoring et la méthode d'apprentissage statistique SVM (séparateurs à vecteurs supports) qui connaît aujourd'hui une large utilisation par les banques. Le travail suivant a pour objet de mettre en évidence le contenu de ces méthodes, leurs apports pour les analystes de management des risques crédits.

Le crédit scoring a vu le jour suite aux travaux pionniers de BEAVER (1966) et d'ALTMAN (1968) et sur la base de ces recherches que le crédit scoring s'est développé partout dans le monde et a évolué au cours de ces vingt dernières années.

De manière générale, la méthode de scoring est une combinaison de plusieurs ratios, exprimées par une fonction. Le but est que le résultat du calcul (le scoring) soit statistiquement différent pour les entreprises ayant un risque de défaillance et pour les entreprises en bonne santé, permettant ainsi de distinguer au mieux les deux catégories de clients. Les scores sont des notes qui permettent de classer les entreprises.

La méthode des machines à vecteurs supports, appelée aussi machines à vaste marge (SVM), est dérivée de la théorie d'apprentissage statistique, dont le fondateur principal est Vladimir Vapnik (1995). Les SVMs ont connu un progrès très intéressant des points de vue théorique et applications. Cette approche a été appliquée dans divers domaines, en particulier

dans la gestion des risques bancaires et financiers.

Notre travail consiste à appliquer des méthodes statistiques, le scoring et les SVMs, pour l'évaluation des risques de crédit bancaire pour les clients d'une banque donnée.

Ce mémoire est constitué d'une introduction générale et de trois chapitres.

Dans le premier chapitre, nous allons tout d'abord définir les notions de base du banques, à savoir : les crédits et les risques de crédit. Le deuxième chapitre décrit la méthode de scoring et celle des SVMs, appliquées dans la gestion des risques dans le domaine bancaire. Ces deux techniques offrent aux décideurs une meilleure prise de décision concernant la classification des demandeurs de crédits en deux catégories : bon et mauvais payeur. La description de ses approches est aussi donné dans ce chapitre

Le troisième chapitre est consacré à l'application des deux méthodes (scoring et SVM) sur un cas pratique ; dont les données ont été récupéré sur Internet. Nous essaierons de mettre en évidence les étapes pratiques qu'il faut respecter pour appliquer ces techniques dans la gestion des risques de crédits.

Nous terminons ce travail par une conclusion générale et une bibliographie.

CHAPITRE I

NOTIONS DE BASE SUR LES BANQUES ET LES CRÉDITS

Introduction

Ce premier chapitre sera consacré à la présentation d'un certain nombre d'éléments et notions fondamentales : la définition de la notion de banque, son rôle et son activité ainsi que les crédits. Dans le cadre de son activité, la banque est sujette à différents risques ; ceux-ci seront brièvement définis et finalement, on présentera les approches pour l'identification et le contrôle des risques.

I.1 Notions de base

I.1.1 Notion d'entreprise

L'entreprise définit comme « un groupement humain hiérarchisé qui met en œuvre des moyens intellectuels, physiques, financiers, pour extraire, transformer, transporter, distribuer des richesses ou produire des services conformément à des objectifs définis par une Direction, personnelle ou collégiale, en faisant intervenir, à des degrés divers, les motivations de profit et d'utilité sociale » ([10],[31])

En effet, ce groupement de différents éléments oblige l'entreprise à entretenir plusieurs relations avec son environnement qui peut être :

1. Un environnement commercial lorsqu'il s'agit des relations avec les

- clients et les fournisseurs ;
2. Un environnement socio-culturel dans le cadre de ses relations avec son personnel ;
 3. Un environnement administratif : c'est l'État ;
 4. Un environnement financier : lorsqu'il s'agit de ses relations avec les banques ;
 5. Un environnement sociétal et politique : relations avec les actionnaires.

Ainsi, nous pouvons affirmer que le développement économique et social et la prospérité d'un pays dépendent essentiellement du niveau de croissance de l'entreprise et l'accomplissement de ses relations par ces dernières.

I.1.2 Notion de banque

Une banque est une entreprise particulière qui s'occupe des dépôts d'argent et des moyens de paiement. Au sens juridique, c'est une institution financière qui dépend du Code monétaire et financier. Une banque a pour fonction de proposer des services financiers : recevoir des dépôts d'argent, collecter l'épargne, gérer les moyens de paiement, accorder des prêts. Une banque fonctionne généralement sous forme d'agences, constituant ainsi un réseau[31].

I.1.3 Notion de profil ou de rentabilité

la rentabilité financière permet d'apprécier le taux d'investissement par le capital à risque. Elle est mesurée par le rapport entre le résultat et les fonds propres. Plus le résultat est positif et élevé, plus la rentabilité financière est bonne. la rentabilité financière fait partie des facteurs de décisions des banquiers en face de demande d'emprunts ou de crédits[31].

I.1.4 Intérêts nets

différence entre les intérêts perçus et les intérêts payés par la banque.[33]

i) **Produit Net Bancaire** : généralement assimilé au chiffre d'affaires d'une banque. Il s'agit de la marge d'intermédiation corrigée par les commissions, des produits et charges sur opérations bancaires diverses.

ii) **Résultat d'exploitation** : il s'agit du Produit Net Bancaire diminué des frais généraux (Charges du personnel, dotations aux amortissements et autres charges d'exploitation).

iii) **Résultat avant impôt** : c'est le résultat d'exploitation corrigé des activités extraordinaires.

iv) **Résultat Net** : est le résultat avant impôt corrigé de l'impôt sur les sociétés .

I.2 Comprendre les banques

Choisir une banque n'est pas une décision anodine. Afin de faire un choix éclairé parmi la multitude d'acteurs, il s'agit avant tout de bien en comprendre le fonctionnement.

I.2.1 La typologie des banques

- Il existe différents types de banques, selon leur statut juridique [23]
- les banques coopératives, ou mutualistes, détenues par leurs clients ;
 - les banques commerciales, généralement cotées en Bourse ;
 - les banques publiques, détenues par l'État ou par des organismes publics.

1- Banque coopérative

Les banques coopératives sont des sociétés qui appartiennent à leurs clients ou sociétaires, c'est-à-dire qu'ils en sont les actionnaires. Ceux-ci peuvent être des personnes morales ou physiques.

Une banque coopérative est une entreprise dont la propriété est collective et dans laquelle le pouvoir est démocratique. Les clients sociétaires ont la caractéristique d'être à la fois associés et usagers, à la fois propriétaires et clients de leur banque.

2- Banque commerciale

Les banques commerciales sont des sociétés constituées d'un capital détenu par des actionnaires extérieurs à leur clientèle, par opposition aux

banques coopératives. Les banques commerciales ont pour but de réaliser des bénéfices commerciaux et peuvent être cotées en bourse.

Une banque commerciale est une entreprise privée qui collecte de l'argent par les dépôts et sur le marché monétaire, et le redistribue sous forme de liquidité ou de crédit. Elle propose ainsi différents produits financiers et non financiers :

Crédits (crédit personnel ; prêt immobilier) ; placement et épargne ; assurances (assurance-vie ; assurances automobile et habitation).

3- Banques publiques

Une banque publique est une société bancaire dont l'État ou des acteurs publics sont propriétaires. Elle se distingue d'une banque commerciale par son type d'actionnariat, mais aussi souvent par certaines missions qui lui sont confiées par la puissance publique.

I.2.2 Financement des banques

En effet, ce ne sont pas les dépôts qui font les crédits, mais le contraire. Parce que l'argent est d'abord créé par les banques centrales sous forme de crédits, pour pouvoir être distribué. Les banques représentent ainsi l'intermédiaire entre les banques centrales, qui créent la monnaie, et l'économie, qui utilise cette monnaie pour son bon fonctionnement[23].

1- Banques nationales

Les banques dites nationales possèdent un seul réseau d'agences sur le territoire. Le traitement du client et des affaires est identique dans l'ensemble de ces agences. Cela concerne les tarifs, mais aussi toutes les procédures. Dans ce système, un client peut déposer un chèque dans une ville alors que son compte se trouve dans une autre, à plusieurs centaines de kilomètres.

2- Banques centrales

Les banques centrales sont des banques particulières, puisque ce sont des organismes étatiques qui ont le monopole de la création monétaire. Seules les banques centrales peuvent créer de la monnaie, qu'elles prêtent ensuite

aux banques commerciales pour que celles-ci puissent à leur tour assurer les moyens de paiement et de crédit de leurs clients.

3- Banques régionales

Les banques dont l'activité ne couvre qu'une seule région, dans certains cas, les agences mêmes sont autonomes et ont des structures qui leur sont propres. Les banques régionales ont pour spécificité de nouer de forts liens de proximité avec leurs clients. Les agences régionales proposent des circuits décisionnels courts et possèdent une très bonne connaissance du tissu économique. Les capitaux collectés par ces agences sont très souvent utilisés pour financer l'économie régionale. Souvent, ce genre de banque revendique une identité régionale forte. Les transferts d'argent et les dépôts de chèque, par exemple, ne peuvent pas se faire dans une autre caisse que celle où le compte est domicilié, et parfois même dans une autre agence de la même banque régionale.

4- Agences bancaires

Les agences bancaires sont des lieux différents du siège social de la banque. C'est un endroit ouvert au public et qui reste le point clé de la relation avec le client (sauf dans le cas des banques en ligne). Dans une agence, il est possible de :

- retirer de l'argent liquide ;
- ouvrir un compte ;
- retirer un chéquier ;
- demander le solde de son compte ;
- faire un virement ;
- rencontrer un conseiller clientèle.

Une agence bancaire est dirigée par un directeur d'agence, lui-même placé sous l'autorité d'un directeur de réseau ou d'un directeur régional. Une agence est également composée de conseillers de clientèle et de guichetiers.

5- Banques en lignes

Avec le développement des nouvelles technologies et notamment d'Internet, les banques ont vu se développer un nouveau système : les banques en ligne ou banques virtuelles. Les services proposés par une banque virtuelle est :

- ouverture d'un compte bancaire ;
- détention d'une carte de paiement et d'un chéquier ;
- obtention de crédits ;
- placements bancaires et financiers.

Une banque en ligne peut même proposer des services supplémentaires :

- comptes rémunérés ;
- offres de cash-back ; (soit un compte crédit en fonction des paiements effectués avec une carte bancaire).

Une banque en ligne propose une tarification des services moins élevée, des services spécifiques et uniques, et des opérations de gestion de compte faisables à distance, 24h/24.

I.2.3 Les métiers de la banque

Il existe plusieurs métiers dans la banque, selon le type de clientèle et le type d'opérations générées [23] :

1. les banques de dépôt gèrent l'argent des particuliers et des petites entreprises (banques de détail) ou des grandes entreprises (banques d'affaires) ;
2. les banques d'investissement s'occupent du financement des grandes entreprises ;
3. les banques privées font de la gestion de fortune pour les particuliers les plus riches.

Les banques proposent également des autres services :

- d'assurance(banc-assurance) ;
- d'assistance et de sécurité ;
- de coffre-fort.

I.2.4 Les moyens de paiement

Le particulier ou l'entreprise a le choix entre de nombreux moyens de paiement lors de ses achats [23] :

- l'argent liquide, ou liquidités, qui peut être transmis par mandat postal, transfert d'argent ou porte-monnaie électronique ;
- le paiement direct de banque à banque (TIP, prélèvement, virement bancaire) ;
- le chèque et le chèque de banque ;

- la carte bancaire ;
- le paiement en ligne.

Les entreprises ont également accès à d'autres moyens de paiement spécifiques :

- les effets de commerce (lettre de change ou billet à ordre), qui ont un intérêt, car ils représentent une forme de crédit ;
- le crédit documentaire et la remise documentaire pour les paiements internationaux.

I.2.5 Rôle de la banque

La banque joue un rôle d'intermédiaire entre les détenteurs et les demandeurs de capitaux. Son activité principale consiste à collecter les capitaux disponibles pour son propre compte et les utiliser sous sa responsabilité à des opérations de crédit. Elle peut également effectuer d'autres opérations de banque : les services bancaires de paiement, les opérations de change etc... [23].

I.2.6 La classification des banques

Le développement de l'économie mondiale, les tendances politiques et la spécialisation des secteurs économiques ont rendu nécessaire l'organisation de la profession bancaire et sa spécialisation. De manière générale, on distingue trois catégories essentielles :

Les banques de dépôts, les banques d'investissement et les banques d'affaires [23].

1- Les banques de dépôts

L'activité principale de ce type de banque consiste à effectuer des opérations de crédits et à recueillir les dépôts de fonds à vue et à terme. Au quotidien, elles gèrent les comptes des particuliers et des entreprises. Elles sont garantes de la sécurité des transactions financières.

2- Les banques d'investissement

Les banques d'investissement sont des banques dont l'activité consiste à accorder des crédits dont la durée est supérieure à deux ans.

3- Les banques d'affaires

En plus de l'octroi des crédits, Les banques d'affaires participent à la prise et la gestion de participations dans des affaires existantes ou en formation. Les opérations de financement engagées par ce type de banques immobilisent des capitaux pour une longue période.

I.2.7 Les ressources de la banque

Il y a deux catégories de ressources : les ressources clientèles et les ressources hors clientèle [23].

1- Les ressources de la clientèle

Ces ressources sont principalement formées par :

1. les dépôts (à vue et à terme) sont des liquidités placées en banque par les clients. Les dépôts à vue peuvent être restitués à la demande ; les dépôts à terme ne peuvent être restitués avant délai.
2. Les bons de caisse (nominatifs ou anonymes) sont des titres émis par la banque contre un placement de fond à rembourser à une échéance définie avec payement d'un intérêt.
3. les bons d'épargne sont des titres émis par la banque pour la collecte de ressources ; ils sont payés en plus des intérêts produits à leur échéance.

