

L'anticipation de la défaillance des entreprises clientes des établissements bancaires marocains via la méthode de la régression logistique

M Abdelilah JEBBARI

Doctorant à la Faculté des Sciences Juridiques, Économiques et Sociales Agdal
Université Mohamed V – Rabat - Maroc

Dr Driss EL ZANATI

Professeur Habilité à la Faculté des Sciences Juridiques, Économiques et Sociales Agdal
Université Mohamed V – Rabat - Maroc

Résumé : La modélisation de la probabilité de défaillance des emprunteurs d'un établissement bancaire marocain par le biais du modèle de la régression logistique interpelle, d'ors et déjà, l'attention des académiciens et des professionnels spécialisés en finance. Le but étant d'étudier et analyser les différents points forts et points faibles de ce modèle dans la prédiction de défaut auprès d'un échantillon de 302 entreprises clientes des principales banques au Maroc, opérant dans différents secteurs d'activités économiques, choisi d'une manière aléatoire. A travers les résultats obtenus, il s'avère que l'indice de performance du modèle est de l'ordre de 0,983, qui est alors proche de 1. Ainsi, la courbe Receiver operating characteristic (ROC) du modèle logistique est au-dessus de celle du modèle aléatoire, ce qui permet de rejeter l'hypothèse nulle qui suppose que le modèle logistique est équivalent au modèle aléatoire. Toutefois, nous avons pu soulever certaines limites de notre modèle semi paramétrique susceptibles d'altérer l'évaluation d'un dossier de crédit dans un établissement bancaire. Par conséquent, on peut dire que le gestionnaire du risque de crédit doit prendre en compte d'autres paramètres afin de pallier aux insuffisances dudit modèle.

Mots-clés : Risque de crédit ; probabilité de défaut ; défaillance des entreprises ; classification ; validation du modèle ; la régression logistique ; la sélection des variables.

Digital Object Identifier (DOI): <https://doi.org/10.52502/ijesm.v1i1.155>



1. Introduction

Avant la prise de conscience de l'ampleur des risques dans les différentes organisations, leur gestion était consacrée uniquement pour se protéger contre les différentes pertes associées aux divers incidents survenus lors de la réalisation des opérations et des activités. C'est qu'en 1970 que la gestion des risques financiers a connu une grande révolution dans le domaine financier au moment où elle est devenue une préoccupation majeure dans la plupart des entreprises et particulièrement dans les banques.

La gestion des risques financiers a pour objectif de permettre aux entreprises de faire front aux risques et incertitudes auxquels elles se heurtent dans l'exercice de leurs activités. Ces risques sont omniprésents dans toutes les activités économiques et financières des banques baptisées « établissements de crédit », alors que le risque de crédit se présente comme étant le risque le plus pertinent qui accapare le pourcentage le plus élevé des risques nets pondérés¹.

A cet égard, suite aux crises financières qui ont touchées le monde, les régulateurs des pays dont le Maroc ont édicté une multitude de règles prudentielles dans le but d'assurer la stabilité et la pérennité du système financier. Ceci avait incité les banques à adopter un ensemble d'approches et de techniques leur permettant de prévoir et de modéliser la gestion des risques financiers, en général, et le risque de crédit en particulier.

Devant l'exacerbation des risques financiers, les établissements de crédit étaient dans l'obligation de développer un système interne d'évaluation, de gestion et de contrôle du risque de crédit en parfaite harmonie avec la réglementation prudentielle, et ce, dans le but de surmonter les points faibles des méthodes adoptées auparavant.

En fait, la présente étude exploratoire revêt une importance capitale pour le monde de la finance, puisqu'elle tente à élucider les axes saillants de la gestion du risque de crédit au niveau des établissements bancaires via des techniques scientifiques.

Afin d'examiner la pertinence du modèle de la régression logistique dans l'évaluation du risque de crédit au sein des établissements bancaires marocains, on a jugé judicieux, dans le cadre du présent papier, de tenter de répondre à la problématique suivante : « Est-ce que l'évaluation du risque de crédit en le modélisant via la méthode de la régression logistique permet-elle de prévoir la défaillance d'un demandeur de prêt lorsqu'il s'agit d'une entreprise ? »

A cet effet, en adoptant une méthodologie descriptive et analytique, l'objectif de ce papier est d'aborder l'approche dite de notation interne à travers l'élaboration d'un modèle fondé sur la régression logistique. Etant donné, que cette approche permet de discriminer les clients selon l'identification des classes de risques et le calcul des probabilités de défaut associées.

