



Le scoring des crédits d'exploitation des PME par le Machine Learning

تنقيط قروض الإستغلال للشركات الصغيرة والمتوسطة باستخدام تعلم الآلة

Scoring of SME operating credits by Machine Learning

Lyne Imene SOUDDA^{1*}, Manel MENARI², Yasser Moussa BERGHOUT³

¹Ecole supérieure de commerce (Algérie), ✉ etd_lsouadda@esc-alger.dz.

²Ecole supérieure de commerce (Algérie), ✉ etd_mmenari@esc-alger.dz.

³ Ecole supérieure de commerce (Algérie), ✉ ymberghout@esc-alger.dz.

Reçu le : 22/11/2022

Accepté le : 23/12/2022

Publié le : 31/12/2022

R É S U M É

Le développement technologique, la révolution numérique et le cumul du Big Data amènent la plupart des banques à moderniser leurs modèles de notation de crédit. Par ailleurs, en intelligence artificielle, le scoring de crédit fut historiquement l'un des premiers champs d'application du Machine Learning. De ce fait, l'objectif de ce travail est d'évaluer quelques modèles d'apprentissage automatique supervisé dans la classification des emprunteurs en emprunteur défaillant et non défaillant. Par le biais de trois modèles : RL, KNN et ANN, 282 dossiers de crédits d'exploitation sont entraînés à l'aide de 31 variables économiques. L'échantillon est composé exclusivement de PME algériennes. Dans la validation et la comparaison des modèles, des mesures comme la matrice de confusion et la courbe ROC sont calculées. Les résultats de l'étude montrent que la technique "neuronale" est meilleure en termes de prévision.

Mots clés : Scoring, Risque de crédit, Machine Learning, PME, Crédit d'exploitation.

Classification JEL : G21, C53.

* Auteur correspondant : Lyne Imene SOUDDA. Email ✉ etd_lsouadda@esc-alger.dz

الملخص

يقود التطور التكنولوجي والثورة الرقمية وتراكم البيانات الضخمة معظم البنوك إلى تحديث نماذج التنقيط القرضي الخاصة بها. علاوة على ذلك، في الذكاء الاصطناعي، كان تنقيط القروض تاريخياً أحد المجالات الأولى لتطبيق تعلم الآلة. لذلك، فإن الهدف من هذا العمل هو تقييم بعض نماذج تعلم الآلة الخاضعة للإشراف في تصنيف المقترضين المتخلفين وغير المتخلفين. من خلال ثلاثة نماذج يتم تدريب 282 ملف قرض استغلال باستخدام 31 متغيراً اقتصادياً. العينة مكونة حصرياً من شركات جزائرية صغيرة ومتوسطة. في التحقق من صحة النموذج والمقارنة، يتم حساب مقاييس مثل مصفوفة الارتباك. أظهرت نتائج الدراسة أن التقنية "العصبية" هي الأفضل في التنبؤ.

الكلمات المفتاحية: تنقيط، مخاطر القروض، التعلم الآلي، الشركات الصغيرة والمتوسطة، قروض الإستهلاك.

تصنيفات JEL: G21, C53

ABSTRACT

The technological development, the digital revolution and the accumulation of Big Data are leading most banks to modernize their credit rating models. Moreover, in artificial intelligence, credit scoring was historically one of the first fields of application of Machine Learning. Therefore, the objective of this work is to evaluate some supervised machine learning models in the classification of defaulting and non-defaulting borrowers. Through three models: RL, KNN and ANN, 282 operating credit files are trained using 31 economic variables. The sample is composed exclusively of Algerian SMEs. In model validation and comparison, measures like confusion matrix and ROC curve are calculated. The results of the study demonstrated that the "neural" technique is best in terms of prediction.

Key words : Scoring, Credit Risk, Machine Learning, SMEs, Operating Credit.

JEL Classification: G21, C53.

1. Introduction

Les banques par la nature de leur activité s'occupent principalement de l'octroi de crédit, elles sont ainsi confrontées au risque de non remboursement. Ce risque, notamment lorsque le crédit est important, peut causer la faillite totale d'une banque. D'une façon générale, le risque de crédit se définit comme étant le risque qu'un emprunteur manque à ses engagements. Ainsi, sa maîtrise et son maintien constituent la principale préoccupation des institutions financières.