2- Les ressources hors clientèle

Ces ressources sont formées principalement par le marché interbancaire, les réescomptes, les avances de la banque centrale ou les fonds d'assainissement du trésor.

I.2.8 Les services bancaires

La banque met à la disposition de ses clients divers outils pour qu'ils puissent gérer leur argent, c'est ce qu'on appelle les services bancaires. Ils comprennent notamment les moyens de paiement. La banque est tenue de fournir à ses clients des services de base ([23],[17]) :

- l'ouverture, la tenue et la clôture d'un compte ;

- la délivrance de relevés d'identité bancaire (RIB) ;
- le changement d'adresse (un par an) ;
- la domiciliation de virements bancaires ou postaux ;
- l'envoi du relevé de compte ;
- l'encaissement des chèques et des virements ;
- les dépôts et les retraits d'espèces au guichet et au distributeur automatique ;
- le paiement par prélèvement ;
- la consultation à distance du compte sur Internet ;
- une carte (de paiement ou de retrait) ;
- les chèques (dont le nombre peut être limité).

I.3 Le crédit bancaire

Un crédit est la mise à disposition d'une somme d'argent, contre un engagement de remboursement avec intérêts. Il s'accompagne de frais, d'intérêts, éventuellement d'une durée et peut prendre la forme d'un prêt d'argent, d'un délai de paiement, d'une garantie bancaire ou encore d'un crédit de caisse (ouverture de crédit, crédit revolving ou autorisation de découvert).

I.3.1 Notion de crédit

Définition économique :

« Le crédit en économie, terme désignant des transactions en nature ou en espèces effectuées en contrepartie d'une promesse de remboursement dans un délai généralement convenu par avance ».

Le crédit est une opération fondée sur la confiance (crédit vient de mot latin « *crédence* » qui signifie « croire ») car le prêteur doit attendre l'exécution de la prestation que l'emprunteur devra réaliser. Dans certain cas, des garanties sont demandées à l'emprunteur pour accroître la confiance qui peut lui être accordée (caution, nantissement, hypothèque...).

Dans le domaine de crédit, les institutions financières jouent un rôle particulier même si des crédits peuvent être accordés par d'autres agents économiques. Comme c'est le cas par exemple des fournisseurs lorsqu'ils consentent de délais de paiement à leurs clients [12].

I.3.2 Objectif du crédit

Le domaine du crédit est extrêmement vaste. Il s'étale dans le temps et l'espace, s'étend à toutes sortes d'activités et répond à de multiples besoins économiques.

Il peut donc avoir pour objet aussi bien le financement des investissements des entreprises et des particuliers que les besoins temporaires de trésorerie. Il permet de faire face à tous les décalages, entre recettes et les dépenses quelle que soit l'origine des unes et des autres [31].

I.3.3 Classification du crédit

Différents critères peuvent être pris en compte pour classer les crédits, les principaux étant la durée (critère le plus utilisé), le bénéficiaire et la destination [7] :

— la durée : elle va dépendre du type d'opération pour laquelle le crédit est utilisé.

On relève :

1. le crédit à très court terme (au jour le jour) qui est utilisé par les banques pour ajuster quotidiennement leur trésorerie.
2. le crédit à court terme, de 3 mois à deux ans, utilisé par les ménages et les entreprises.
3. le crédit à moyen terme, entre deux et sept ans.
4. le crédit à long terme, plus de sept ans, concernant les ménages, les entreprises et les collectivités locales (communes, département...).

— les bénéficiaires : ce sont essentiellement les ménages, les entreprises et les administrations publiques

— la destination : il s'agit de l'utilisation qui va être faite des sommes mises à disposition.

Les banques sont les principaux organismes qui délivrent des crédits. Le service crédit occupe une place privilégiée dans une agence bancaire car il constitue l'organe qui contribue le plus à la formation du produit de la banque en général et de l'agence en particulier de par son rôle prépondérant qui consiste à développer le portefeuille clientèle et fructifier les ressources de la banque en ayant un volume satisfaisant d'engagements sains.

I.3.4 Types d'amortissement d'un crédit bancaire

L'amortissement ou le remboursement d'un crédit bancaire peut prendre plusieurs formes. En général, les échéances de paiement sont mensuelles[7].

On recense plusieurs types de remboursements :

- le remboursement à mensualités constantes, ou remboursement progressif du capital (les mensualités sont toujours les mêmes, mais au début elles comportent une part majoritaire d'intérêts, et à la fin une part majoritaire de capital) ;
- le remboursement à mensualités dégressives, ou remboursement constant du capital (tous les mois, le même montant de capital est remboursé, ce qui fait que le montant mensuel des intérêts associés décroît dans le temps) ;
- le remboursement in fine (on ne paye tous les mois que les intérêts, et on rembourse la totalité du capital au terme du crédit).

I.3.5 Types de crédits bancaires

Avant de d'emprunter à une banque, il faut prendre en compte le Taux Effectif Global (TEG), qui est le taux définissant le coût final pour l'emprunteur.

Il comprend les frais de dossier, les diverses commissions et le montant des garanties particulières. Le TEG est le vrai taux du crédit. C'est uniquement sur cette base que l'on peut comparer deux crédits bancaires. Suivant la raison pour laquelle vous empruntez de l'argent à une banque, un crédit différent vous sera proposé.

Le crédit bail est un crédit pour lequel une partie donne jouissance d'un bien à une autre, qui paie ce bien périodiquement. Il comporte une option d'achat au terme du contrat ; il s'agit du leasing, ou de la location vente[7].

I.3.6 Financement d'exploitation

Il existe différents modes de financement de l'exploitation à savoir [12] :

1. Les crédits par caisse ; qui seront présentés en deux volets :
 - **Les crédits globaux** : découvert, facilité de caisse, etc...
 - **Les crédits spécifiques** : avances sur titres, sur factures, sur marché public, etc...

2. Les crédits par signature tels que les cautions et crédits documentaires.

1- Les crédits d'exploitation par caisse (direct)

Ce sont des crédits qui donnent lieu à des décaissements certains de la part du banquier. Lorsque l'objet de ces crédits n'est pas précisé par l'entreprise, on parle de crédits « globaux ». Par contre, quand leur objet est clairement avancé par l'entreprise, et qu'il porte sur des postes précis de l'actif circulant, les crédits d'exploitations sont dits « spécifiques ».

i) Les crédits d'exploitation globaux : Ces crédits sont destinés à financer globalement les actifs cycliques du bilan sans être affectés à un objet précis.

Ils sont généralement appelés crédits par caisse ou crédits en blanc car, d'une part, ils sont utilisables par le débit d'un compte, et d'autre part, ils ne sont assortis d'aucune garantie sauf la promesse de remboursement du débiteur. Ce sont donc, des concours à risque très élevé.

De ce qui suit, on peut distinguer parmi les crédits d'exploitation globaux :

1. La facilité de caisse ;
2. Le découvert ;
3. Le crédit de campagne ;
4. Le crédit relais.

ii) Crédits d'exploitation spécifiques : Ce sont des crédits destinés à financer un poste bien déterminé de l'actif circulant d'une entreprise. Autrement dit, ils ont un objet bien précis et particulier qui constitue en lui-même une garantie de remboursement.

Les risques que prend le banquier lors de ces opérations sont d'ordre commercial ; le dénouement du crédit est tributaire de la capacité de l'entreprise à mener à bien les opérations traitées avec les tiers, d'une part, et des qualités morales et financières des mêmes clients de l'entreprise, d'autre part. Les crédits par caisse peuvent revêtir les formes suivantes :

1. Crédits de financement des stocks ;
2. Crédits de mobilisation de créances ;
3. Crédits de financement des marchés publics.

2- Les crédits d'exploitation par signature (indirects)

Mis à part les crédits comportant des décaissements ou bien des mobilisations de fonds, le banquier peut aider ses clients en leur permettant vis-à-vis de leurs créanciers de décaler leur paiement dans le temps ou bien de leur éviter les sorties de fonds.

En effet, cette aide se pratique par le moyen des engagements par signature (c'est-à-dire garantir l'insolvabilité du client).

Dans cette hypothèse, la banque prête seulement sa signature. Ce prêt peut intervenir sous forme de caution, aval, acceptation ou crédit documentaire.

I.4 Le risque de crédit

I.4.1 La nature du risque de crédit

Le risque de crédit est le risque particulier naissant d'une opération de prêt. Il correspond à la probabilité qu'un événement négatif affecte le service de la dette sur lequel le débiteur s'est engagé ([6],[11]).

i) **La notion de risque :** Un risque financier est un risque de perdre de l'argent à la suite d'une opération financière (sur un actif financier) ou à une opération économique ayant une incidence financière (par exemple une vente à crédit ou en devises étrangères).

ii) **Le risque naissant des opération de crédit :** Le risque peut être considéré comme l'incertitude affectant les montants et les dates auxquels les paiements du débiteur seront effectués. Il est lié aux aléas qui pèsent sur l'évolution de la situation économique et financière de la contrepartie.

On appréhende alors le risque de crédit comme une possibilité de pertes consécutives à cette évolution du débiteur, ce qui renvoie à deux états de la nature :

1. **la dégradation de la situation de l'emprunteur :** le gain original espéré ne rémunère pas l'incertitude grandissante à laquelle le créancier est exposé, c'est une forme de perte d'opportunité,
2. **la cessation des paiements du débiteur :** s'il ne peut payer, son créancier constate un défaut de paiement dont les conséquences

seront plus ou moins graves selon qu'il y a, ou non, faillite de la contrepartie.

I.4.2 Typologie des risques de crédit

Trois relations principales de crédit se distinguent du fait de leur fréquence et importance dans l'économie [6].

1. La dette publique.
2. Le crédit interentreprises.
3. Le crédit bancaire.

i) **La dette publique** : C'est l'ensemble des engagements financiers pris sous formes d'emprunts par un État, ses collectivités publiques et ses organismes qui en dépendent directement (certaines entreprises publiques, les organismes de sécurité sociale, etc...).

ii) **Le crédit interentreprises** : Le crédit inter-entreprises se définit macro économiquement comme la somme des crédits fournisseur que les entreprises industrielles, commerciales et de services se consentent entre elles. Au niveau micro-économique, c'est-à-dire au niveau d'une entreprise, il est la différence entre le crédit fournisseur que cette entreprise accorde à ses clients et celui qu'elle a obtenu de ses fournisseurs.

I.4.3 Évolution de la perception du risque de crédit

les banques ont analysé le risque de crédit par des méthodes intuitives et empiriques, relativement efficaces. Depuis une trentaine d'années, l'évolution de l'économie globale a totalement modifié la notion même de risque de crédit. Afin d'éviter la survenance de crises systémiques et de faillites bancaires, les régulateurs ont défini un ensemble de règles pour discipliner les banques dans l'octroi du crédit et la gestion de leurs risques. Elles leur imposent des procédures et des modèles, une transparence accrue et des règles formelles de couverture des risques. Les approches réglementaires de « Bâle II » et « Bâle III » ont redéfini la notion de risque de crédit et imposé un nouveau paradigme pour sa gestion ([6],[16]).

I.4.4 Formalisation du risque de crédit

On admet classiquement que les composantes du risque de crédit sont les suivantes [6] :

- **le défaut** : événement par lequel l'emprunteur n'honore pas une échéance fixée, c'est un « accident de crédit » ;
- **l'exposition à la date du défaut** : c'est le montant pour lequel la banque est en risque et qui inclut le capital restant dû ;
- **la perte en cas de défaut** : elle correspond à la fraction de l'exposition qui ne pourra être récupérée ; elle dépend fortement du taux de recouvrement (ou de récupération) en cas de défaut, lui-même lié à la situation de l'entreprise, à la législation et à la présence d'éventuelles garanties en faveur du créancier financier ;
- l'horizon du défaut, c'est-à-dire le moment futur où le défaut peut se produire.

L'approche contemporaine du risque de crédit permet de préciser ces composantes.

En simplifiant, supposons qu'une banque accorde un crédit à une entreprise.

Ce crédit est d'un montant M remboursable à une échéance t . Il génère un taux d'intérêt i .

S'il n'y a pas de défaut, à maturité (t) la valeur de ce crédit est :

$$V_t = M(1 + i)^t.$$

En cas de défaut à cet horizon, la banque ne « récupère » qu'une fraction R du crédit, exprimant le taux de recouvrement.

Alors la valeur du crédit à maturité (t) est :

$$V_t = M(1 + i)^t R.$$

Au temps 0, à la signature du contrat, la probabilité de défaut au temps t et le taux de recouvrement R en cas de défaut sont incertains (aléatoires). On appelle p la probabilité de défaut au temps t , sa valeur est comprise entre 0 (pas de défaut) et 1 (défaut).

En conséquence, la valeur du contrat à l'échéance est :

$$V_t = M(1 + i)^t(1 - p(1 - R)).$$

Cette approche simplifiée montre que le risque de crédit dépend de :

1. **la probabilité de défaut** : (expected default frequency) (EDF) ou probability of default (PD) ;
2. **l'exposition au défaut** : (exposure at default)(EAD), perte maximale en cas de défaut ;
3. **la perte en cas de défaut** : (loss given default) (LGD) égale à 1 moins le taux de recouvrement : $1 - R$.

En effet, la perte attendue sur un crédit (expected loss, EL) est égale à :

$$EL = EAD \times PD \times LGD.$$

La perte attendue sur un crédit est une variable aléatoire qui, associée à l'incertitude sur l'horizon du défaut, constitue le risque de crédit.

Le défaut (de paiement) correspond généralement à un événement objectif et mesurable par le créancier financier : la non-tenu d'un engagement de crédit, l'emprunteur se trouvant incapable d'honorer une échéance de sa dette financière.