Pour se faire, il y a lieu de mettre en lumière les techniques de la modélisation du risque de crédit; de constituer un modèle semi-paramétrique relatif au calcul du score des demandeurs de prêts en choisissant comme échantillon de notre analyse 302 emprunteurs du groupe bancaire marocain et d'analyser, d'une manière détaillée, les résultats obtenus par le biais de la modélisation de la régression logistique

2. Aperçu général sur les techniques de quantification du risque de crédit

La quantification du risque de crédit était d'emblée une priorité pour les banques désirant d'accroître leur production en termes de crédits distribués. Afin de se fixer, d'une manière synthétique, sur l'évolution des principales recherches dans ce domaine, on a jugé opportun de faire un tour d'horizon sur les

¹ Rapport annuel Bank Al Maghrib sur la supervision bancaire, exercice 2019 p : 85

différentes techniques de quantification du risque de crédit comme étant des modèles ayant suscité plus de débat.

2.1 Les premiers jalons des modèles statistiques

Depuis les travaux de Mr Tamari(1964), plusieurs auteurs ont évalué le risque de défaut des entreprises en faisant appel à l'analyse financière. En 1966 Mr Beaver a mis en place un modèle qui recommande aux établissements de crédit d'analyser l'état de santé d'un demandeur de prêt en examinant séparément cinq ratios financiers, c'est une technique qui était pertinente au début mais elle présentait certaines limites. Certes, délibérer les ratios d'une manière indépendante ne peut pas donner une visibilité totale sur la santé financière d'une entreprise.

Afin de pallier aux insuffisances des méthodes précédentes, Mr Edward Altman(1968) a fait naissance à un modèle multidimensionnel qui permet de calculer la probabilité de défaut à deux ans, c'est une approche qui recourt à l'analyse discriminante de Fisher fondée sur un critère métrique. En exploitant simultanément plusieurs ratios à la fois, on arrive à construire une fonction appelée « Z-score » afin de classer une entreprise parmi l'un des deux groupes d'entreprises : « défailtantes » ou « non-défailtantes ».

Cependant, la méthode élaborée par Mr Edward Altman présente des insuffisances, dont notamment :

- Lorsque la base de données financières n'est pas significative, elle laisse outrepasser d'autres indicateurs comptables et extracomptables ;
- L'interprétation du Z-score n'est pas fondée sur une approche scientifique ;
- L'échantillon du modèle est très petit alors qu'une décision valable doit prendre en considération un nombre important d'entreprises défailtantes et saines.

En outre, les hypothèses sur lesquelles se base l'analyse discriminante de Fisher, comme la normalité des distributions des variables retenues, l'homogénéité des matrices variances-covariances entre les deux groupes et l'absence de corrélations entre moyenne et variance, présentent des limites majeures à l'utilisation de cette méthode pour le calcul du scoring bancaire.

Afin de pallier aux différentes carences des anciennes méthodes d'anticipation des risques financiers, les gestionnaires du risque de crédit se sont orientés vers les approches statistiques dont principalement la régression logistique.

2.2. La notion de régression logistique

En 1980 Mr Ohlson a recommandé l'application d'un modèle dichotomique appelé la régression logistique dans l'anticipation de la défaillance des entreprises, ce modèle modélise la probabilité de défaut à posteriori contrairement aux modèles paramétriques. En fait, les modèles probabilistes présentent un grand intérêt par rapport aux anciennes méthodes d'anticipation du risque de crédit et en particulier l'analyse discriminante linéaire. Ces modèles abolissent la condition de distribution normale multivariée des variables, ne font aucune différence entre les informations choisies c.-à-d. aucune hypothèse de normalité, n'exigent pas la condition d'homogénéité des variables, et permettent l'utilisation des variables explicatives dichotomiques et/ou des variables explicatives continues.

Le modèle dichotomique est un modèle statistique dont la variable expliquée y_i ne peut prendre que deux valeurs (0 ou 1) qui représentent la survenance ou non d'un événement donné, d'agissant dans notre cas, de la défaillance ou non d'un demandeur de prêt bancaire.

On suppose un échantillon de N entreprises, indicés $i=1, \dots, N$, on observe pour chaque entreprise si l'évènement de défaillance s'est réalisé dans le passé ou pas, et on note :

$$y_i = \begin{cases} 0 & \text{entreprise défailtante} \\ 1 & \text{entreprise solvable} \end{cases}$$

Ce codage aide à déterminer la probabilité de survenance de l'évènement de défaillance comme l'espérance de la variable y_i puisque :

$$E(y_i) = [Proba(y_i = 1)] * [1 + Proba(y_i = 0)] * 0 = Proba(y_i = 1) = p_i$$

Ceci nous permet de dire que l'espérance de la variable dichotomique y_i donne la probabilité de défaut d'une entreprise objet de l'étude en fonction de plusieurs variables explicatives (X_{i1}, \dots, X_{in}).

La régression logistique binaire se présente comme le modèle le plus utilisé et livrable pour l'analyse des variables discrètes et l'estimation de probabilités, il est utile pour prédire l'effet de plusieurs variables sur une variable dichotomique.

Cette technique vient pour corriger le problème de la régression linéaire multiple dans la mesure où cette dernière est définie théoriquement dans \mathbb{R} alors que la variable à expliquer ne prend que deux valeurs.