Habituellement, l'approche générique de l'évaluation du risque de crédit consiste à analyser les dossiers des anciens clients de la banque, afin de trouver les caractéristiques d'un profil défaillant soit par des ratios financiers et comptables ou par des modèles linéaires, puis d'attribuer différentes notes de crédit aux entreprises ayant différentes probabilités de défaut, c'est la notation de crédit.

Aujourd'hui, l'intelligence artificielle est une composante existentielle de la finance. Ses progrès dans ce domaine sont remarquables. Elle a rendu la finance moins chère, plus rapide, plus grande et plus rentable, en particulier une de ses branches ; le Machine Learning. Cette discipline contribue à plusieurs applications : tarification des assurances, trading, détection des fraudes et gestion de portefeuille. Elle est notamment efficace dans la gestion de crédits, elle élabore des modèles fiables capables de détecter et de prédire avec précision les profils défaillants.

En Algérie, nous assistons à une croissance notable du nombre de PME, 4,8% entre 2020 et 2021.¹ Cette croissance a été possible grâce au financement du secteur bancaire. Néanmoins, seulement la moitié des crédits sollicités a été accordée². Cette méfiance des institutions financières est due en partie à

¹ Bulletin PME N° 39 : novembre 2021, Bulletins d'information statistique de la PME, <https://www.industrie.gov.dz/> 24/04/2022 à 14 : 00.

² Bulletin PME N° 33 : novembre 2018, Bulletins d'information statistique de la PME, <https://www.industrie.gov.dz/> 24/04/2022 à 14 : 15.

l'absence de marché financier, d'où un manque d'informations et un manque de garanties. Il en résulte, la nécessité de développer un outil de gestion de ce risque.

1.1. La problématique de recherche

L'objectif de la présente étude, est l'adaptation de différents modèles d'apprentissage automatique supervisé dans la classification des demandeurs de crédit. Ainsi, notre article s'inscrit dans le contexte de développement de méthodes modernes et puissantes pour la gestion du risque auquel sont confrontées les banques, plus précisément, le risque de non remboursement des PME. Dans cette perspective, la problématique de recherche est formulée de la manière suivante :

« Dans quelle mesure les algorithmes du Machine Learning sont capables de modéliser le risque de crédit d'exploitation des PME ? »

1.2. Méthodologie de la recherche

Nous avons observé dans l'ensemble des études antérieures que les chercheurs s'intéressent à la comparaison de différents modèles automatiques supervisés dans le domaine de scoring des crédits. Toutefois, certains les ont complétés par l'optimisation de ces modèles par des méthodes de réduction de dimensionnalité et différentes techniques de prétraitement de données. Cette dernière démarche est celle que nous souhaitons adopter afin de vérifier nos hypothèses.

D'abord, une base de données constituée de 282 dossiers de crédit d'exploitation est recueillie auprès de la direction des petites et moyennes entreprises du CPA. Par la suite, ces données sont traitées et l'analyse en composantes principales (ACP) est introduite pour trouver le nombre optimal de caractéristiques pour une prédiction précise, cela permet d'utiliser plus efficacement les ressources limitées disponibles.¹³ modèles sont ainsi configurés, testés et évalués. Une fois arrivé à la phase de modélisation, trois algorithmes d'apprentissage supervisé sont mis en place sur un ensemble de 19 variables explicatives, il s'agit des algorithmes de

régression logistique, les k voisins les plus proches et les réseaux de neurones artificiels. Avant leurs implémentations, ces modèles de notation de crédit doivent être évalués pour leurs cohérence et précision. Ce processus est appelé la validation, il est généralement entrepris par une partie indépendante du processus de développement et est réalisé à partir de données qui n'ont pas fait partie de la phase d'apprentissage du modèle. Dans cette validation, nous avons opté pour les mesures ROC ASC et la matrice de confusion.

1.3. Structure globale de l'article

Cet article est structuré comme suit : la section 1 fournit une revue détaillée de la littérature. La section 2 décrit les principales étapes méthodologiques utilisées dans notre recherche, celle-ci comprend une présentation de la base de données, des variables utilisées et des méthodes appliquées. La section 3 sera consacrée à l'application des méthodes RL, KNN et ANN dans la classification des bénéficiaires de crédit d'exploitation par le biais de l'organisme bancaire algérien, le CPA. Elle vise à mesurer et comparer la performance des différents modèles ainsi qu'à leur optimisation et sera conclue par une discussion des résultats obtenus.

2. Évaluation des performances du ML dans la gestion des risques de crédit : Revue de la littérature

Dans le milieu des années 80, de nombreuses études académiques ont cherché à évaluer la performance prédictive des méthodes de ML par rapport aux méthodes plus classiques comme la régression logistique dans la gestion du risque de crédit.