Cependant, le défaut peut prendre une acception plus large comme la violation d'un covenant, la restructuration de la dette ou encore une dégradation sensible du rating d'une société.

Le défaut de paiement, au sens strict, est un événement confidentiel connu des seuls débiteurs et créanciers.

I.4.5 La mesure du risque de crédit

Mesurer le risque de crédit d'un emprunteur vise à évaluer, plus ou moins formellement, et plus ou moins quantitativement, la probabilité qu'il rencontre des difficultés financières et soit incapable d'honorer ses engagements, c'est-à-dire la probabilité que sa possible détresse financière génère un « incident de crédit » [6].

1- Contingences de la mesure du risque de crédit

L'analyse du risque de crédit a plusieurs objectifs :

- évaluer le niveau de risque présenté par un débiteur pour son créancier à l'instant t (spot), et anticiper son évolution probable (forward). En l'occurrence, il s'agit d'apprécier le degré d'aléa qui pèse sur sa capacité à respecter ses engagements de crédit ;

- exprimer ce risque de manière claire et intelligible sur une échelle par un système de notation au sens large, qualitative ou quantitative, permettant à la fois de l'exprimer clairement et de le situer dans l'absolu et de le relativiser dans le temps et dans l'espace ;
- examiner sa causalité et ses composantes afin de pouvoir expliquer le niveau de risque ;
- chercher à mesurer une probabilité mathématique de défaut ou de défaillance, donc traduire la note en probabilité d'occurrence.

I.4.6 Les risques bancaires

Un risque bancaire est un risque auquel s'expose un établissement bancaire lors d'une activité bancaire. L'activité bancaire, par son rôle d'intermédiation financière et ses services connexes, expose les établissements bancaires à de nombreux risques. On présente les différents risques bancaires [16],[21],[6].

1- Le risque de liquidité

La liquidité correspond à la capacité, pour une entité donnée, de faire face à ses obligations au temps t , autrement dit la capacité de sa trésorerie à supporter les flux négatifs résultant de ses obligations envers les tiers.

La liquidité est donc un concept ponctuel, à distinguer de la solvabilité qui recouvre la capacité la société à faire face à l'ensemble des obligations futures : est liquide une entreprise capable de faire face, à tout moment de son existence future, à des flux de trésorerie négatifs.

Autrement dit, le risque de liquidité est le risque de ne pas trouver de contrepartie pour vendre ou acheter un produit financier.

Plus le nombre d'acteur, sur le marché, est grand, plus les contreparties potentielles et donc la possibilité d'échanger rapidement son produit financier est grand.

2- Le risque de marché

Ce risque concerne les activités de négociation sur les marchés des capitaux face à une variation des prix de marché. Il correspond aux pertes éventuelles découlant de la fluctuation des valeurs des instruments financiers par suite des variations des taux d'intérêts, des cours de change des

devises et des cours des actions et des prix des marchandises. Ces risques tendent à s'amplifier avec la volatilité des marchés et de ce fait, peuvent se concrétiser d'une manière extrêmement rapide.

3- Le risque de taux d'intérêt

Le risque de taux d'intérêt global est le risque encouru en cas de variation des taux d'intérêt du fait de l'ensemble des opérations de bilan et hors-bilan, à l'exception, des opérations soumises aux risques de marché. Le risque de taux fait partie des risques majeurs et naturels auxquels sont confrontées les banques.

4- Le risque de change

C'est un risque qui intervient lors des investissements à l'étranger (emprunt en Dollar par exemple) et pour les produits financiers en devise étrangère. Une hausse de la devise par rapport à sa monnaie est un coût pour l'établissement, il peut néanmoins couvrir ce risque par des instruments financiers de couverture.

5- Le risque opérationnel

Les risques opérationnels se définissent comme les risques de pertes directes et indirectes résultant de l'inadaptation ou de la défaillance des procédures, des personnes, des systèmes ou découlant d'événements extérieurs.

C'est un risque pur en ce sens qu'il ne comporte que des pertes. La réduction des pertes découlant de ces risques pourrait être obtenue grâce à différents niveaux de contrôle avec une définition précise des tâches, une formation appropriée du personnel aux méthodes de contrôle et un recours à des pratiques sécurisées...

6- Le risque de crédit

Selon le dictionnaire financier « Lexique finance », Le risque de crédit se définit comme étant la probabilité qu'un débiteur soit dans l'incapacité de faire face au remboursement du crédit octroyé par une institution financière.

Le risque de crédit est bien évidemment est une fonction de la qualité de l'emprunteur.

De plus, ce risque est naturellement fonction de trois paramètres : le montant de la créance, la probabilité de défaut et la proportion de la créance qui ne sera pas recouvrée en cas de défaut.

Au sens large, le risque de crédit ou le risque de contrepartie désigne le risque de dégradation de la situation financière d'un emprunteur ou aussi le risque de taux de recouvrement, qui se présente dans le cas où il y a un recouvrement après défaut.

Dans ce qui suit, on va voir l'apport du crédit scoring dans la limitation du risque de crédit.

Conclusion

Pour conclure ce chapitre, on peut dire que les banques qui sont les principaux fournisseurs de crédit, tant aux entreprises qu'aux particuliers, offrent plusieurs types de crédit. La détermination du type de crédit sollicité par un client est fonction de son besoin de financement. Ainsi, la gestion bancaire répartit les crédits en 4 grands groupes tels que les crédits à l'exploitation, au financement, au commerce extérieur, et aux particuliers. Par ailleurs, le banquier avant d'octroyer un crédit, une décision sera prise à savoir un avis favorable ou défavorable. Par conséquent, l'activité d'octroi des crédits génère bien évidemment des risques majeurs qui pourront avoir un impact sur la rentabilité de la banque. La banque met donc en œuvre une politique de gestion de couverture de ces différents risques.

CHAPITRE II

MÉTHODES D'ÉVALUATION DU RISQUE DE CRÉDIT BANCAIRE

L'évaluation du risque de crédit revient à se poser la question de la solvabilité du client. Pour la prise de décision d'octroi de crédit, il est indispensable que le décideur dispose d'outils d'appréciations quantitatives qui vont lui permettre d'évaluer cette solvabilité afin d'assurer une probabilité minimale de défaillance du client.

L'analyse du risque crédit conduit à l'utilisation d'un certain nombre de méthodes, à savoir la méthode scoring et les méthodes de machines learning.

II.1 Méthode de scoring

II.1.1 Crédits scoring

« Le credit-scoring est une méthode statistique utilisée pour prédire la probabilité qu'un demandeur de prêt ou un débiteur existant fasse défaut » [25][22].

Le crédit scoring est un outil d'aide à la décision, basé sur l'utilisation des techniques statistiques pour prédire la probabilité de défaillance d'un demandeur de prêt. Elle vise à associer à chaque demande de crédit une note proportionnelle à la probabilité de l'emprunteur de faire défaut.

II.1.2 Principe de la méthode

La méthode de scoring consiste à calculer une combinaison linéaire d'un certain nombre de paramètres retenus comme les plus significatifs dans la discrimination entre les emprunteurs défaillants et emprunteurs non défaillants. Cette fonction score se présente sous la forme suivante :

$$F(\text{score}) = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_P X_P,$$

où $X_i, i = 1, 2, \dots, p$ sont les paramètres choisis et les $\alpha_i, i = 1, 2, \dots, p$ sont les p coefficients à estimer.

Les fonctions scores sont relativement nombreuses, elles présentent plusieurs atouts pour le secteur bancaire([21][22]).

1. **la simplicité** : L'utilisation du score est obtenu à partir d'un certain nombre d'informations synthétisées et offre une rapidité dans la prise de décision, ce qui constitue un double avantage : une charge de travail réduite et une réponse rapide pour le client.
2. **l'homogénéité** : Le crédit scoring donne la même décision quelque soit l'agence ou le temps de la prise de décision.

II.1.3 Les étapes de la méthode de scoring

Le crédit scoring est une méthode de prévision statistique utilise dans l'évaluation de la défaillance des entreprise et les clients. Dans notre cas, l'objectif principal de l'application de cette méthode est la discrimination des nouveaux clients demandeur de crédits en deux classes(bon payeurs/mauvais payeurs)[34].

La méthode de crédit scoring se base sur les quatre étapes suivantes :

- i) **Étape 1** : Dans cette première étape, il faut déterminer les éléments suivants :
 1. Un échantillon des clients demandeurs de crédit qui sont de la classe "bon payeurs",
 2. Un échantillon des clients demandeurs de crédit qui sont de la classe mauvais payeurs,
 3. Un ensemble de ratios.

ii) **Étape 2 :** Cette étape contient les phases suivantes :

1. La sélection des ratios les plus significatifs par la méthode discriminante,
2. La détermination des coefficients de pondération cohérents à chaque ratio.

iii) **Étape 3 :** L'objectif est de déterminer une fonction score qui effectue une meilleure combinaison des ratios retenus :

$$Z_t = a_1R_1 + a_2R_2 + \dots + a_nR_n;$$

Avec R_i : Ratio i de l'entreprise ;

a_i : coefficient de pondération du ratio R_i .

iv) **Étape 4 :** Dans cette étape, on définit un score limite servant de seuil de classement et de validation de la fonction.

II.1.4 Panorama des techniques de scoring appliquées au crédits aux particuliers

Parmi les modèles les plus classiques, on citera

1- Modèle d'Altman

Altman (1968) a développé un modèle de score établi sur la base d'un échantillon de 66 entreprises, dont 33 sont considérées comme défailtantes et 33 comme saines. Nous remarquons que la taille de l'échantillon utilisée dans le modèle d'Altman est faible ce qui peut altérer la qualité des résultats.

La technique statistique adoptée dans ce modèle est celle de l'analyse discriminante multivariée.

Elle repose sur une fonction de score combinaison linéaire des cinq ratios financiers jugés les plus pertinents pour départager au mieux les deux groupes d'entreprises (saines ou défailtantes).

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

$$Z = 1.2X_1 + 1.4 X_2 + 3.3 X_3 + 0.6 X_4 + 0.999 X_5,$$

où les variables explicatives sont définies comme suit :

X_1 = Capital Circulant/Total des Actifs(Liquidité),

X_2 = Profits Retenus/Total des Actifs(Rentabilité cumulative),

X_3 = Profits avant frais financiers et impôts/Total des Actifs (Rentabilité),

X_4 = Capitalisation Boursière/Valeur Comptable des Passifs (Structure du capital),

X_5 = Ventes/Total des Actifs(Efficacité),

Z : Résultat discriminant.

Le risque encouru par la banque varie dans le sens contraire de Z , avec 3 comme valeur critique tel que résumé par la figure (II.1).

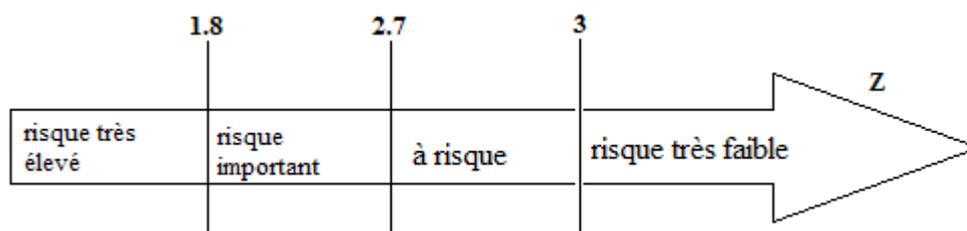


FIGURE II.1 – Règles de décision pour le modèle d'Altman [2]

Si le Z-score est supérieur à 3, l'entreprise à peu de risque de faire défaut, et s'il se situe entre 2.7 et 3, l'entreprise est à risque.

Pour un score compris entre 1.8 et 2.7, la probabilité de faire défaut est importante et l'entreprise est jugée à haut risque.

Enfin pour un score inférieur à 1.8 la probabilité d'un problème financier est très élevée.

L'utilisation d'un seuil unique pour procéder au classement des entreprises entraînait par conséquent des erreurs de classification. Dans cet échantillon, l'utilisation d'un seuil de 2.675 était celui qui minimisait ces erreurs.

L'application du modèle d'Altman est simple. Lorsqu'on désire prévoir l'état futur d'une entreprise, on calcule les ratios précédents et on applique l'équation Z . Si le résultat obtenu est inférieur à 2.675, on prévoit la faillite. S'il est plus grand, l'entreprise est selon toute vraisemblance, saine [8].

2- Modèle de la banque de France BDFI (1995)

La Banque de France a élaboré son modèle, en 1985, à partir des documents comptables et financiers de 3000 entreprises industrielles de moins de 500 salariés, qui ont été départagées en trois classes : saines, défailtantes et vulnérables. Les données sont formées de ratios.

Ce modèle célèbre de prévision des risques de défailtances par l'analyse statistique. Le score BDFI, s'intéresse plus particulièrement à l'endettement financier. Le score est calculé automatiquement chaque année en retenant une dizaine de ratios pertinents et les moins corrélés entre eux. Il est calculé, classiquement, par une analyse discriminante linéaire ([24],[2]).

La formule de ce score se présente comme suit :

$$\begin{aligned} Z = & -1.251X_1 + 2.003X_2 - 0.824X_3 + 5.221X_4 \\ & - 0.689X_5 - 1.164X_6 - 0.106X_7 + 1.408X_8 - 85.544. \end{aligned}$$

Les ratios qui ont été pris en considération sont les suivants :

- X_1 = Frais Financiers / Résultat Économique Brut,
- X_2 = Ressources Stables / Actif économique,
- X_3 = CA / Endettement,
- X_4 = EBE / CA HT,
- X_5 = Dettes Commerciales / Achats TTC,
- X_6 = Taux de Variation de la Valeur Ajoutée,
- X_7 = (Stocks + Clients- Avances Clients) / Production TTC,
- X_8 = Investissements Physiques / Valeur Ajoutée

La règle de décision de cette méthodologie est la suivante :

- Si $Z > 0,125$, alors l'entreprise est saine.
- Si $-0,250 < Z < 0,125$, alors l'entreprise est en zone d'incertitude.
- Si $Z < -0,250$, alors l'entreprise est défailtante.