Dans le cas d'une entreprise i on suppose que :

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

Avec :

β : une constante

α : est le vecteur des coefficients d'une combinaison linéaire à estimer

ε_i : les erreurs supposées indépendantes, avec une moyenne nulle et une variance égale à 1

Si ε_i suit une loi logistique de fonction de répartition F on obtient le modèle Logit qui correspond à une fonction de densité logistique, la formule de calcul se présente comme suit :

$$F(x) = [1 + e^{-x}]^{-1}$$

Avec $F(x)$: une fonction de répartition

ou : x sont les données concernant les caractéristiques de l'emprunteur et la nature du prêt et $\theta_i = \alpha + \beta X_i + \varepsilon_i$

Si on suppose que μ_i suit une loi normale, on obtient donc un modèle PROBIT, cette fois-ci la fonction de répartition est celle de la loi normale centrée réduite $N(0,1)$.

$$P(Y) = \int_{-\infty}^Y \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

3. La modélisation du risque de crédit à travers la régression logistique

Dans cet axe nous allons étudier la modélisation du risque de défaut via la régression logistique pour des fins de prévision et d'évaluation de cette approche à prédire la défaillance d'un demandeur de prêt.

3.1. La Description des données de l'étude

Les données de l'étude sont extraites de plusieurs entreprises marocaines défaillantes et non défaillantes. Nous avons pris en considérations plusieurs informations quantitatives et qualitatives qui sont nécessaires

dans l'étude d'une demande de crédit afin d'éliminer toute subjectivité dans ce choix, sachant que plusieurs variables peuvent être déterminantes de la faillite d'une entreprise sans que cela soit intuitivement prévisible.

3.1.1. La distribution de l'échantillon

Dans notre modèle, nous avons collectés les indicateurs de 302 entreprises marocaines, clientes des banques marocaines sur le plan national dont 29 défailtantes et 273 saines qui exercent leurs activités dans plusieurs secteurs d'activités, détaillées dans le tableau ci-dessous :

Tableau 1: Répartition de l'échantillon selon la nature des entreprises

Nature des entreprises	Entreprises solvables	Entreprises en défaut
Industrie	27	7
Service	111	9
BTP	52	2
Commerce	73	10
Autres secteurs	10	1
Total	273	29

Source : Etabli par nos soins sur la base des données collectée

Notre travail a fait l'objet d'une étude sur les petites et moyennes entreprises pour la simple raison que cette catégorie est la plus touchée par la défaillance et mérite d'être analysée minutieusement.

3.1.2. La sélection des variables

Pour un meilleur résultat, notre étude a été réalisée à travers l'utilisation de 30 variables qualitatives et quantitatives (voir les annexes 1 et 2).

3.2. La construction d'un modèle logistique (logit) pour la prédiction de la probabilité de défaut

Cette étape consiste à construire un modèle logistique, ensuite nous allons effectuer un ensemble de tests de validation et de performance.

3.2.1. L'analyse du modèle de référence

Le modèle de référence n'inclut que la constante et néglige tous les autres paramètres de prédiction du risque de défaillance chez les demandeurs de prêts bancaires dans le but de déterminer les chances qu'un emprunteur fasse défaut en l'absence des facteurs de risques

Ce tableau décrit un modèle de référence qui exclut nos variable explicatives, il met en évidence le niveau de signification pour illustrer un récit de mise en garde.

Tableau 2: le classement des entreprises

	Observations	Prévisions		
		Décision		Pourcentage correct
		0,00	1,00	
Etape 0	Décision 0,00	0	29	0,0
	Décision 1,00	0	273	100,0
	Pourcentage global			90,4

Source : Etabli par nos soins sur la base des données collectées

Le tableau suivant nous donne l'estimation d'un seul paramètre (constante)

Tableau 3 : Variables dans l'équation

	A	E.S.	Wald	ddl	Sig.	Exp(B)
Etape 0 Constante	2,242	0,195	131,793	1	0,000	9,414

Source : Etabli par nos soins sur la base des données collectées

On remarque que le coefficient de la constante β est égale à odds ratios= 2,249 et son exponentielle est $e^\beta = 9,414$ qui est égale 273/29, cela veut dire que les chances qu'un emprunteur fasse défaut sans tenir en compte des variables est 9 fois plus élevé que les chances de ne pas faire défaut et un sig<5%, donc on peut dire que le modèle avec la constante seulement est un prédicteur statistiquement significatif.

3.2.2.L'analyse du modèle en fonction des variables

Nous allons analyser le modèle en introduisant les paramètres sus mentionnés et en faisant appel au coefficient de détermination qui est calculé selon ce qui est indiqué au niveau du tableau ci-dessous selon deux façon, la première de Cox et Snell et la deuxième concerne R² Nagelkerke.

Tableau 4: statistique de Cox et Snell

Etape	-2log-vraisemblance	R-deux de Cox & Snell	R-deux de Nagelkerke
1	53,087 ^a	0,367	0,782

Source : Etabli par nos soins sur la base des données collectées

Les valeurs de R² nous indiquent approximativement dans quelle mesure la variation dans le résultat est expliquée par le modèle, ce coefficient signifie que les paramètres du modèle affectent d'un pourcentage compris entre 37% et 78% le risque de défaillance chez un client demandeur de prêt.