Dans l'objectif de synthétiser les modèles modernes de ML, Thomas (2000) propose la première étude comparative des modèles de scoring supervisés. L'auteur reporte le pourcentage de classification correcte de six méthodes dont : les arbres de décision, les réseaux de neurones, la régression logistique, la régression linéaire, ... Les résultats indiquent qu'aucune

méthode ne domine les autres et que les différences entre les méthodes sont très faibles. Par ailleurs, une étude similaire réalisée par Baesens et al., (2003) à donner d'autres résultats. La publication propose une analyse de 17 algorithmes de classification à partir de huit ensembles de données sur des crédits de banques internationales. Les machines à vecteurs de support (SVM) et les réseaux de neurones ont affichés les meilleures performances prédictives avec des ASC de 66% et 91% respectivement.

Dans une autre étude sur le pouvoir prédictif des algorithmes ML dans la gestion du risque de crédit, Moscatelli et al. (2020) ont comparé la performance de classification de modèles traditionnels comme la régression logistique avec celle des modèles de ML. L'étude a révélé que les modèles d'apprentissage automatique procurent des gains significatifs en puissance et en précision discriminatoires par rapport aux modèles plus classiques.

En 2009, Yeh et Lien ont comparé le pouvoir discriminant de la probabilité de défaut de paiement entre six techniques d'exploration de données : LDA, RL, ANN, KNN, naïf bayes et DT. Les résultats ont montré que la prévision de la probabilité de défaut par l'ANN est la seule méthode qui pourrait être utilisée pour représenter la probabilité réelle de défaut. Par conséquent, cette technique devrait être utilisée pour déterminer les scores des clients plutôt que d'autres techniques d'exploration de données.

En Tunisie, Matoussi a analysé 1435 dossiers de crédit octroyés aux entreprises industrielles tunisiennes entre 2003 et 2006. Les modèles de scoring, RL et ANN sont comparés. Les résultats montrent la supériorité des réseaux de neurones artificiels par rapport aux autres méthodes classiques avec un taux de bon classement global de 89.9%.¹

Dans de plus petites dimensions, plusieurs éléments favorisent l'implémentation des techniques de scoring intelligent par les institutions bancaires lors de l'étude des crédits aux PME. Selon Berger et al. (2011), en

¹ MATOUSSI Hamadi et ABDELMOULA Aida Krichène : « *La prévision du risque de défaut dans les banques tunisiennes : Analyse comparative entre les méthodes linéaires classiques et les méthodes de l'intelligence artificielle : les réseaux de neurones artificiels* », Crises et nouvelles problématiques de la valeur, 2010.

1998, 62 % des plus grandes banques américaines utiliser déjà le scoring pour faire l'évaluation des demandes de financement des petites entreprises.

En 2020, Wang et al., ont publié une étude sur le processus de décision d'octroi de crédit aux petites entreprises en utilisant des modèles classiques et des modèles supervisés d'apprentissage automatique, l'étude consiste à comparer les modèles : score- carte, RF, ANN ainsi que d'autres méthodes. Selon l'article, le modèle des forêts aléatoires booster à présenter le plus de précision.

Dans leurs étude, Ampountolas et al. (2021) ont procédé à une étude comparative entre différentes méthodes d'apprentissage supervisé sur un échantillon de 4450 dossiers de crédit d'exploitation sur une période de 4ans, en guinée. Plusieurs mesures de performances sont utilisées. Dans l'ensemble, les résultats ont montré que l'algorithme des arbres de décision était le plus efficace.

De nos jours, les solutions de l'IA s'activent à changer l'expérience client et diversifier l'offre de produits bancaires. Avec cette technologie, les banques auront plus de moyens pour se concentrer sur leurs clients et leur offrir plus de proximité. De grandes sociétés bancaires comme la BNP Paribas réfléchissent actuellement à la mise en place d'assistants virtuels pour mieux répondre aux besoins des clients, les assister dans leurs démarches et les orienter vers des produits mieux adaptés à leurs besoins.¹

3. Démarche méthodologique

Cette section présente la démarche adoptée dans cette recherche et énonce les outils de travail utilisés dans la notation de crédit. Bien qu'il n'y ait pas de meilleures techniques ou processus d'IA utilisées dans l'ensemble pour construire des modèles de notation, ce qui est le mieux dépendra du

¹ MILANA Carlo et ASHTA Arvind : « *Artificial intelligence techniques in finance and financial markets : A survey of the literature* », Strategic change, 2021.

problème en question, de la structure des données, des variables utilisées et de l'objectif de la classification.¹

3.1. Ensemble de données

Afin de réaliser une étude comparative en appliquant les algorithmes ML, des données secondaires ont été exploitées.