3- Le modèle de Conan et Holder

Ce modèle est basé sur un échantillon de 190 petites et moyennes entreprises industrielles dont 95 considérées comme saines et 95 comme défailtantes. Les deux auteurs ont observé les valeurs de 31 ratios financiers se rapportant à toutes les entreprises de l'échantillon [2].

Ils ont conclu que parmi ces 31 ratios seuls 5 ratios sont les plus significatifs suivants :

$R_1 = \text{Excédent Bruit d'Exploitation/Total des Dettes,}$
 $R_2 = \text{Capitaux permanents/Total d'Actif,}$
 $R_3 = \text{Réalizable et disponible/Total d'Actif,}$
 $R_4 = \text{Frais financiers/Chiffre d'Affaires Hors Taxes,}$
 $R_5 = \text{Frais du personnel/Valeur Ajoutée.}$

L'application de la procédure d'analyse discriminante linéaire multiple a conduit les auteurs à proposer la fonction discriminante optimale à 5 variables suivante :

$$Z = 0.24R_1 + 0.22R_2 + 0.16R_3 - 0.87R_4 - 0.10R_5.$$

Pour affiner le Z-score et perfectionner la règle de décision des banquiers, Conan et Holder (1979) proposent une probabilité de défaillance selon la valeur du score Z . En fonction de ce score l'entreprise est considérée comme saine ou défaillante [2].

Le calcul de cette loi de probabilité de défaillance, en plus du score Z , permet de classer les entreprises selon leurs niveaux de risque présumé comme le montre la figure II.2.

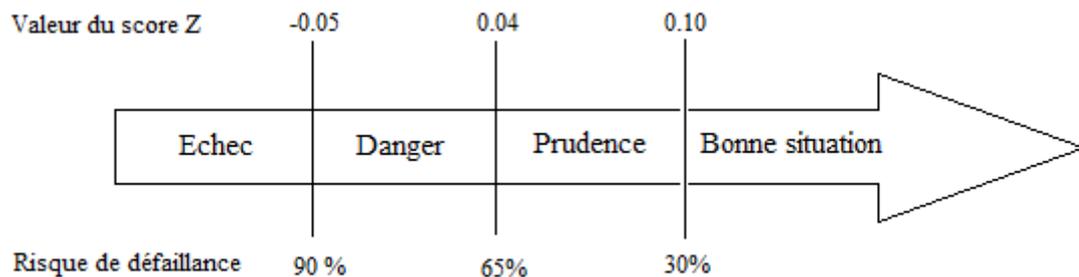


FIGURE II.2 – Règle de décision pour le modèle Conan et Holder [2]

Si le score $Z > 0.10$: Très bonne situation financière ;
risque de défaillance inférieur à 30%.

Si $0.04 < Z < 0.10$: Zone d'alerte ;
probabilité de défaillance de 30% à 65%. Pour l'entreprise

Si $-0.05 < Z < 0.04$: Zone de danger ;
probabilité de défaillance de 65% à 90%.

Si $Z < -0.05$: Entreprise classée défailante ;
probabilité de défaillance dépassant 90%.

II.1.5 Application du crédit scoring aux particuliers

Actuellement, la plupart des institutions financières sont dotées d'un système de notation interne ou font recours à des organismes de notation externes pour évaluer et prédire si un emprunteur sera bon ou mauvais payeur [2].

1- Préparation des données

Le processus de crédit, commence d'abord par une phase de collecte d'informations auprès du client, et auprès de sources externes, afin de former le dossier de crédit. Ces informations portent sur la forme (la qualité de l'emprunteur, l'équilibre du montage financier, le respect de la réglementation, etc), le fond (la capacité de remboursement des emprunteurs, les éléments d'appréciation du risque, etc.) et les garanties (cautions, hypothèques, etc.)

2- Quantification et la détermination des variables du crédit

Le banquier étudie le dossier du client et prend sa décision d'accorder ou non ce crédit en se basant sur le score calculé. Avant la phase de modélisation, les données collectées sont analysées pour déterminer les variables les plus pertinentes et discriminantes des clients. la plupart des variables sont de type qualitatif et ne peuvent être intégrés directement dans un modèle statistique.

3- La phase de modélisation

lorsque la phase de préparation des données terminée, on commence la phase de modélisation qui permet le calcul des scores. On utilise les différentes techniques de scoring qui sont explicitées déjà, on peut l'appliquer pour la construction de la fonction de score.

Dans le cas pratique, une fois l'étape de collecte de données terminée, des scores partiels sont attribués à chaque variable du dossier et le score final est obtenu par leur sommation.

Dans la dernière étape du processus consiste à comparer ce score au seuil fixé par la banque. Si le score dépasse ce seuil alors la banque accorde

le crédit sinon la demande est rejetée.

Exemple du score d'un demandeur de crédit immobilier

On suppose pour un demandeur de crédit immobilier, sept variables sont considérées comme suit (l'âge, l'ancienneté dans la profession, la situation matrimoniale, le nombre d'enfants, l'endettement, l'apport personnel et l'état du bien)[2].

Le demandeur de crédit est âgé de 59 ans, marié, ayant 3 enfants et qui travaille depuis 13 ans comme médecin.

Il désire acquérir un appartement neuf pour 300 000 €, avec une capacité d'autofinancement de 15%, et il n'a jamais contracté de crédit.

1.Age du client		2.Ancienneté professionnelle	
Attribut	Note	Attribut	Note
< 20 ans	0	≤ 1an	0
[20,30[0.5	2 à 3 ans	1
[30,50[5	3 à 5 ans	3
[50,60[3	5 à 7 ans	4
Plus de 60 ans	1	Plus de 7 ans	5

3.Situation matrimoniale		4.Nombre d'enfants	
Attribut	Note	Attribut	Note
Divorcé	1	≥ 3enfants	1
Célibataires	2	2 enfants	3
Marié	4	≤ 1enfants	5
Veuf	3		

5.Endettement		6.Apport personnel	
Attribut	Note	Attribut	Note
< 20%	0	< 15%	1
[20%,30%[1	[15%,40%[3
≥ 30%	3	≥ 40%	5

7.État du bien	
Attribut	Note
Nécessitant des travaux	1
Ancien en bonne état	3
neuf	5

TABLE II.1 – Un exemple simple de grille de score.

Le score attribué au dossier de crédit est égal à 21, calculé comme suit :
 Score = 3 + 5 + 4 + 1 + 0 + 3 + 5 = 21.

Avec l'hypothèse que, selon la banque, le seuil d'admission d'un dossier est de 20 points, le crédit sera donc accordé au client.

II.1.6 Les différentes techniques de discrimination et de prédiction appliquées au scoring

Les méthodes de scoring les plus utilisées par les banques, à cause de leur simplicité d'interprétation et leur grande fiabilité, sont généralement de type linéaire telles que l'analyse discriminante linéaire ou encore la régression logistique[2].

1- Analyse discriminante

L'Analyse discriminante de Fisher à été élaborée en 1936 par Ronald Aylmer Fisher. l'analyse discriminant est un modèle de classification fondé sur l'analyse des données, il est une technique statistique qui consiste à reclasser les emprunteurs en deux groupes (défaut et absence de défaut) et à rechercher l'ensemble des variables (ratios) qui permettent de prévoir le mieux, qui à fait défaut(dans le passé).

Dans le cas d'une classification à deux groupes, l'analyse discriminante peut être réduite à une analyse de régression.

La fonction discriminante se présente comme une combinaison linéaire de ces variables R_i , $1 \leq i \leq n$ d'où la relation suivante :

$$Score = a_0 + a_1 R_1 + a_2 R_2 + \dots + a_n R_n = \sum_{i=1}^n a_i R_i.$$

Les a_i représentent les coefficients ou pondérations associés aux ratios R_i .

La technique de l'analyse discriminante linéaire permet de trouver les valeurs des coefficients qui discriminent le mieux les deux groupes d'entreprises.

L'analyse discriminante de Fisher est basée sur l'étude des covariances inter classes ainsi que sur la construction d'un classifieur bayésien optimal au sens de la probabilité de l'erreur. Cet outil est à mi-chemin entre la modélisation et la classification, il faut en général passer par une méthode d'apprentissage pour valider la règle décisionnelle.

i) **La fonction discriminante du classifieur bayésien :** Notons, $\forall k \in [1, 2], \Sigma_k$ la matrice de variance-covariance et μ_k le centre de gravité du nuage d'individus associé à la classe k de Y , soit à $\mathbf{x}|\mathbf{Y}=\mathbf{k}$, d'effectif n_k .

Supposons que les observations de chaque classe soient générées selon une loi normale de paramètres (μ_k, Σ_k) :

$$P(Y = k/\mathbf{X} = x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)}.$$

Si les matrices de variance-covariance sont identiques où $\Sigma_1 = \dots = \Sigma_K = \Sigma$ (la matrice de variance-covariance toutes classes confondues), les fonctions discriminantes du classifieur bayésien deviennent [3] :

$$\delta_k(\mathbf{x}) = (x - \mu_k^T) \cdot \Sigma^{-1} \cdot (x - \mu_k) - 2 \log\left(\frac{n_k}{n}\right).$$

ii) **Le calcul des coefficients** : Le vecteur des coefficients associé à la règle décisionnelle se retrouve à travers de la formule de la fonction discriminante décrite ci-dessus,

$$\beta = (\mu_k^T \Sigma_k^{-1} \mu_k, \Sigma_k \mu_k).$$

Le premier terme du vecteur représentant le coefficient constant [3].

iii) **Règle décisionnelle** : Nous cherchons à savoir de quelle classe $Y = k$ le nouvel individu i que nous voulons classer est le plus près, via la distance définie par la matrice Σ_k et qui correspond la règle descriptive de Mahalanobis-Fisher, de définition,

$$d^2(x, \mu_k) = x^T \cdot \Sigma_k^{-1} \cdot x - 2 \cdot (x \cdot (\beta_2, \dots, \beta_k) - \beta_1).$$

La règle décisionnelle basée sur le théorème de Bayes nous donne :

$$P(Y = k/X = x) = \frac{P_k \cdot P(x/k)}{\sum_{k=1}^2 P_k \cdot P(X = x/k)} = \frac{P_k \cdot e^{-\frac{1}{2} \cdot d^2(x, \mu_k)}}{\sum_{k=1}^2 P_k \cdot e^{-\frac{1}{2} \cdot d^2(x, \mu_k)}}, P_k = \frac{n_k}{n}.$$

L'analyse discriminante renseigne sur la vraisemblance du défaut à court terme pour cet emprunteur : un score très faible, en au-delà un certain seuil déterminé par le modèle, indique une forte probabilité de défaut alors qu'un score élevé signifie au contraire un faible risque de défaut ([20], [3]).

2- Régression logistique

Est un modèle multivarié qui permet d'expliquer sous forme de probabilité, la relation entre une variable dépendante Y qualitative et une

ou plusieurs variables indépendantes X qui peuvent être quantitatives ou qualitatives.

Le modèle fournit la probabilité qu'un événement se produise ou non (dans notre cas bon payeurs ou non) et les variables indépendantes X sont celles susceptibles d'influencer la survenue ou non de l'événement.

Pour l'entreprise i , on suppose que :

$$Y = \begin{cases} 0 & \text{Si } \beta + \alpha X + u_i \leq 0, \\ 1 & \text{Si } \beta + \alpha X + u_i > 0. \end{cases}$$

où β est une constante et α est le vecteur ligne des coefficients d'une combinaison linéaire à estimer.

Les u_i sont les perturbations supposées indépendantes, de moyenne nulle, de variance 1 ; elles sont supposées suivre une loi logistique de fonction de répartition F :

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

pour déterminer les paramètres β et α on utilise la méthode du maximum de vraisemblance qui vise à fournir une estimation des paramètres qui maximise la probabilité d'obtenir les valeurs réellement observées sur l'échantillon

la méthode du maximum de vraisemblance s'écrit :

$Max_{\alpha} \Lambda = \prod_{i=1}^n p_i^{Y_i} (1-p)^{1-Y_i}$, avec $n = n_D + n_N$ la taille de l'échantillon, où

$$p_i = P(Y_i = 1/X = x_i) = P(u_i > -\beta - \alpha x_i) = 1 - F(-\beta - \alpha x_i) = \frac{1}{1 + e^{-\beta - \alpha x_i}}.$$

Donc on a :

$$Max_{\alpha} \Lambda = \prod_{i=1}^n \left(\frac{1}{1 + e^{-\beta - \alpha x_i}} \right)^{Y_i} \left(\frac{e^{-\beta - \alpha x_i}}{1 + e^{-\beta - \alpha x_i}} \right)^{1-Y_i}.$$

Pour l'optimisation, il est parfois plus facile d'optimiser la fonction log-vraisemblance

$$Max_{\alpha} \log(\Lambda) = Y_i \log(p_i) + (1 - Y_i) \log(1 - p_i).$$

[28].

II.1.7 Condition d'utilisation efficace de la méthode :

Le modèle doit contenir le maximum d'informations [34].