3.2.1.La classification du modèle

L'évaluation du modèle logistique est estimée selon son classement ce qui est explicité par le tableau ci-après :

Tableau 5 : le taux du bon classement du modèle logistique

	Observations	Prévisions		
		Décision		Pourcentage correct
		0,00	1,00	
Etape 1	Décision 0,00	22	7	75,9
	Décision 1,00	4	269	98,5
	Pourcentage global			96,4

Source : Etabli par nos soins sur la base des données collectées

Ce tableau est utile pour la classification du modèle il est équivalent à celui du bloc 0. Lorsqu'on inclus les variables explicatives, on constate une amélioration remarquable, étant donné que notre modèle classe correctement le résultat pour 96,4% des cas comparativement à 90,4% dans le modèle nul.

3.2.1.La construction de l'équation et Interprétation des coefficients du modèle

Afin de se fixer sur les déterminants de la relation entre les variables explicatives et leurs coefficients, il s'avère nécessaire de les mettre en exergue au niveau du tableau suivant :

Tableau 6 : les variables dans l'équation

	A	E.S.	Wald	ddl	Sig.	Exp(B)
EBE_	1,085	0,527	4,248	1	0,039	2,960
Autonomie financière	0,232	0,059	15,623	1	0,000	1,262
Chiffre d'affaire	0,876	0,187	21,859	1	0,000	2,400
BFR	-3,839	1,093	12,333	1	0,000	0,022
Taux de défaut par secteur(1)	-0,947	0,371	6,527	1	0,011	2,579
Constante	-6,739	1,282	27,466	1	0,000	0,001

Source : Établi par nos soins sur la base des données collectées

Ce tableau permet de déterminer la relation entre les variables explicatives et leurs coefficients, on remarque que les variables EBE, autonomie financière et le chiffre d'affaire ont tous un signe positif. Ce qui signifie que toute augmentation des coefficients de ces variables augmentera la probabilité de l'entreprise à faire partie du groupe non défaillant. Alors que, les signes des coefficients du BFR et le taux de défaut par secteur d'activité sont négatifs cela signifie que toute augmentation du coefficient de ces variables augmentera la probabilité de l'entreprise à faire partie du groupe défaillant.

Par ailleurs, il faut signaler que les autres paramètres ont été éliminés du modèle pour la simple raison qu'ils ne sont pas statistiquement significatifs au seuil de 5%, c'est ce qui explique que ces ratios n'ont pas un grand impact sur l'équation de la régression logistique, alors que dans certains cas ils peuvent influencer négativement les prévisions. On peut résumer les résultats obtenus du tableau ci-dessus dans l'équation de la régression logistique suivante :

$$Y = \frac{\ln(p_i)}{(1-p_i)} = -6,739 + 1,085(\text{EBE}) + 0,232(\text{autonomie financière}) + 0,876(\text{chiffre d'affaire}) - 3,839\text{BFR} - 0,947(\text{Taux de défaut par secteur})$$

3.3. la validation du modèle de la régression logistique

L'objectif de cette étape est de tester la capacité de notre modèle à prédire l'état de santé des entreprises à travers la méthode ROC et Air sous la courbe.

3.3.1. La courbe ROC (Receiver Operating Characteristic)

La courbe ROC, qui est utilisée dans plusieurs domaines notamment la finance, permet de déterminer et de comparer la performance de plusieurs modèles, elle est constituée à travers :

- un axe vertical qui représente la sensibilité, c'est-à-dire le taux du vrai positif ;
- un axe horizontal qui désigne le 1-specificité, en d'autre terme, le taux du faux positif,
- Le classificateur aléatoire est représenté par la diagonale.

La sensibilité ou la proportion des vrais positifs indique le pourcentage des cas positifs correctement identifiés, tandis que la spécificité ou la proportion des vrais Sp indique les cas négatifs correctement identifiés. La formule de calcul de ces derniers se présente comme suit :

$$S_e = \frac{VP}{VP + FN} \quad \text{ET} \quad S_p = \frac{VN}{VN + FP}$$

Avec :

- VP : les résultats positifs chez les entreprises défaillantes
- FP : les résultats positifs chez les entreprises défaillantes

- VN : les résultats négatifs chez les entreprises non défaillantes
- FP : les résultats négatifs chez les entreprises défaillantes