3.1.1. Présentation de la base de données

L'ensemble de données a été obtenu auprès de l'institution bancaire Algérienne « le CPA », il se compose de 282 demandes de crédit d'exploitation d'entreprise de taille PME entre la période 2015-2019.

La collecte de données a été réalisé comme suit : d'abord une base de données numérisée contenant des dossiers de crédit d'exploitation a été mise à notre disposition. Par la suite, un échantillonnage stratifié est conduit sur Excel. L'échantillon est composé de 80% de clients non défaillant et 20% de clients défaillant proportionnellement à la composition originale de la population.

La base de données comporte une multitude de variables financières et comptables. Nous avons synthétisé 26 ratios inspirés de la littérature dans un tableau représenté dans l'annexe n° 1.

3.1.2. Caractéristiques de l'ensemble de données

Les entreprises de l'échantillon utilisées dans cette étude sont des PME actives dans divers secteurs d'activités. Ainsi, les deux critères retenus lors de la formation de l'échantillon se présentent comme suit :

- Taille de l'entreprise : PME.
- Nature du crédit sollicité : crédit d'exploitation.

L'échantillon se compose exclusivement de PME car elles sont généralement sous capitalisées, elles ne disposent principalement que d'une source externe

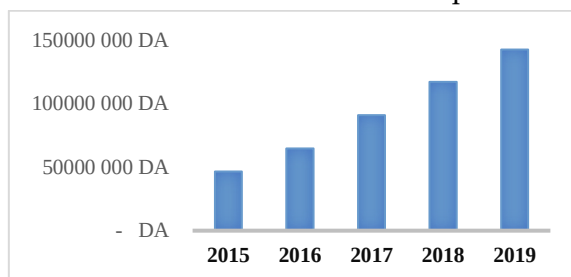
¹ MHLANGA David : « *Financial inclusion in emerging economies : The application of machine learning and artificial intelligence in credit risk assessment* », International journal of financial studies, 2021.

de financement : la banque. Par ailleurs, elles font face à la contrainte de production de l'information, cette contrainte surgit à cause du caractère familial d'une grande partie de ces entreprises ainsi qu'au défaut d'accès aux marchés financiers producteurs d'informations publiques et le poids négligeable de ce dernier en Algérie.

De plus, les PME ont une probabilité de défaillance nettement plus importante que les grandes entreprises.¹ Dans leurs articles « Estimation du risque de crédit et qualité de l'information comptable en Algérie » Gliz et Touati-Tliba (2012) annoncent : « le scoring est utilisé de façon intensive pour le crédit à la consommation et de plus en plus pour le crédit aux entreprises, en particulier celles de petite taille ».²

En effet, on constate une forte évolution des crédits d'exploitation octroyés par l'organisme bancaire, le CPA, entre la période 2015-2019.

Figure N° 1 : Évolution des crédits d'exploitation CPA



Source : Elaboré par les auteurs à l'aide d'une documentation interne du CPA.

3.1.3. Critère de défaillance des entreprises

Les clients de la banque sont classés en catégories de risque par le système de notation comportementale de la banque.

¹ NAHMIAS Laurent : « Impact économique des défaillances d'entreprise », Bulletin de la banque de France n°137, 2005.

² GLIZ Abdelkader et TOUATI-TLIBA Mohamed : « Estimation du risque de crédit et qualité de l'information comptable en Algérie », Les cahiers du cread n°98-99, 2012.

Afin d'obtenir deux classes d'entreprise (entreprise saine et entreprise défaillante), nous avons arrêté comme critère de défaut tout retard de remboursement de 90 jours à partir de la date de survenance de l'incident de paiement, et ce conformément aux exigences du comité de Bâle et au règlement 2014-03 relatif au classement et au provisionnement des créances.

3.2. Présentation des variables

Dans un premier lieu, 26 ratios économiques déduits d'un traitement comptable et 5 variables qualitatives sont retenus conformément à la littérature.

Au total, nous disposons de 31 variables explicatives résumées dans le tableau en annexe.