- Les données historiques doivent couvrir une période assez longue pour couvrir un cycle économique (autour de 7 ans).
- Les coefficients doivent être significatifs et conformes à la logique comptable et économique.
- L'utilisation des scores en dynamique : il est nécessaire d'examiner un peu plus en détail la situation financière du client afin de lutter contre la dérive temporelle.
- L'échantillon de construction doit contenir un grand nombre d'individus pour qu'il soit représentatif du portefeuille de crédit.
- Le modèle doit prévoir le défaut : le taux de bon classement doit être le plus élevé possible.

II.1.8 Les avantages et les inconvénient de scoring

Les techniques statistiques de scoring présentent un certain nombre d'avantages et limites[2] .

1- Les avantages

La méthode du scoring présente plusieurs atouts pour le secteur bancaire, pour l'établissement qui l'utilise et qui sont essentiellement 34 :

◇ **La rapidité :**

Il permet un traitement de masse de populations nombreuses d'emprunteurs et cela en un temps réduit. La durée de traitement des dossiers est passée de 15 jours à quelques heures pour les crédits standards, favorisant ainsi une économie de coût.

◇ **La simplicité :**

Utilisation du score s'obtient généralement à partir d'un certain nombre d'information.

◇ **L'homogénéité :**

Il est difficile de définir une politique de crédit homogène, par contre le crédit scoring donne la même décision quelle que soit l'agence ou le temps de la prise de décision.

◇ **Diminution des impayés :**

La méthode du Scoring est fondée sur une analyse statistique et ob-

jective des critères de risque. Il se révèle d'une efficacité supérieure aux méthodes classiques; elle permet de détecter de façon précoce les défauts de paiement des entreprises, d'estimer les pertes ainsi que d'évaluer les probabilités de défaillance.

◇ **La productivité :**

Une appréciation rapide relativement fiable qui permet en quelques minutes un grand nombre de cas qui ne présentent aucun problème et laisser les techniques traditionnelles opérer les dossiers tangents.

2- Les inconvénients

Les limites d'une démarche de scoring sont nombreuses :

- ◇ Au niveau de l'échantillonnage Le biais de la sélection : une décision valable doit être basée sur un échantillon représentatif, il faut donc prendre en compte tous les dossiers (acceptés et rejetés).
- ◇ La disponibilité de l'information : d'une part, pour élaborer une évaluation statistique, une base de données présentant un nombre suffisamment grand de prêts est nécessaire.
D'autre part, cette base de données n'est pas toujours sous une forme électronique, du moins pour certaines banques en voie de développement, ce qui est coûteux en matière de temps pour élaboration du modèle.
- ◇ Changement du comportement Les problèmes de score ne peuvent détecter les changements relatifs au comportement des emprunteurs vis-à-vis du défaut.
- ◇ Les méthodes statistiques de scoring supposent comme toute autre méthode statistique que le futur est identique au passé;
- ◇ Le risque est expliqué par les seules variables disponibles;
- ◇ Il existe un vrai problème de biais de sélection dans l'élaboration d'une méthode de crédit scoring.

II.2 Méthode des séparateurs à vaste marge

Parmi les méthodes à noyaux, inspirées de la théorie statistique de l'apprentissage de Vladimir Vapnik, séparateurs à vaste marge (SVM) constituent la forme la plus connue.

SVM est une méthode de classification binaire par apprentissage supervisé, elle fut introduite par Vapnik en 1995. Cette méthode est donc une alternative récente pour la classification. Elle repose sur l'existence d'un classifieur linéaire dans un espace approprié. Puisque c'est un problème de classification à deux classes, cette méthode fait appel à un jeu de données d'apprentissage pour apprendre les paramètres du modèle. Elle est basée sur l'utilisation de fonctions dites noyau (kernel) qui permettent une séparation optimale des données.

II.2.1 Définition des SVMs

C'est une technique de data mining, ayant le but de SVM est de trouver un classifieur qui va sépare au mieux les données et maximiser la distance entre ces deux classes.

Avec SVM, ce classifieur est un classifieur linéaire appelé hyperplan. des performances meilleures que les méthodes statistiques traditionnelles en matière de classification.

Elle est d'utilisation récente dans le domaine de crédit scoring. Elle peut être utilisée dans les cas de séparation linéaire ou non linéaire entre classes. Pour le cas de la classification linéaire, et si les observations sont linéairement séparées,

la méthode permet de séparer les individus en deux classes par une frontière linéaire. Cette frontière est un hyperplan optimal qui garantit une grande marge de séparation entre deux classes (figure II.3) [29].

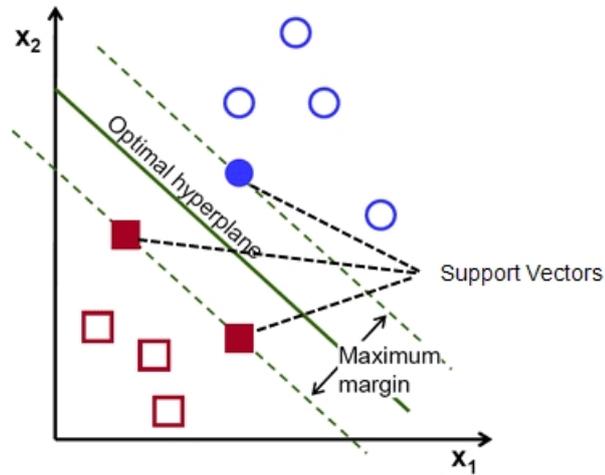


FIGURE II.3 – Exemple support vecteurs machines SVM.

II.2.2 Objectif d'un SVM

Soit un nuage de points de nature différente (points rouges et bleus). L'objectif recherché d'un classifieur SVM est de trouver un hyperplan séparateur (frontière de décision) qui permet de séparer l'ensemble des points en deux régions avec un minimum d'erreur, trouver l'hyperplan optimal.

La figure(II.4) montre qu'il existe plusieurs hyperplan séparateur dont les performances d'apprentissage sont équivalentes. Pour éviter ce problème Vapnik montre qu'il existe un seul hyperplan optimal qui maximise la marge [30].

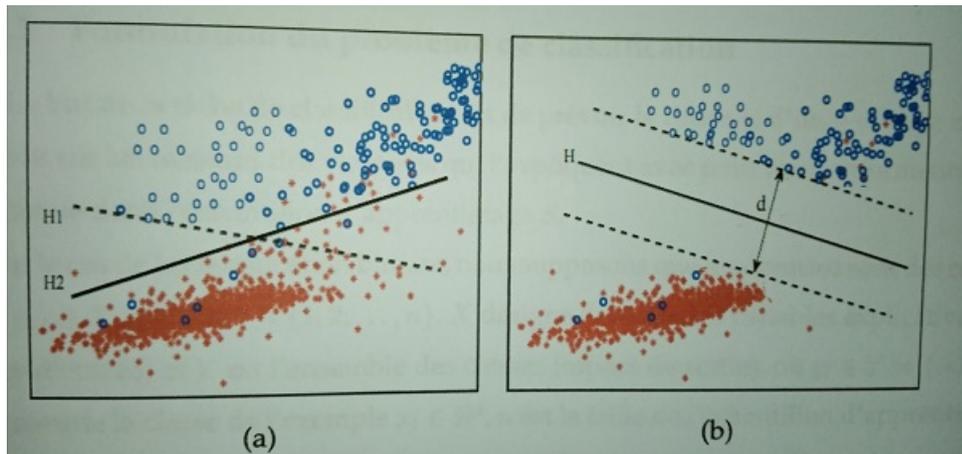


FIGURE II.4 – Principe d'hyperplan séparateur, il existe plusieurs celui qui correspond au minimum d'erreur est hyperplan optimal.

II.2.3 Notion de base

1- Apprentissage statistique

La notion d'apprentissage étant importante, nous allons commencer par effectuer un rappel. Il se divise en apprentissage supervisé et non supervisé. dans les cas qui concerne les SVMs est l'apprentissage supervisé.

2- Apprentissage automatique

L'apprentissage automatique est une méthode utilisée en intelligence artificielle.

Il s'agit d'algorithmes (procédures traduites en langages informatiques) qui analysent un ensemble de données afin de déduire des règles qui constituent de nouvelles connaissances permettant d'analyser de nouvelles situations.

3- La marge

La distance d'un point x_i à l'hyperplan (H) est :

$$d(x_i) = \frac{|w^T x_i + b|}{\|w\|}.$$

ou w est le vecteur orthogonal à l'hyperplan et b est le déplacement par rapport à l'origine.

Les points x_i les plus proches de (H) sont situés sur (H_1) et (H_2) , qui sont les hyperplans canoniques, ayant donc une distance $d(x_i) = \frac{1}{\|w\|}$.

Ainsi, la marge géométrique entre les deux classes est la distance euclidienne entre (H_1) et (H_2) , qui vaut $\gamma = \frac{2}{\|w\|}$.

i) Pourquoi maximiser la marge ? Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple.

De plus, si l'on trouve le classifieur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans la figure II.5, qui suit, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé

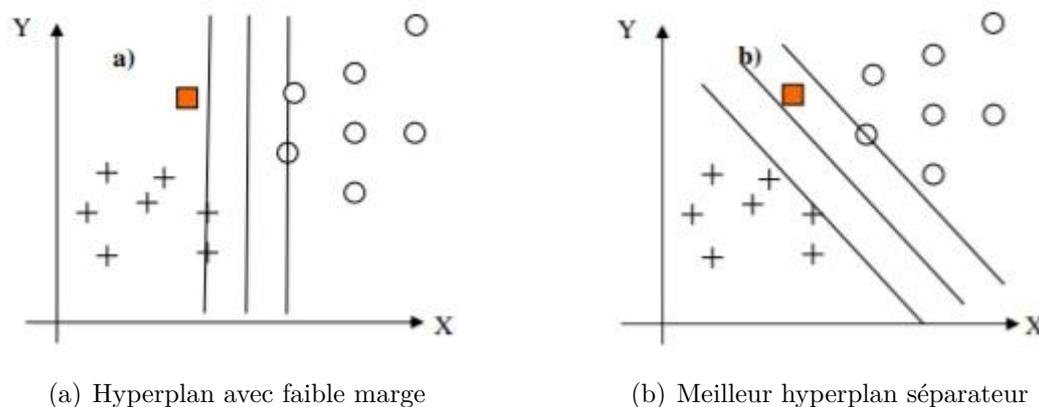


FIGURE II.5 – Classification d'un nouvel exemple inconnu.

II.2.4 classifieurs linéaires

Les classifieurs linéaires sont une famille d'algorithmes de classement statistique. Le rôle d'un classifieur est de classer dans des groupes (des classes) les échantillons qui ont des propriétés similaires, mesurées sur des observations.

Un classifieur linéaire est un type particulier de classifieur, qui calcule la décision par combinaison linéaire des échantillons.

Dans cette sous-section, on décrit l'utilisation des SVM pour un problème de classification linéaire. On représente la méthode générale de construction de hyperplan qui sépare des données appartenant à deux classes, On traite le cas de données séparable et par la suite on présent le cas de données non séparables [15].

1- SVM à marge dure

On se place ici dans le cas où les exemples de l'échantillon S sont linéairement séparable en deux classe par un hyperplan.

Même dans le cas simple le choix d'hyperplan séparateur n'est pas évident, car il existe une infinité d'hyperplan séparateur. Pour résoudre se problème, Vapnik à montre qu'il existe un unique hyperplan optimal.

Il s'agit de trouver le couple (w, b) qui maximise la marge pour détermine l'équation de l'hyperplan optimal (H) :

$$(H) : w^T x + b = 0$$

Soient x^+ et x^- deux point classées dans les frontière positives et négative. on considère les x^+ et x^- situe respectivement sur les hyperplan canoniques (H_1) et (H_2) .

$$(H_1) : w^T x^+ + b = +1 \quad \text{et} \quad (H_2) : w^T x^- + b = -1$$

correspondant à : $f(x^+) = +1$ et $f(x^-) = -1$

On sait que la distance entre x et (H) définie :

$$d_{xH} = \frac{|w^T x + b|}{\|w\|}.$$

La distance entre (H) et x^+ et x^- est $\frac{1}{\|w\|}$ donc la marge est égale à $\frac{2}{\|w\|}$ donc pour maximiser la marge revient à minimiser $\|w\|$. l'équation $w.x + b = 0$ revient à trouver les paramètres (w, b) qui maximise la marge. Il s'agit d'un problème d'optimisation d'une fonction objectif sous contrainte :

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ \text{s.c } y_i [w^T x_i + b] \geq 1; \\ i \in I = \{1, 2, \dots, n\}. \end{cases} \quad (\text{II.1})$$

Le problème (II.1) transforme en problème dual, qui introduit les multiplieurs de Lagrange :

$$L(w, b) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i (w^T x + b) - 1].$$

Le problème doit satisfaire les conditions de KKT qui consistent à annuler les dérivées partielles du Lagrange par rapport aux variables primales w et b :

$$\begin{cases} \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^n \alpha_i y_i = 0, \\ \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^n \alpha_i y_i x_i. \end{cases}$$

Le problème $\frac{1}{2} \|w\|^2$ est convexe. En ajoutant les valeurs obtenues par les conditions de KKT, donc on obtient la forme duale équivalente du problème (II.1) suivant :

$$\begin{cases} \min_{\alpha} L(\alpha) = \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n (y_i^T y_j x_i^T x_j) \alpha_i \alpha_j - \sum_{i=1}^n \alpha_i \\ \text{S.c : } \sum_{i=1}^n y_i \alpha_i = 0, \\ 0 \leq \alpha_i, i \in I \end{cases}$$

Lorsque on obtient la solution optimale du dual $\hat{\alpha}$ le vecteur de poids \hat{w} de l'hyperplan de marge maximale est

$$\hat{w} = \sum_{i=1}^n \hat{\alpha}_i y_i x_i = \sum_{i \in I_{sv}} \hat{\alpha}_i y_i x_i$$

2- SVM à marge souple

Souvent il arrive que même si le problème est linéaire, les données sont affectées par un bruit (par ex. de capteur) et les deux classes se retrouvent mélangées autour de l'hyperplan de séparation. Pour gérer ce type de problème on utilise une technique dite de marge souple, qui tolère les mauvais classements :

- Rajouter des variables de relâchement des contraintes ξ_i
- Pénaliser ces relâchements dans la fonction objectif.