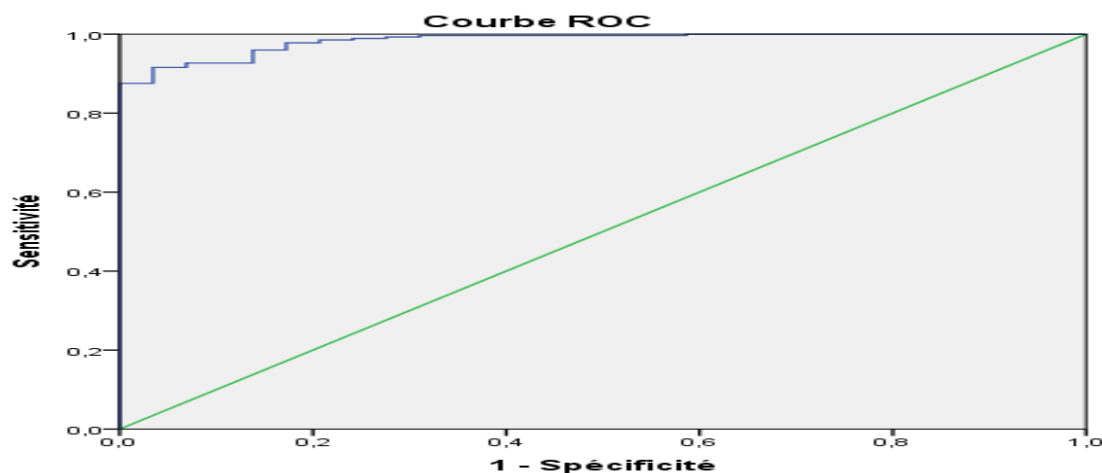


Figure 1: la courbe ROC du modèle logistique

A travers cette figure, on aperçoit que le modèle de la régression logistique est largement supérieur à la diagonale, et par voie de conséquence notre modèle est très performant par rapport modèle aléatoire.

3.3.2.L'air sous la courbe (AUC) et le ratio d'exactitude(AR)

Tableau 7: la zone sous la courbe

Zone	Erreur Std. ^a	Signif. asymptotique ^b	Intervalle de confiance 95% asymptotique	
			Borne inférieure	Borne supérieure
0,983	0,007	0,000	0,970	0,997

Source : Etabli par nos soins sur la base des données collectées

Le tableau ci-dessus reproduit la même conclusion obtenue via la courbe ROC, On remarque que l'air sous la courbe est de 0.983 avec une P-value inférieur à 5%, donc on rejette l'hypothèse nulle de ce modèle qui est équivalent au modèle aléatoire, ainsi le ratio d'exactitude est :

$$AR=2*AUC-1$$

$$AR=2*0.983-1=0.966=97\%$$

A travers les résultats précédents, on peut dire que la discrimination selon notre modèle est très excellente.

4. Interprétation des résultats et critiques de la régression logistique

Cet axe sera consacré à l'interprétation des résultats obtenus à travers notre modèle sur l'échantillon choisi, ensuite nous allons décortiquer les différents points faibles qui laissent du doute sur l'utilisation de cette méthode prédictive dans l'analyse des demandes de prêts au niveau bancaire.

4.1. Les interprétations des résultats

A travers les résultats obtenus, on peut dire que la régression logistique incite les établissements de crédit à l'utilisation de l'excédent brut d'exploitation, l'autonomie financière, le chiffre d'affaire, le besoin en fonds de roulement et le taux de défaut par secteur qui s'avèrent très déterminants dans l'explication du phénomène de la défaillance des entreprises.

On constate, aussi, par le biais de notre analyse que le score obtenu pour prédire la défaillance d'une entreprise dépend en grande partie du modèle adopté, de la taille de l'échantillon des entreprises choisi et de l'historique des données utilisées dans l'étude.

En effet, afin de tester la robustesse de notre modèle, on a fait appel à deux techniques importantes notamment la courbe ROC et l'aire sous la courbe. Ainsi, on remarque que la courbe relative à la probabilité de prédiction logistique est largement supérieure à la courbe de la diagonale, ce qui est confirmé par le coefficient de l'aire sous la courbe du modèle logistique qui est équivalent à 0,98, qui est proche de 1. Ceci nous a permis de conclure que le modèle logistique prédit mieux la défaillance des emprunteurs (entreprises).

4.2. Les critiques concernant la régression logistique

La fonction du score obtenu à travers l'utilisation de la régression logistique ne prend pas en considération diverses variables qui sont très discriminantes lors de leurs utilisation pour quantifier le risque de crédit, tels que : les fonds propres, les dettes fournisseurs, la capacité de remboursement, la trésorerie nette en plus des autres variables qualitatives.

Ainsi, ces variables peuvent être très révélatrices de la dégradation de la situation financière des entreprises demandant des crédits auprès des établissements bancaires. Ceci se présente comme un inconvénient pour l'adoption de cette méthode.

Dans le même sens, la régression logistique présente aussi la limite de présumer que les réponses sont non reliées.

Si les variables résultantes sont relatives à la période pendant laquelle les mesures sont prises (avant et après traitement), ou si les variables proviennent d'un groupe par couplage (chaque sujet du groupe expérimental est jumelé à un sujet du groupe contrôlé), la régression logistique n'est plus adéquate compte tenu des risques probables liés aux corrélations.