3.2.1. Prétraitement de l'ensemble de données

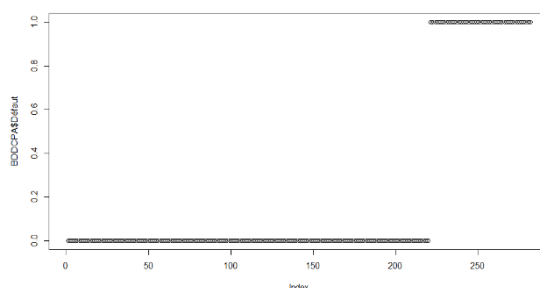
Ce processus préalable à l'apprentissage des modèles permettra une qualité supérieure des données de formation. Par conséquent, un modèle plus apte à effectuer des classifications correctes.

3.2.2. Codification de variables qualitatives

Une variable qualitative réfère à une caractéristique qui n'est pas quantifiable. Cette donnée n'a pas le même type d'analyse statistique qu'une donnée numérique, elle n'a pas d'échelle d'intervalle standardisée. Il existe une variété de système de codage. Dans notre cas, nous avons fait correspondre un chiffre pour chaque variable qualitative. De cette façon, elles peuvent être considérées comme entrées dans nos modèles.

Par ailleurs, le point de coupure sélectionné pour permettre de distinguer les deux groupes, qui est la variable cible, est intitulé défaut. Elle prend la valeur 0 si l'entreprise est saine et 1 si elle est défaillante. C'est la mesure adoptée par Altman (1968) ainsi que divers autres auteurs.

Figure N° 2 : Distribution de la variable défaut



Source : Elaboré par les auteurs à l'aide du logiciel R.

3.2.3. Valeurs manquantes (NA)

Les valeurs manquantes sont les observations manquantes de l'ensemble de données, leur total est de 12. La méthode de prétraitement choisie consiste en leur suppression.

3.2.4. Sélection de variables

Afin de sélectionner les caractéristiques de l'étude empirique, nous avons eu recours à l'avis d'un expert du secteur bancaire. Ainsi, le nombre de variables a été réduit de 31 à 19 variables explicatives, puis nous avons opté pour une réduction de dimensionnalité par la méthode ACP.

L'ensemble de variables qui sera utilisés par la suite est restreint à 19 variables, les motifs de ce choix seront discutés dans un tableau représenté dans l'annexe n° 5.

3.2.5. Création d'un ensemble de données d'entraînement et de test

L'étape suivante consiste à répartir les observations en un ensemble de formation et un ensemble de test. Le premier ensemble (80 % de l'ensemble de données) est utilisé pour concevoir divers modèles et élaborer des règles d'affectation d'un individu en fonction de ses caractéristiques. Quant au

deuxième (20 % de l'ensemble de données), il sert à vérifier si le modèle basé sur l'échantillon de formation est statistiquement fiable.¹

3.3. Modélisation

Nous allons présenter quelques aspects des trois méthodes utilisées dans notre étude ainsi que leurs principaux avantages.

Cette recherche va adopter une démarche expérimentale, du fait que plusieurs modèles seront testés et discutés et les meilleurs seront retenus pour la comparaison (annexe 5).

3.3.1. Méthodes de classification

Cette étape concerne la sélection d'un algorithme d'apprentissage automatique supervisé et l'adaptation de nos données d'entraînement à celui-ci. Cette analyse prédictive aura recours à différents algorithmes, et sera réalisée sur R.

Il est toujours conseillé d'avoir une variété d'algorithmes pour les différents problèmes, le même algorithme ne sera pas toujours le meilleur.²

A. La méthode de RL

La régression logistique permet d'étudier la relation entre une variable qualitative dépendante et des variables explicatives indépendantes à travers une transformation sigmoïde des probabilités.

La régression logistique sera utilisée en raison de ses aspects intéressants tels que la performance, le faible coût computationnel, la documentation, le background littéraire, la rapidité et surtout la transparence.

¹ ADDO Peter et al. : « *Credit risk analysis using machine and deep learning models* », Maison des sciences économiques, 2018.

² MHLANGA David : op.cit.

- **Réduction des dimensionnalités (ACP) :**

Afin de simplifier l'interprétation du modèle statistique RL et ses résultats, nous proposons la méthode ACP, dans laquelle les variables d'origine sont intégrées dans des vecteurs de plus faibles dimensions.