L'idée est de modéliser les erreurs potentielles par des variables d'écart positives ξ_i associées aux observations $(x_i, y_i), i = 1, \dots, n$.

Si un point (x_i, y_i) vérifie la contrainte de marge $y_i (w^T x_i + b) \geq 1$ alors la variable d'écart (qui est une mesure du coût de l'erreur) est nulle. Nous avons donc deux situations :

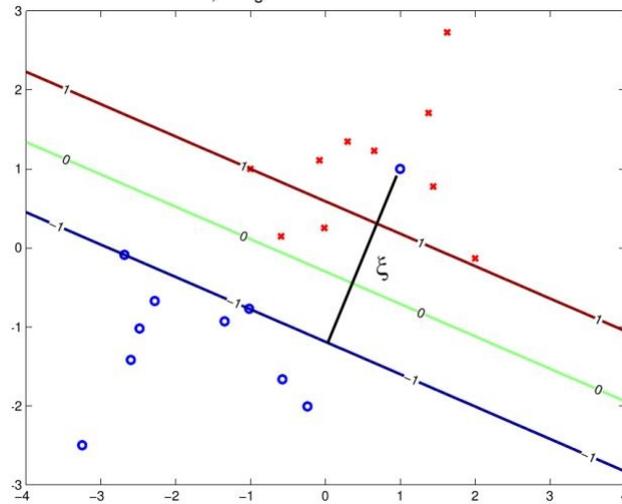


FIGURE II.6 – Plus en exemple est éloigné du mauvais coté du séparateur (point bleu), plus la variable de relâchement ξ_i a une valeur importante.

- Pas d'erreur : Si $y_i(w^T x_i + b) \geq 1$, donc $\xi_i = 0$
- Erreur : Si $y_i(w^T x_i + b) < 1$, donc : $\xi_i = 1 - y_i(w^T x_i + b)$

On associe à cette définition une fonction coût appelée « coût charnière » : $\xi_i = \max(0, 1 - y_i(w^T x_i + b))$.

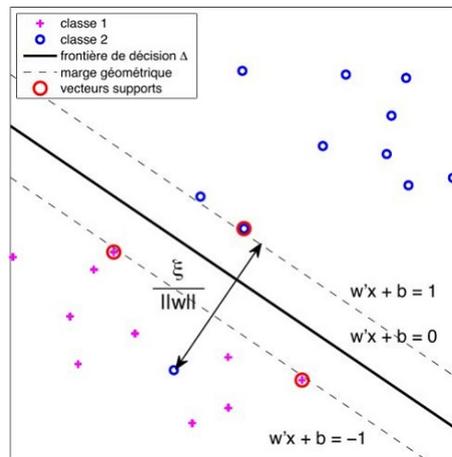


FIGURE II.7 – Un seul point est mal classé (point bleu). L'écart mesure la distance du point à la marge numérique de l'hyperplan séparateur.

Le problème d'optimisation dans le cas des données non-séparables est :

$$\begin{cases} \min_{w, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i, \\ \text{s.c : } y_i(w^T x_i + b) \geq 1 - \xi_i, \\ \xi_i \geq 0, b \in \mathbb{R}, y_i \in [-1, 1], \\ i = 1, \dots, n \end{cases}$$

Si toutes les variables d'écart $\xi_i = 0$, on retrouve le problème linéairement séparable traité plus tôt.

C est une variable de pénalisation des points mal classés faisant un compromis entre la largeur de la marge et les points mal classés.

Les variables ξ_i s'appellent aussi variables ressort (anglais : slack variables). Par la même procédure qu'avant, on obtient le problème dual :

$$\begin{cases} \min_{\alpha} L(\alpha) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (y_i y_j x_i' x_j) \alpha_i \alpha_j - \sum_{i=1}^n \alpha_i \\ \text{S.c : } \sum_{i=1}^n \alpha_i y_i = 0, \\ 0 \leq \alpha_i \leq C; \\ i = 1, \dots, n. \end{cases}$$

La fonction de Lagrange associée est la suivante :

$$L(\alpha, b, s, \xi) = \frac{1}{2} \alpha' D \alpha - e' \alpha + b y \alpha - s' \alpha + \xi' (\alpha - C e).$$

puis on calcule \hat{w} de l'hyperplan séparateur (H) de la même façon que le cas précédent.

II.2.5 La non linéarité (cas non séparable/ marge molle)

On part du problème primal linéaire et on introduit des variables « ressort » pour assouplir les contraintes.

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \forall i, y_i (w^T x_i + b) \geq 1 - \xi_i. \end{cases}$$

On pénalise par le dépassement de la contrainte.

On en déduit le problème dual qui à la même forme que dans le cas séparable

$$\begin{cases} \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum \alpha_i \alpha_j y_i y_j x'_i x_j \\ \forall i, 0 \leq \alpha_i \leq C; \\ \sum_{i=1}^n \alpha_i y_i = 0. \end{cases}$$

La seule différence est la borne supérieure C sur les α ([15],[30]).

i) **Fonction noyau (kernel) :** Dans les cas linéaire, on pouvait transformer les données dans un espace où la classification serait plus aisée.

Dans ce cas, l'espace de redescription utilise le plus souvent est \mathbb{R} (ensemble des nombres réels).

Il se trouve que pour des cas non linéaires, cet espace ne suffit pas pour classer les entrées. On passe donc dans un espace de grande dimension

$$\phi : R^d \longrightarrow Fx \longrightarrow \phi(x) \text{ avec } \text{card}(F) > d.$$

Exemple :

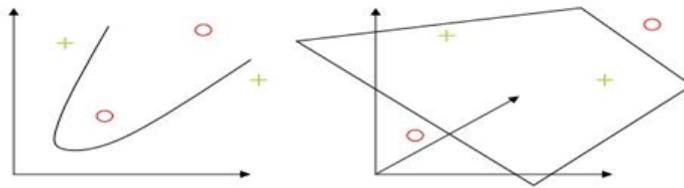


FIGURE II.8 – Illustration de passage de \mathbb{R}^3 .

Le passage dans $F = \mathbb{R}^3$ rend possible la séparation linéaire des données. On doit donc résoudre :

$$\begin{cases} \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \phi(x_i) \phi(x_j), \\ \forall i, 0 \leq \alpha_i \leq C, \\ \sum_{i=1}^n \alpha_i y_i = 0. \end{cases}$$

Et la solution à la forme :

$$f(x) = \sum_{i=1}^n \alpha_i^* y_i \phi(x_i) \phi(x) + b.$$

Le problème et sa solution ne dépendent que du produit scalaire $\phi(x_i) \cdot \phi(x)$. Plutôt que de choisir la transformation non-linéaire $\phi : X \rightarrow F$, on

choisit une fonction $K : X \times X \rightarrow \mathbb{R}$ (nombres réels) appelée *fonction noyau*.

Elle représente un produit scalaire dans l'espace de représentation intermédiaire. Du coup k est linéaire (ce qui nous permet de faire le rapprochement avec le cas linéaire des paragraphes précédents).

Cette fonction traduit donc la répartition des exemples dans cet espace $k(x, x') = \phi(x) \cdot \phi(x')$ Lorsque k est bien choisie. on n'a pas besoin de calculer la représentation des exemples dans cet espace pour calculer ϕ .

Exemple :

$$\text{Soit } x = (x_1, x_2) \text{ et } \phi(x) = (x_1^2, \sqrt{2} x_1 x_2, x_2^2)$$

Dans l'espace intermédiaire, le produit scalaire donne :

$$\begin{aligned} \phi(x) \cdot \phi(x') &= x_1^2 x_2'^2 + 2x_1 x_2 x_1' x_2' + x_2^2 x_2'^2 \\ &= (x_1 x_1' + x_2 x_2')^2 \\ &= (x \cdot x')^2 \end{aligned}$$

On Peut donc calculer $\phi(x) \cdot \phi(x')$ sans calcule $\phi : k(x, x') = (x, x')^2$
 k représentera donc le noyau pour les entrées correspondantes mais devra néanmoins remplir certaines conditions.

ii) **Condition de Mercer :** Une fonction k symétrique est un noyau si $(k(x_i, x_j))_{i,j}$ est une matrice définie positive.

Dans ce cas, il existe un \mathbb{F} et une fonction ϕ tels que $k(x, x') = \phi(x) \cdot \phi(x')$.

Problèmes

- Cette condition est très difficile à vérifier
- Elle ne donne pas d'indication pour la construction de noyaux
- Elle ne permet pas de savoir comment est ϕ

Exemples de noyaux :

- Linéaire $k(x, x') = xx'$,
- Polynomial $k(x, x') = (x \cdot x')^d$ ou $(c + x \cdot x')^d$,
- Gaussien $k(x, x') = e^{-\|x - x'\| \frac{2}{\sigma}}$,
- Laplacien $k(x, x') = e^{-\|x - x'\| \frac{1}{\sigma}}$.

On remarque en général que le noyau gaussien donne de meilleurs résultats et groupe les données dans des paquets nets.

En pratique, on combine des noyaux simples pour en obtenir de plus complexe.

II.2.6 Les domaines d'applications

SVM est une méthode de classification qui montre de bonnes performances dans la résolution de problèmes variés. Cette méthode a montré son efficacité dans de nombreux domaines d'applications tels que le traitement d'image, la catégorisation de textes ou le diagnostics médicales et ce même sur des ensembles de données de très grandes dimensions.

La réalisation d'un programme d'apprentissage par SVM se ramène à résoudre un problème d'optimisation impliquant un système de résolution dans un espace de dimension conséquente.

L'utilisation de ces programmes revient surtout à sélectionner une bonne famille de fonctions noyau et à régler les paramètres de ces fonctions. Ces choix sont le plus souvent faits par une technique de validation croisée, dans laquelle on estime la performance du système en la mesurant sur des exemples n'ayant pas été utilisés en cours d'apprentissage.

L'idée est de chercher les paramètres permettant d'obtenir la performance maximale. Si la mise en œuvre d'un algorithme de SVM est en général peu coûteuse en temps, il faut cependant compter que la recherche des meilleurs paramètres peut requérir des phases de test assez longues [30].

II.2.7 Les avantages et les inconvénients du SVM

i) Les avantages

- ✓ Capacité à traiter de grandes dimensionnalités (variables élevé) ;
- ✓ Robuste même quand le rapport (observations / variables) est inversé ;
- ✓ Traitement des problèmes non linéaires avec le choix des noyaux ;
- ✓ Robuste par rapport aux points aberrants (contrôlé avec le paramètre C) ;

- ✓ Points supports donne une bonne indication de la complexité du problème traité;
- ✓ Souvent performant dans les comparaisons avec les autres approches ;
- ✓ Paramétrage permet de la souplesse (ex. résistance au sur-apprentissage avec C).

ii) **Les inconvénients**

- ✓ Difficulté à identifier les bonnes valeurs des paramètres (et sensibilité aux paramètres) ;
- ✓ Difficulté à traiter les grandes bases avec observations très élevé (capacité mémoire avec matrice de Gram – si implémentation naïve, temps de calcul) ;
- ✓ Problème lorsque les classes sont bruitées (multiplication des points supports) ;
- ✓ Pas de modèle explicite pour les noyaux non linéaires (utilisation des points supports) ;
- ✓ Difficulté d'interprétations (ex. pertinence des variables) ;
- ✓ Le traitement des problèmes multi-classes reste une question ouverte ;

Conclusion

Le crédit scoring est généralement considéré comme une méthode d'évaluation du niveau du risque associé à un dossier de crédit potentiel. Cette méthode implique l'utilisation de différentes techniques statistiques pour aboutir à un modèle de scoring basé sur les caractéristiques du client, il sera classé par le modèle comme : Bon Payeur ou Mauvais Payeur.

Par contre la méthode support vecteur machine consiste à projeter les données de l'espace d'entrée non-linéairement séparables (marge dure et marge souple) dans un espace de plus grande dimension appelé espace de caractéristiques de façon à ce que les données deviennent linéairement séparables. Dans cet espace, la technique de construction de l'hyperplan optimal est utilisée pour calculer la fonction de classement séparant les deux classes.

CHAPITRE III

APPLICATION DES MÉTHODES

Introduction

Dans ce chapitre, nous allons présenter la méthode scoring avec modèle Altman et la méthode statistique du score, aussi la méthode SVM appliquée à notre problème, à savoir le risque de crédit bancaire et implémenter par la suite sur langage de programmation MATLAB 2014 version 8.0.3

III.1 La méthode des scores

III.1.1 Modèle Altman

La mesure d'Altman ou le z score d'Altman est un instrument de synthèse permettant la prévision de la défaillance d'une entreprise à partir d'un certain nombre de ratios. [1]

Cet instrument permet de répondre facilement à la question : « Est-ce que l'entreprise est susceptible de faire faillite ? ». Son avantage est la simplicité, ses résultats permettent une réponse par oui ou non. Cette méthode est très intéressante dans la mesure où elle considère des mesures financières jumelées et non séparées.

La mesure d'Altman suit la fonction suivante :

$$Z = a_1X_1 + a_2X_2 + \dots + a_nX_n,$$

où les ratios X_i sont pondérés par des coefficients $a_i, i = 1, 2, \dots, n$. Pour juger la situation financière d'une entreprise, la fonction de régression Z

prend en compte plus d'un aspect, à savoir : la liquidité, la solvabilité, la rentabilité, l'activité et la croissance.