En outre, le modèle basé sur la régression logistique dépend principalement de l'échantillon initial pris en considération afin d'estimer les coefficients de la fonction discriminante. En conséquence, ce type de validation conduit très souvent à des résultats trop optimistes d'où la nécessité de choisir l'échantillon suivant une démarche scientifique laissant peu de place au hasard.

Partant de ce fait, plusieurs banques internationales considèrent que le score obtenu à travers la régression logistique ne permet pas d'avoir un jugement définitif sur la capacité d'un emprunteur à honorer ou pas ses engagements vis-à-vis de son créancier, la note obtenue fait l'objet d'un grand débat entre les personnes responsables pour décider de servir ou non le demandeur de prêt.

En effet, on peut comprendre pourquoi dans le contexte marocain les établissements bancaires s'intéressent dans leur évaluation des demandes de prêts à l'analyse financière.

Cependant, la méthode de l'analyse des paramètres qualitatives et quantitatives d'une manière isolée augmente le délai de traitement des dossiers et peut engendrer une incohérence entre les décisions prises par le même groupe bancaire, car ils sont, dans ce cas, individuelles et susceptibles de varier en fonction de la personnalité de chacun, ce qui ne permet pas de faire une estimation correcte du risque et pourrait augmenter le taux des créances en souffrance.

La méthode de la régression logistique est une substitution des techniques traditionnelles qui repose sur certaines hypothèses qui peuvent se présenter comme étant un grand obstacle qui empêche une quantification du risque de crédit. Ces modèles probabilistes présentent un grand intérêt par rapport aux anciennes méthodes, mais ils restent limités quant à la quantification de la probabilité de défaillance des entreprises.

4.3. Les perspectives de la quantification du risque de crédit

Selon notre étude, le résultat obtenu indique clairement que la régression logistique est une meilleure approche de classification des entreprises dans le but de prévoir le risque de crédit. Cependant, malgré la robustesse de cette technique, elle présente plusieurs limites qui peuvent impacter négativement l'analyse des données, ce qui explique que dans la pratique, le gestionnaire du risque de crédit complète les techniques de modélisation par une analyse financière de chaque paramètre d'une manière isolée, d'où la nécessité de réfléchir à d'autres techniques qui combinent à la fois la quantification du risque de crédit à travers les modèles statistiques et l'analyse financière proprement dite.

5. Conclusion

Nous avons remarqué, qu'en pratique, les gestionnaires du risque de crédit donne une grande importance à l'analyse financière proprement dite afin d'examiner une demande de prêt. Ainsi, le but de la modélisation par la technique de la régression logistique est d'étudier la possibilité de faire une confiance totale à cette méthode sans intervention de l'être humain afin de prédire une éventuelle défaillance d'un futur débiteur.

A travers les résultats obtenus, nous avons constaté que le modèle logistique est solide et capable d'offrir un score qui reflète au mieux la situation future des entreprises. Notre modèle a considéré comme pertinentes cinq variables explicatives à savoir : EBE, l'autonomie financière, le chiffre d'affaire, le BFR et le taux de défaut par secteur d'activité. En effet, ces résultats s'accordent parfaitement avec les recommandations de l'analyse financière classique.

A l'encontre de ce qui vient d'être dit, ce modèle de quantification du risque de crédit écarte plusieurs éléments essentiels à l'étude, conséquemment, l'abus de confiance aux modèles statistiques peut engendrer des dérèglements au niveau de l'étude d'un dossier de crédit dans un établissement de prêt.

BIBLIOGRAPHIE

- [1] A. BOUZZARA and al, "Evaluation du risque de défaillance de solvabilité des PME: une application du modele de la regression logistique," *Dirassat*, vol. 11, 2020, pp. 491–505.
- [2] Alexa lebel, "Le risque de crédit dans le marché hypothécaire du canada," *Memoire de maitrise a l'université du quebec en outaouais*, 2020.