En réalité, ces composantes principales n'ont pas beaucoup de sens pour un spécialiste financier. Une solution possible est d'étudier la fonction de projection et indirectement relier la structure de chaque vecteur principal aux variables originales. Une interprétation de ces résultats (annexe 3) donne lieu à cinq vecteurs principaux :

Vecteur de trésorerie, vecteur/variable de rentabilité des actifs, vecteur de rentabilité commerciale, vecteur de liquidité et un vecteur de structure.

Interprétation du modèle

L'ensemble des variables sélectionnées étaient non significatives, à l'exception de trois variables et une constante : la variables mouvements confiés à le plus de signification statistique à hauteur de 0,1%, suivi par la constante à 1%, puis PC4 pour 5%, et finalement la variable Activité7 : import-export à 10% (annexe 4). Or, en raison de leur importance potentielle (annexe 1) et de l'objectif du machine learning qui tente de montrer que le pouvoir prédictif du modèle est prépondérant aux statistiques, nous les avons conservés dans le modèle.

Pour la validation statistique du modèle, la statistique Log vraisemblance (LR) est calculé :

$$LR = -2 \ln \left(\frac{\text{vraisemblance du modèle courant}}{\text{vraisemblance du modèle saturé}} \right) = 92,61$$

Cette statistique suit une distribution khi-deux à 10 degrés de liberté, elle est donc comparé à la valeur tabulée du khi-deux :

$\chi^2_{10}(0,05) = 18,31 < 92,61 \rightarrow$ Rejet de H_0 ¹ $\rightarrow$ Modèle valide statistiquement.

La statistique McFadden's ou Pseudo- R^2 est une mesure de la capacité d'un modèle d'apprentissage automatique à se généraliser à des données similaires à celles sur lesquelles il a été formé. Un modèle bien ajusté produit des résultats plus précis. Un modèle sur-ajusté correspond trop étroitement aux données. Enfin, un modèle sous-ajusté ne correspond pas assez étroitement.²

$$\text{Pseudo} - R^2 = 1 - \frac{\text{Ln}(\text{vraisemblance du modèle saturé})}{\text{Ln}(\text{vraisemblance du modèle courant})} = 0,364$$

Une valeur Pseudo - R^2 entre 0.2 et 0,4 indique un excellent ajustement du modèle³, cette statistique est de 0,364 dans notre RL, cela prouve l'ajustement du modèle.

B. La méthode KNN

L'algorithme KNN repose sur le calcul des distances entre des paires d'observations, il est utilisé dans les problèmes de classification où les données de formation sont disponibles avec des valeurs cibles connues. La valeur de k indique le nombre de voisins les plus proches, une façon de le trouver consiste à tracer un graphique contenant différentes valeurs de k et leurs taux d'erreur correspondant. Le nombre optimal des k voisins semble être autour de 33 voisins les plus proches.

Parmi les motivations du choix de ce modèle, nous avons notamment la simplicité de son exécution et de son interprétation. De plus, il est libre de toutes hypothèses de départ. Ainsi, l'avantage majeur de cette approche est qu'il n'est pas nécessaire d'établir un modèle prédictif avant la classification, d'où sa désignation de modèle paresseux.

¹ H_0 : cette hypothèse suppose que le modèle composé uniquement de la constante est plus significatif que le modèle complet.

² <https://www.datarobot.com/wiki/fitting/#:~:text=Model%20fitting%20is%20a%20measure,doesn't%20match%20closely%20enough>, 13/05/2022 à 7 : 50

³ <https://stats.stackexchange.com/questions/82105/mcfaddens-pseudo-r2-interpretation>, 11/05/2022 à 10 : 45.

C. La méthode ANN

Les réseaux de neurones artificiels utilisent des équations pour développer successivement (d'une couche à une autres) des relations significatives entre les variables d'entrée et de sortie à travers un processus d'apprentissage. Nous avons appliqué des réseaux de rétropropagation pour classer les données.

Les ANN peuvent facilement gérer les effets non linéaires et interactifs entre les variables. Ils seront utilisés en raison de leur nature numérique, de l'absence de toute exigence concernant les hypothèses de distribution des données (pour les entrées) et leurs capacités de mettre à jour les données sans retraiter les anciennes.

Lors de cette recherche, plusieurs variétés de réseaux de neurones sont testées en faisant varier le nombre de couches et le nombre de neurones cachés dans chaque couche afin de choisir la meilleure architecture qui présente un taux d'erreur minimal, puisqu'il n'existe aucune loi ou théorème qui permettrait de déterminer le nombre de couches cachées et le nombre de neurones optimal à placer dans la couche cachée.¹

En effet, dans notre programme, nous avons fixé un nombre de couches cachées égal à 2 composés de 5 et 1 neurones respectivement. Car suite à quelques simulations, ce modèle a montré le plus de précision (annexe 2).