Altman voulait que son modèle ne néglige aucun de ces aspects de la situation financière d'une entreprise. Il a considéré un échantillon de 66 sociétés, où Appliqué sur un échantillon de 66 sociétés américaines, ce modèle a réussi à classer 95% des sociétés en faillite un an avant à leur faillite.

Pour les firmes canadiennes, la fonction discriminante finale est la suivante :

$$Z \text{ score} = 1,2.X_1 + 1,4.X_2 + 3,3.X_3 + 0,6.X_4 + 0,999.X_5$$

Sachant que les différentes variables sont définies comme suit :

X_1 = Capital Circulant/Total des Actifs,

X_2 = Profits Retenus/Total des Actifs,

X_3 = Profits avant frais financiers et impôts/Total des Actifs,

X_4 = Capitalisation Boursière/Valeur Comptable des Passifs,

X_5 = Ventes/Total des Actifs.

Ceci étant pour déduire la probabilité qu'une entreprise d'exploitation privée fasse faillite.

Un score de 1.10 ou moins indique que la faillite est probable, alors qu'un score de 2.60 ou supérieur est un indicateur de la bonne santé de l'entreprise.

La zone grise est le score compris entre 1.10 et 2.60. Les probabilités de faillite pour les entreprises sont de 95% pendant la première année et de 70% pour la deuxième année.

Un score plus élevé est évidemment un signe d'une situation financière saine.

Désignation	2005	2006	2007
Capital Circulant	67 039.00	34 625.00	17 281.00
Total des Actifs (immobilisations)	175 813.00	325 454.00	300 519.00
Profits Retenus	1 516.00	4 592.00	1 569.00
Profits avant frais financiers et impôts	2 435.00	4 874.00	12 517.00
Capitalisation Boursière	10 000.00	10 000.00	10 000.00
Valeur Comptable des Passifs (Dettes)	86 769.00	208 588.00	155 297.00
Ventes	969 135.00	1 029 283.00	1054 204.00

TABLE III.1 – Données financière Entreprises Fondatechiques

Ratio X_i	2005	2006	2007	Coefficient a_i
X_1	0.3813085	0.1063898	0.0575039	1.2
X_2	0.0086228	0.0141095	0.005221	1.4
X_3	0.0138499	0.014976	0.0416513	3.3
X_4	0.1152485	0.0479414	0.0643927	0.6
X_5	5.5123057	3.1626067	3.5079446	0.999
Formule Altman	$Z = 1.2.X_1 + 1.4.X_2 + 3.3.X_3 + 0.6.X_4 + 0.999.X_5$			
Z-Score d'Altman	6.0913	3.3851	3.7568	

TABLE III.2 – Calcul les ratios et le score d'Altman.

III.1.2 Méthode statistique de construction du score

i) Score à une variable explicative [13]

Considérons d'abord une fonction score à une seule variable explicative (un seul ratio).

On étudie un échantillon aléatoire d'entreprises de tailles et d'activités comparables dont certaines ont fait faillite et d'autres ont survécu.

On cherche à définir une fonction discriminante $Z = a \cdot R + b$ en donnant à Z la valeur 0 si l'entreprise a été défaillante et la valeur 1 si elle s'est révélée saine.

On teste le pouvoir discriminant du ratio R parmi plusieurs ratios Possibles relevés un à trois ans avant l'époque clé faillites.

On essaie en premier lieu le ratio qui présente la meilleure corrélation avec Z .

Exemple numérique : La matrice A ci-dessous donne, pour un échantillon de 10 entreprises les valeurs affectées à trois ratios (R_1, R_2, R_3) ainsi que la défaillance ($Z = 0$) ou la non-défaillance de l'entreprise ($Z = 1$) observées posteriori.

Entreprises	Z	R_1	R_2	R_3
A	1	2	1	4
B	0	3	1	2
C	1	2	3	0
D	0	0	1	3
E	1	3	4	4
F	1	4	1	1
G	1	2	3	1
H	0	0	2	0
I	1	1	3	4
J	0	4	0	4
Moyenne	0.6	2.1	1.9	2.3
Ecart-type	0.490	1.375	1.221	1.616

TABLE III.3 – Matrice A

La meilleure corrélation entre les quatre variables prises chacune deux à deux figurent sur le tableau suivant :

	Z	R_1	R_2	R_3
Z	1.000	0.208	0.602	0.025
R_1	0.208	1.000	-0.232	0.167
R_2	0.602	-0.232	1.000	-0.137
R_3	0.025	0.167	-0.137	1.000

TABLE III.4 – Matrice des corrélation

La meilleure corrélation (coefficient = 0.602) est obtenue entre Z et R_2 . L'équation de la droite de régression (ou droite des moindres carrés) de Z en R_2 est :

$$\frac{Z-0.6}{0.490} = 0.602 \frac{R_2-1.9}{1.221} \rightarrow Z = 0.242R_2 + 0.141$$

La fonction $Z = 0.242R_2 + 0.141$ permet de prévoir la défaillance d'une entreprise. On calcule le score Z de l'entreprise en fonction de son ratio R_2 .

On établit ainsi une règle de décision :

- Si Z est supérieur à sa moyenne ($Z > 0.6$), on présume que l'entreprise est saine ;
- Dans le cas contraire, on présume qu'elle sera défaillante.

Vérifions la validité de cette règle de décision un sur les dix entreprises de l'échantillon. Nous constatons que la règle est validée dans 7 cas sur 10.

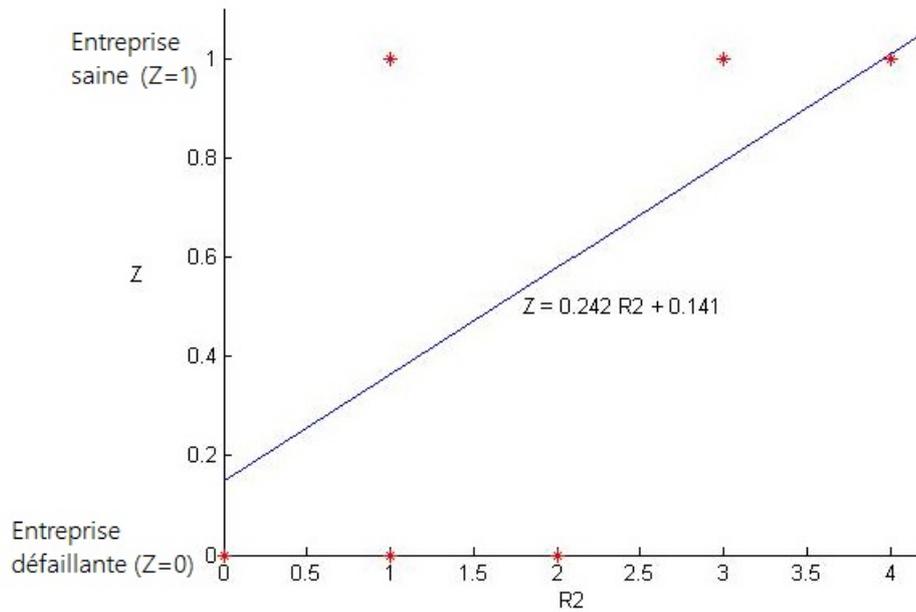


FIGURE III.1 – Régression de Z en R_2

ii) **Score à plusieurs variables explicatives** Les scores à une variables (voir ci-dessus) considèrent chaque ratio sans relation avec les autres.

En reprenant les données de l'exemple précédent déterminant l'équation d'un score qui soit fonction des trois ratios :

$$Z = a_1 \cdot R_1 + a_2 \cdot R_2 + a_3 \cdot R_3 + a_4.$$

. Il s'agit de trouver la valeur de chacun des quatre coefficient (a_1, a_2, a_3, a_4) qui minimise l'expression :

$$\sum_{i=1}^{10} (Z_i - a_1 R_{1i} - a_2 R_{2i} - a_3 R_{3i} - a_4)^2$$

(Méthodes des moindres carrés).

En annulant simultanément les quatre dérivées partielles

$$\frac{\partial Z}{\partial a_1} = 0; \frac{\partial Z}{\partial a_2} = 0; \frac{\partial Z}{\partial a_3} = 0; \frac{\partial Z}{\partial a_4} = 0.$$

On obtient un système de quatre équation à quatre inconnues dont les racines sont :

$$a_1 = 0.128; a_2 = 0.278; a_3 = 0.018; a_4 = -0.238.$$

La fonction score est $Z = 0.128R_1 + 0.278R_2 + 0.018R_3 - 0.238$ une entreprise est présumée saine si, pour cette entreprise, Z_i est supérieur à la moyenne de Z dans l'échantillon de dix entreprises ($Z_i > \bar{Z}$).

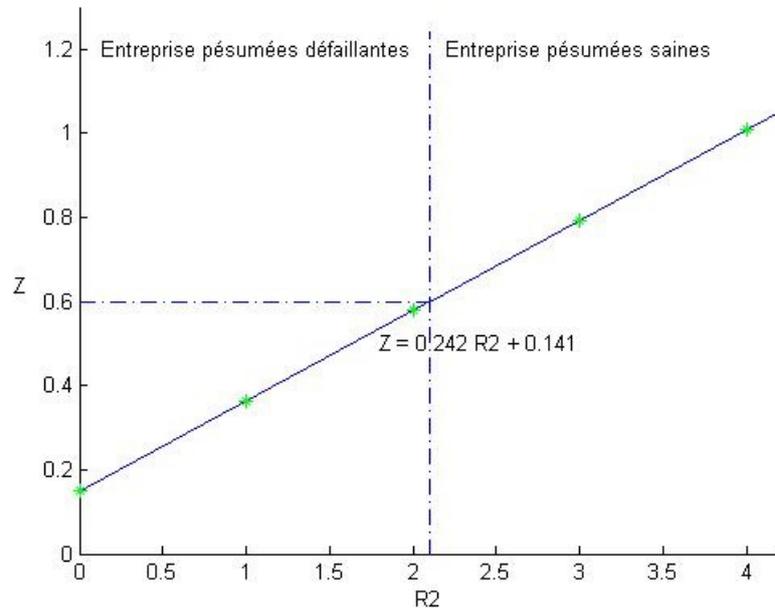


FIGURE III.2 – Application de la règle de décision

Comme $\bar{Z} = 0.602$ une entreprise est présumée saine si $Z_i > 0.602$

On remarque que les entreprises (A) et (G) sont mal classées par la fonction discriminante. Leur score est inférieure à 0.6 alors qu'en réalité, elles sont saines. La fonction score à trois variables ; 8 entreprises à classé correctement sur 10.

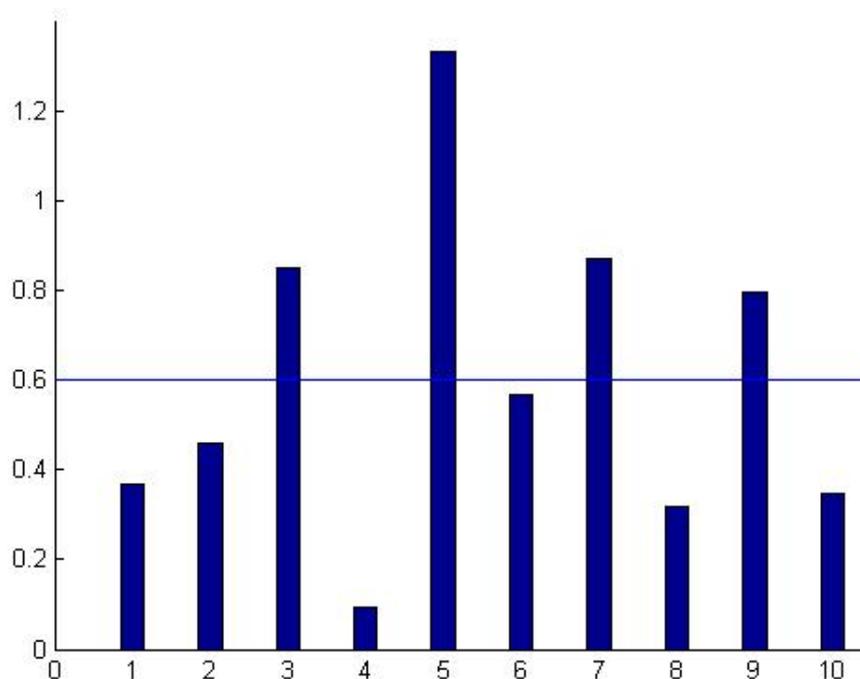


FIGURE III.3 – Le classement des entreprises d’après la valeur du score

III.2 Application méthode SVM

III.2.1 Ensemble des données

Les ensembles de données de crédit German sont disponibles sur le site <https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html.german.nu>. Sont adoptées ici pour évaluer le précision prédictive.

Les données crédit de German de notation il se compose de 600 cas de demandeurs solvables et 400 cas où le crédit ne devrait pas être prolongé. Pour chaque candidat, 24 variables d’entrée décrivent les antécédents de crédit,

III.2.2 Modèle mathématique

Parmi les 24 attributs on choisit deux attributs ; On à choisi le 1 ère et 2 ème attribut, En utilisant 500 exemples d’apprentissage,.