- [3] Altman, E.I, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *Journal of finance*, vol. 23, pp. 589–609, Septembre 1968.
- [4] Antônio Alves Tôres Fernandes and al, "Read this paper if you want to learn logistic regression," *Revista de sociologia e politica*, vol. 28, pp. 1–19, 2020.
- [5] Arnaud de Servigny et Ivan Zelenko, "Le risque de crédit face à la crise," Edition Dunod, 2010.
- [6] Arnaud de Servigny et Ivan Zelenko, "Le risque de crédit, nouveaux enjeux bancaires," Edition Dunod, 2001.
- [7] Bardos, M, "Detecting the risk of company failure at the Banque of France," *Journal of banking and finance*, vol. 22, No. 10-11, 1998, pp. 1405–1419.
- [8] BASTIN Jean, "La défaillance de paiement et sa protection, l'assurance-crédit," 2ème édition L.G.D.J, 1999.
- [9] Berrada Mohamed azzedine, "Les techniques de banque de crédit et de commerce extérieur au Maroc," édition SECEA, 2007.
- [10] Comité de Bâle sur le contrôle interne, "Dispositif réglementaire mondial visant à renforcer la résilience des établissements et systèmes bancaires," version révisé, 2011.
- [11] Comité de Bâle sur le contrôle interne, "Ratios de levier et exigences de publicité," 2014.
- [12] Comité de Bâle sur le contrôle interne, "Révision de l'approche standard pour le risque de crédit," 2015.
- [13] DIETSCH Michel et PETEY joel, "Mesure et gestion du risque de crédit dans les institutions financières," RB Edition, 2003.
- [14] Douw Gerbrand Breed and all, "Developing an Impairment Loss Given Default Model Using Weighted Logistic Regression Illustrated on a Secured Retail Bank Portfolio," *Journal of risk*, vol. 13, pp. 47–62, 2019.
- [15] Ernest Yeboah Boateng and Francis T. Oduro, "Predicting Microfinance Credit Default: A Study of Nsoatreman Rural Bank, Ghana," *Journal of Advances in Mathematics and Computer Science*, vol. 26, pp. 1–9, 2018.
- [16] Hosmer, D. W. and Lemeshow, S, "A goodness-of-fit test for the multiple logistic regression model," *statistics in medicine*, vol. 16, pp. 965–980, 1997.
- [17] Hosmer, D. W. Lemeshow, S, "Applied Logistic Regression," 3rd Ed., John Wiley & Sons, 2013.
- [18] Idriss Abdelmajid Idriss Ahmed, Weihu Cheng, "The Performance of Robust Methods in Logistic Regression Model," *Open Journal of Statistics*, vol. 10, pp. 127–138, 2020.
- [19] Jean- Micheal Rocchi Michel Ruimy, "Les dérivés de crédit : risque de crédit solutions et évolutions," Edition Broché, 2004.
- [20] John Hull, Christophe Godlewski et Maxime Merli, "Gestion des risques et institutions financières," Edition Pearson, 2007.
- [21] KHAROUBI Cécile et THOMAS Philippe, "Analyse du risque de crédit : banque & marchés," RB Edition, 2013.
- [22] Manuel Díaz-Pérez, Ángel Carreño-Ortega, José-Antonio Salinas-Andújar and Ángel-Jesús Callejón-Ferre, "Application of Logistic Regression Models for the Marketability of Cucumber Cultivars," *Agronomy*, vol. 9, pp. 1–19, 2019.
- [23] Marco Tronzano, "Safe-Haven Assets, Financial Crises, and Macroeconomic Variables: Evidence from the Last Two Decades (2000–2018)," *Journal of risk and financial management*, vol. 13, pp. 49–69, 2020.
- [24] Mireille Bardos, "Analyse discriminante : application au risque et scoring financier," Edition DUNOD, 2001.
- [25] Ohlson, J. A, "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, vol. 18, pp. 109–131, 1980.
- [26] Rasa Kanapickiene, and Renatas Spicas, "Credit Risk Assessment Model for Small and Micro-Enterprises: The Case of Lithuania," *Risk*, vol. 10, pp. 1–23, 2019.
- [27] thierry Roncalli, "La gestion des risques financiers," édition Economica, 2004.
- [28] Tobias Berg an al, "On the Rise of FinTechs: Credit Scoring Using Digital Footprints," *The review of financial studies*, vol. 33, pp. 2845–2845, 2897.
- [29] Zhang, Lili, Herman Ray, Jennifer Priestley, and Soon Tan, "A Descriptive Study of Variable Discretization and Cost-Sensitive Logistic Regression on Imbalanced Credit Data," *Journal of Applied Statistics*, vol. 47, pp. 568–581, 2020.

Les annexes :

Annexe N°1 : Liste des variables quantitatives initialement retenus pour l'étude

Paramètres	Définition
Fonds propres	Il s'agit des sommes versées par les associés au profit de l'entreprise en ajoutant les profits générés annuellement et qui ne sont pas distribués sous forme de dividendes
Dettes fournisseurs	il s'agit du décalage qui peut exister entre la constatation d'un achat enregistré dans le compte de résultat et le décaissement du cash correspondant
Trésorerie nette	Il s'agit du liquide qui est à la disposition d'une entreprise pour satisfaire ses besoins d'achat a cours terme ce paramètre donne une idée sur le pouvoir d'une entreprise à couvrir ses besoins à court terme sans avoir recours au financement externe
EBE/IBE	Il représente le résultat provenant du cycle d'exploitation et constitue le premier solde significatif en termes de rentabilité
Autonomie financière	Mesure le degré d'autonomie financière de l'entreprise vis-à-vis des créanciers
Chiffre d'affaire	C'est le total des ventes
Endettement à CT	Il s'agit d'un montant emprunté pour une durée court (généralement inférieure à 1 an) dans le but de financer les emplois à court terme (éléments d'actifs à moins d'un an)
BFR	L'ensemble des besoins non financés par le passif circulant et indispensables à être financés pour le bon fonctionnement de l'entreprise
Actif net	L'actif net représente la valeur de tous les éléments qui constituent le patrimoine de l'entreprise
Capacité de remboursement	Mesure le nombre d'année nécessaires à l'entreprise pour assurer le remboursement total de ses dettes à terme au moyen de sa CAF
Résultat net	C'est la mesure comptable de l'enrichissement ou de l'appauvrissement de l'entreprise
Fond de roulement	Correspond au montant des ressources stables qui, après le financement de l'actif immobilisé, demeure disponible pour couvrir le besoins de financement
Rotation du stock	Ce ratio détermine la vitesse de rotation du stock. En effet, une rotation élevée peut être synonyme d'une entreprise dynamique sur son marché
résultat financier	Il représente la capacité de l'entreprise à équilibrer ses opérations de financement
Dettes bancaires et financière	L'ensemble des dettes bancaires à long terme ainsi que les dettes financières de l'entreprise auxquelles une entreprise fait appel, parallèlement au financement dont elle bénéficie par ses capitaux propres, pour financer son cycle d'exploitation ainsi que son cycle d'investissement.
Effectif	Nombre total des employés au sein de l'entreprise