3.3.2. Mesure de performance

C'est une étape très importante dans l'analyse prédictive. Il s'agit de la vérification de l'exactitude des valeurs prédites dans l'échantillon test pour savoir si elles correspondent aux valeurs connues dans l'échantillon d'apprentissage. Il existe plusieurs mesures pour comparer la performance des modèles, notamment ASC, Gini, fonction logLost, matrice de confusion

¹ ADDO Peter et al. : op.cit.

et RMSE. Dans cette recherche, nous mettrons d'avantage l'accent sur les critères de la matrice de confusion et la courbe ROC/ASC.

La comparaison du pouvoir prédictif des trois modèles, RL, KNN, ANN montre la performance de la technique neuronale par rapport aux deux autres modèles. En effet, le pourcentage de bon classement issu de l'application des réseaux de neurones artificiels est meilleur que celui présenté par RL et KNN. Leurs pourcentages respectifs sont de : 62,7%, 80,3% et 87,1%.

Par ailleurs, l'étude des matrices de confusion révèle que les modèles RL et ANN se ressemblent en précisions, ainsi, ils affichent une Accuracy de 81%. Alors que le modèle KNN peut trouver les positifs des positifs dans 74% des cas.

4. Discussion des résultats

Dans ce qui suit, l'analyse qui a permis de faire le choix de l'ensemble des variables correspondant le mieux à nos données.

Nous allons donc passer en revue 13 configurations de modèles (annexe 5). Ils seront examinés en termes de méthodes utilisées, de mesures de performance ainsi que du nombre de variables et de prétraitement. Deux classes de clients sont prédites et 3 algorithmes : M2, M5, M13, sont sélectionnés. En effet, ces algorithmes ont montré les meilleures précisions.

Dans un modèle KNN, on remarque que la surface sous la courbe ROC atteint son optimum quand le nombre de variables est de 19. Les résultats pour chaque groupe de variables sont respectivement : 80%, 77% et 71% pour un nombre de variables et/ou vecteurs de 19, 31 et 10.

Tandis que, la matrice de confusion indique que la précision varie entre 76% et 90%, avec plus de précision pour le modèle M1. Cela signifie que, pour un seuil de 50%, le nombre de client classé défaillant quand ils sont réellement défaillant par apport au total des observations est de 90%. Donc, le pouvoir

discriminant de l'algorithme KNN est plus pertinent avec 19 variables explicatives seulement.

Les modèles M4, M5 et M6 sont des algorithmes de réseaux de neurones, les deux mesures, ASC et Accuracy, indiquent que le modèle le plus performant en prédiction est M5, pour un ensemble de 19 variables explicatives.

Remarquons que tous les modèles de KNN de ANN ont subi un prétraitement de valeurs manquantes. Ce prétraitement est une étape incontournable du fait que les packages utilisés dans leur modélisation n'acceptent les NA.

La régression logistique est ensuite appliquée sur 6 modèles d'ensembles de données différents. On remarque que le premier modèle, M7, n'a pas convergé, probablement en raison de l'hétérogénéité des intervalles des caractéristiques, car, une fois normalisé, il aboutit à des résultats dans M8.

On remarque une grande amélioration de performance entre les modèles 9 et 10, suite au prétraitement des NA. Leurs suppressions à nettement amélioré le modèle.

Malgré la performance du modèle M12, c'est M13 qui est sélectionné. Et ce, à cause de la nécessité du traitement des NA pour permettre la comparaison des 3 méthodes.¹

Par ailleurs, on observe que KNN a montré la précision (Accuracy) supérieure dans la majorité des cas pour les 3 ensembles de variables explicatives.

Pour la mesure ASC, ANN présente une meilleure performance dans le modèle M5, de 87,1%. Mais perd en pouvoir prédictif dans M6 une fois l'ACP utilisée. Dans le cas contraire, l'introduction de la méthode de réduction des dimensionnalités sur le modèle RL a amélioré tous ses résultats pour les deux

¹ Rappelons-nous que les algorithmes ANN et KNN n'acceptent pas les NA.

méthodes de validation, par exemple, une ASC optimale dans M12 suivi d'une haute précision.

Bien que RL a été considéré comme le modèle le moins prédictif en termes de mesure ASC, on remarque qu'il a montré des pourcentages élevés et réguliers de précision.