L’objectif de l’apprentissage par SVM est d’aboutir à une fonction de

Attributs	Nature	Description
Statut du compte courant	Qualitative	
Genre	Qualitative	Male, Female
Marié ou non	Qualitative	Oui, Non
l'objet du prêt	Qualitative	
Enf. à charge	Qualitative	0, 1, 2, 3+
renseignements personnels	Qualitative	Graduate, Not Graduate
Activité	Qualitative	Oui, Non
Revenu	Numérique	
le montant du prêt	Numérique	
Echéance du prêt	Numérique	
Historique Du crédit	Numérique	
Zone d'habitation	Qualitative	Urban, Semi-urban, Rural
Statut d'emploi	Qualitative	Oui, Non
Téléphone	Qualitative	Oui, Non

TABLE III.5 – La liste des attributs de l'ensemble de données à la base German .

décision de la forme :

$$H(x) = \sum_{i \in S} \alpha_i y_i K(x_i; x) + b$$

Où S représente l'ensemble des vecteurs supports, les α_i leurs coefficients, b est un biais et K le noyau utilisé.

Dans notre travaille, on à choisi des valeurs appropriées pour le paramètres de régularisation C et le paramètre du noyau RBF σ .

En utilisant une validation croisée par la recherche de la grille les valeurs de paramètres $C = 10$, $\gamma = 0.5$ et $\sigma = 0.707$.

On à utilisé le noyau RBF(Radial Basis Function) défini par : $k(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{\sigma}}$, où σ est une constante positive.

Le temps d'apprentissage 14.16s, obtenus à l'optimum de les coefficient de lagrange α , et le biais $b = 0.5583$.

La méthode SVM a donné un taux de reconnaissance de 70%. et la figure ci-dessus présente la phase d'apprentissage

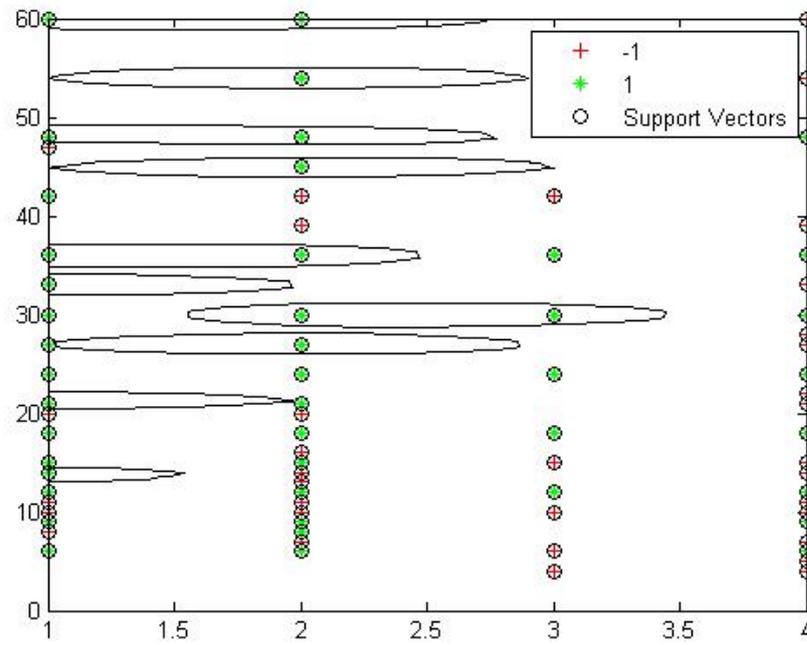


FIGURE III.4 – Représentation graphique de la surface séparatrice entre les deux classes par la base German.

D’après la méthode SMO(Sequential Minimal Optimization), existant dans la boîte outils SVM de MATLAB, on a obtenu la classification figure [III.2.2](#).

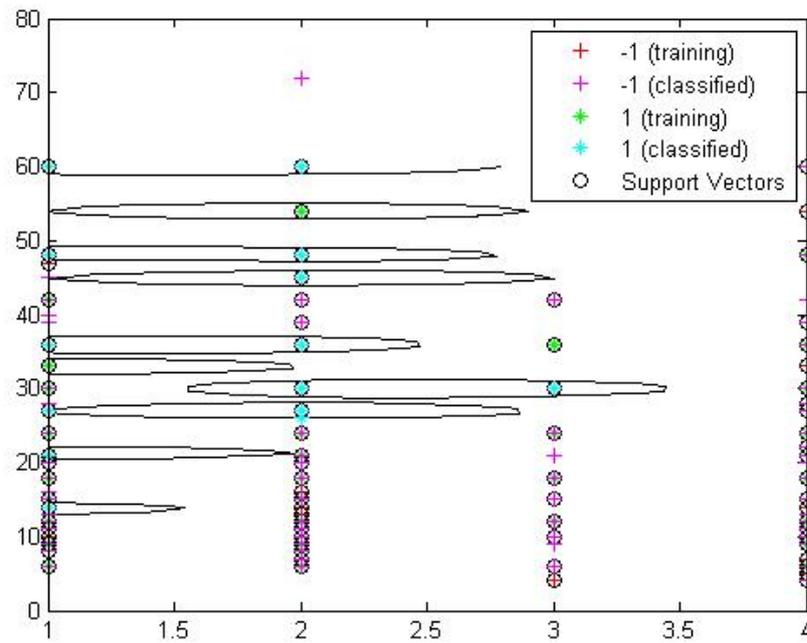


FIGURE III.5 – Classification des exemples de la base German.

Conclusion

Dans ce chapitre, on a présenté la méthode de scoring de modèle Altman et la méthode de statistique de construction le score dans MATLAB ; Ces méthode de scoring présente l'avantage de tester de les combinaisons de ratio les plus efficaces, pour prévoir la défaillance d'entreprise.

Et la méthode SVM applique dans crédit bancaire (Ex : German) intéressé à plusieurs attribues pour prédire les bon payeurs et les mauvais payeurs, donne des résultat plus précis.

CONCLUSION GÉNÉRALE

Le credit scoring est généralement considéré comme une méthode d'évaluation du niveau du risque associé à un dossier de crédit potentiel. Cette méthode implique l'utilisation de différentes techniques statistiques pour aboutir à un modèle de scoring basé sur les caractéristiques du client, il sera classé par le modèle comme : Bon Payeur ou Mauvais Payeur.

Par contre la méthode support vecteur machine consiste à projeter les données de l'espace d'entrée non-linéairement séparables (marge dure et marge souple) dans un espace de plus grande dimension appelé espace de caractéristiques de façon à ce que les données deviennent linéairement séparables. Dans cet espace, la technique de construction de l'hyperplan optimal est utilisée pour calculer la fonction de classement séparant les deux classes.

Les résultats de classification, montrent que plusieurs exemples sont bien reconnus par la méthode SVM et mal reconnus par la méthode scoring ceci nous a mène à dire que les SVMs sont meilleurs que Scoring.

BIBLIOGRAPHIE

- [1] A. BOUAZABIA, *Assistance, conseil et diagnostic de la PEM Fondatechnique. Mise en place du plan d'affaires et élaboration des outils décisionnels*, thèse de Master Méthodes Quantitatives pour l'Expertise et la Décision Économique, Université Lyon 2, 2008.
- [2] A. GUIZANI. *Traitement des dossiers refusés dans le processus d'octroi de crédit aux particuliers. Thèse de doctorat en Sciences de Gestion, L'institut Supérieur de Gestion, Université Sousse, Tunisie, 2014.*
- [3] A. REMACHE. *Conception et proposition de modèle de risque de crédit* mémoire master en Informatique, Université de Oum El Bouaghi, 2019.
- [4] B. LEGROS. *Mini-Manuel de Mathématiques pour la Gestion*, 2011.
- [5] B. COLASSE *L'analyse financière de l'entreprise*, La Découverte, 2008.
- [6] C. KHAROUBI, P. THOMAS. *Analyse du risque de crédit, Banque et Marchés* 2016 RB édition.
- [7] E.I. ALTMAN, A. SAUNDERS. *Credit risk measurement : Developments over the last 20 years-(1998)* , p.1721-1742.
- [8] E.I. ALTMAN. *Financial ratios discriminant analysis and the prediction of corporate bankruptcy*. Journal of Finance, 1968, 589-609.
- [9] E. MANCHON. *Analyse bancaire de l'entreprise*, Economica, 5ème édition (Collection Economica – Institut Technique de Banque), paris 2001.

- [10] F. BOUYACOUB. *L'entreprise et le financement bancaire*, édition casbah, 2000, p 17.
- [11] G.P. DUTAILLIS, J Hamel, P Roghi, L Logé. *Le risque du crédit bancaire*, Éditions Riber, 1967, pp 45.46.
- [12] M. Rouach, G. Naulleau, B. Tournier. *le contrôle de gestion bancaire et financière*, revue de banque, 1994, p 310.
- [13] G. LANGLOIS, M. MOLLET. *Manuel de gestion financière*, Édition FOUCHER, 2000.
- [14] G. SAPORTA. *Actualité de l'analyse des données*. Conservatoire National des Arts et Métiers, Paris, RCE-Multimedia, 2015
- [15] J. KHARROUBI. *Etude de techniques de classement Machines à vecteurs supports*, pour la vérification automatique du locuteur. Doctoral dissertation, 2002.
- [16] J. HULL, C. GODLEWSKI, M. MERLI. *Gestion des risques et institutions financières*, université de Toronto, Canada, Édition Pearson Éducation France, 2010.
- [17] J. FERRONIERE, E.CHILLAZ. *Les opérations de banque*, Dunod, 1963, page 187/190-192/193-196.
- [18] JP. NAKACHE, G. Celeux. *Analyse Discriminante sur Variables Qualitatives*. - Polytechnica, 1994.
- [19] L. ST-CYR, D. PINSONNEAULT. *Mesure et analyse du risque d'insolvabilité* École des hautes études commerciales, 1997.
- [20] M. BARDOS. *Analyse discriminante, application au risque et scoring financier*, Dunod, 2001 Ouvrage de niveau 2^{ème} cycle universitaire.
- [21] M. DIETSCH, J. PETEY. *Mesure et gestion du risque de crédit dans les institutions financières*, Édition Revue Banque, Paris 2003.
- [22] M.F. VIDAL and F. BARBON. *Credit scoring in financial inclusion*, 2019.
- [23] M.M. Guigal. *Le guide de la banque, Comptes, carte bancaire, services*. Edition ComprendreChoisir.com, Paris 2011.
- [24] N. WELLINK. *Au delà de la crise : la réponse stratégique du comité de Bâle*. Banque de France, Revue de la stabilité financière, 2009, 13, 131-141.

- [25] P.L. GONZALEZ. *Calcul d'un score (scoring) Application de techniques de discrimination* 2019.
- [26] P. VINCENT. *Modèles à noyaux à structure locale*, grade de Philosophie Doctor , Université de Montréal, Canada, 2003 .
- [27] P. NARASSIGUIN. *Monnaie, banques et banque centrales dans la zone euro*, De Boeck Supérieur, 2004.
- [28] R. RAKOTOMALALA. *Pratique de la Régression Logistique- Régression Logistique Binaire et Polytomique*. Université Lumière Lyon 2, (Version 2.0), 2009.
- [29] R. SUBLET. *La gestion du risque de crédit bancaire sur les portefeuilles professionnels et particuliers*, Mémoire de Bachelor of Business 2016, École de commerce de Lyon.
- [30] S. DJEMAI. *Résolution des problèmes de classification SVM par la méthode adaptée. Thèse de doctorat en Mathématiques Appliquées, Université de Bejaïa, 2016.*
- [31] S. COUSSERGUES, G. BOURDEAUX, T. PÉРАН *Gestion de la banque-8e éd. : Normes et réglementation à jour, Nouvelles stratégies bancaires. Vol. 1. Dunod, 2017.*
- [32] V. BRUNEL. *Le Risque de crédit Bancaire*, Département Ingénierie financière - 2016.
- [33] W.B. English. *Risque de taux d'intérêt et marges d'intérêt nettes des banques*. Rapport trimestriel BRI, décembre 2002.
- [34] Z. CHIBEL, Z. BAMOUSSE, M. ELKABBOURI. *Prévision du risque de crédit : ambition du scoring analyse comparative des pratiques de crédit scoring*. Article International Journal of Management et Marketing Research (MMR), Université Hassan Premier.

Résumé : Notre travail consiste à appliquer des méthodes statistiques, afin d'évaluer le risque de crédit pour les particuliers et les entreprises. Parmi ces approches, la méthode de scoring estime le risque de crédit en prévoyant la solvabilité du demandeur de crédit. Les institutions financières utilisent ce modèle pour estimer la probabilité de défaut qui sera utilisée pour affecter chaque client à la catégorie qui lui correspond le mieux : bon payeur ou mauvais payeur.

La méthode d'apprentissage statistique SVM a connu ces deux dernières décennies un large développement des points de vue théorie et applications. Elle repose sur un fondement théorique solide et utilise le principe de maximisation de la marge. Les SVMs ont été utilisées avec succès dans plusieurs domaines, notamment dans l'évaluation des risques des crédits bancaires.

L'application de ces méthodes nous a permis à déterminer les variables qui séparent au mieux entre un bon et un mauvais client, et cela permet aux banques de faciliter la prise de décision concernant l'octroi de crédit.

Mots clés : Évaluation de risque, crédit, scoring, apprentissage statistique, SVM.

abstract

Our thesis consists in applying statistical methods in order to assess the credit risk for individuals and businesses. Among these approaches, the scoring method estimates the credit risk by predicting the creditworthiness of the credit applicant. Financial institutions use this model to estimate the probability of default that will be used to assign each customer to the category that best suits them : good payer or bad payer.

The SVM statistical learning method has undergone a large development in theory and application over the past two decades. It is based on a solid theoretical foundation and uses the principle of margin maximization. SVMs have been used successfully in several areas, notably in the risk assessment of bank loans.

Applying these methods allowed us to determine the variables that best separate a good and a bad customer, and this allows banks to facilitate decision-making about granting credit.

Key words : Risque assessment, credits, scoring, statistical learning, SVM.