Annexe N°2 : Liste des variables qualitatives initialement retenus pour l'étude

Paramètres	Définition
Taux de défaut par secteur	Ce paramètre exprime le pourcentage de défaut des entreprises par secteur d'activité, car un taux de défaut élevé d'un secteur impact négativement le taux de défaut des entreprises du même secteur
intensité concurrentielle par secteur	L'intensité concurrentielle est estimée par celle de chaque force au sein du secteur. Elle fait à la fois la rivalité entre les concurrents de l'entreprise, ses fournisseurs et ses clients qui ont un pouvoir de négociation
nombre barrière à l'entrée du secteur	Il existe des secteurs d'activités dont l'accès est conditionné par le respect de certaines exigences comme l'expertise technique, un niveau de capital donné, la réglementation et l'autorisations ce qui peut influencer le taux de défaut pour une entreprise
exposition risque nature et environnement	L'exposition aux risques naturels et/ou environnementales peut créer une instabilité de l'activité de la firme en particulier ceux du secteur agroalimentaire.
position concurrentielle	Une entreprise peut être un leader, un prétendant, un suiveur, nouvel arrivant, sortant ou un spécialiste ainsi sa situation dans le marché peut être le signe d'une solidité ou d'une fragilité de ses bénéficiers.
Concentration clientèle	L'entreprise peut être un faiseur de prix ou un meneur de prix et cette relation face à ses clients et fournisseurs est un élément essentiel pour assurer un bon fonctionnement de l'activité en conséquence, ce thème peut influencer la probabilité de défaut.
concentration fournisseurs	La consultation des clients peut être cruciale pour la survie de l'entreprise. En effet, l'entreprise peut avoir un seul client, un petit nombre de clients ou une diversification des clients. Dans le premier cas, l'entreprise dépend de son client car elle peut cesser son activité suite à la défaillance de son client, la non-conformité de ses produits et services aux exigences de son client ou suite à un conflit avec lui. En conséquence, la concentration du client peut influencer la probabilité de défaut
entreprise face à ses fournisseurs	L'entreprise peut être un fabricant ou un preneur de prix et de conditions de paiement avec ses clients et fournisseurs, ce qui lui permet de définir un ensemble de situations définies par le couple (position par rapport au client, position par rapport au fournisseur). Chaque situation peut avoir un impact sur la liquidité et le coût de produits et services. En conséquence, cette variable peut influencer la probabilité de défaut
existence plan de succession	L'existence d'un plan de continuité et de succession dans une entreprise se présente comme une garantie vis à vis de ses intervenants ce qui explique qu'elle est capable d'honorer ses engagements à long terme.
expérience PDG	Cette variable est exprimée en nombre d'ans, est un facteur important de la survie de l'entreprise. Une expérience élevée du PDG peut donner une idée sur la gouvernance au sein de l'entreprise notamment les conditions de conflit commercial, recouvrement de créances, relations bancaires etc. En conséquence, cela peut avoir un impact sur la probabilité de défaut
Ancienneté des principaux opérationnels	Cette variable inclut les tops managers dans l'évaluation de la probabilité de défaut. En effet, une expérience confirmée de ce dernier peut assurer les partenaires sur l'avenir de l'entreprise, notamment en l'absence du président et un plan de relève

Annexe N°2: (suite)

pourcentage du capital détenu par le dirigeant	Ce pourcentage dépend de la taille de l'entreprise, ainsi la concentration du capital dans une seule main ou dans une famille pour une grande entreprise représente un risque élevé car l'entreprise dépendra de la personne majoritaire.
respect remise document comptable	Le dépôt à temps des documents de synthèses peut être synonyme de fiabilité et de transparence pour le créancier.
nombre d'incidents provoqués 12 mois	Cette variable représente le nombre d'incident de paiement non régularisé. Un nombre élevé montre que l'entreprise a un problème de liquidité