La précision des modèles de machine learning peut avoir des comportements très différents, ce qui rend difficile de juger quel est le modèle le plus performant. De nombreux classificateurs ont une ASC élevée (c'est-à-dire supérieure à 70 %), mais leur précision peut varier considérablement. Par exemple : M1, M2, M3, M4. Par ailleurs, le contraire est vrai, plusieurs modèles présentent des précisions très favorable, mais leurs ASC ne le sont pas toujours : M5, M6, M8, M10, M11, M12, M13.

Cela peut être due au seuil appliqué dans ces deux méthodes. Tandis que la précision issue de la matrice de confusion applique par défaut un seuil de 50%, les probabilités issues de chaque modèle seront comparées avec ce seuil et seront ainsi classé en 0 ou 1. La courbe ROC ASC applique une variété de seuil (toutes les combinaisons comprises entre 0 et 1) pour comparer cette probabilité. Le seuil optimal, tranchant les classes, varie d'un secteur d'activité à un autre. Il est difficile de dire quel est le meilleur seuil à appliquer au risque de crédit d'exploitation. Une manière de se faire est de procéder à une validation croisée sur différents seuils.

Ainsi, les réseaux de neurones artificiels semblent constituer un outil de prévision puissant en matière de gestion de risque de crédit d'exploitation. Ce travail confirme ainsi les études empiriques antérieures (Oden et Sharada (1990), Kerling et Podding (1994), Abdou et al, (2008)).

Les modèles des réseaux de neurones artificiels sont de plus en plus utilisés en scoring. Cependant, bien que ces nouvelles méthodes soient intéressantes et parfois plus performantes que les techniques statistiques traditionnelles, elles sont moins déchiffrables, les réseaux de neurones sont incapables

d'expliquer les résultats qu'ils fournissent, ils se présentent comme des boîtes noires dont les règles de fonctionnement sont inconnues.¹

En termes d'interprétation des pondérations, la régression logistique semble être plus performante.

En effet, dans un réseau de neurones artificiels, les liaisons internes n'ont pas de signification économique. Les pondérations des ratios figurant dans les fonctions logistiques sont par contre transparentes et faciles à interpréter du point de vue de l'analyste financier.

Ainsi, la comparaison finale s'est effectuée sur les modèles optimisés : M2, M5, M13, opérant avec un ensemble de variables de 5 composantes principales et 5 variables qualitatives dans la RL et un ensemble de 19 variables explicatives dans les deux autres algorithmes. A noter que les 19 variables sont exploitées dans les 3 modèles, même dans l'ACP, la dimension est réduite car une rotation de l'espace de données est faite. Mais rappelons-nous que toutes les variables rentrent dans la composition des 5 vecteurs principaux, ce sont les pourcentages de contribution qui varient.

5. Conclusion

Le financement bancaire est la principale source de financement externe sollicitée par les PME. Ce financement n'est toutefois pas facilement accessible, l'évaluation des institutions bancaires du risque de crédit est souvent plus sévère pour cette taille d'entreprise en raison de leur susceptibilité à refléter une éventuelle faillite, et l'instabilité de leurs chiffres. Ainsi, la maîtrise de ce risque demeure une préoccupation majeure et un objectif recherché par les emprunteurs de crédit. En effet, la mise en place des meilleurs mécanismes, les plus modernes, doit être envisagée. Pour y parvenir, les modèles ML peuvent constituer un outil d'aide à la décision pour le conseiller bancaire. Étant donné que l'objectif de cet article est de vérifier

¹KENNEDY Kenneth : « *Credit scoring using machine learning* », Doctoral thesis, Technological University Dublin, 2013.

la capacité des modèles d'intelligence artificielle à prédire le risque de crédit vis-à-vis des demandeurs de prêt d'exploitation afin d'atteindre la probabilité de défaut de chaque demandeur, des modèles intelligents de l'apprentissage supervisé sont sélectionnés, entraînés, validés et évalués à travers différentes métriques d'évaluations. Par ailleurs, l'importance de ce crédit est reflétée par son rôle vital tout au long du cycle d'exploitation des entreprises. Il représente 10% de l'ensemble des crédits accordés par le CPA en 2019.¹

A travers 13 configurations de modèles supervisés de classification appliquer sur les données de l'organisme bancaire algérien le CPA, 282 demandes de crédit sont examinées à travers 31 variables économiques et financières. Les techniques de prédiction utilisées dans l'étude empirique sont RL, KNN et ANN. Les résultats indiquent que le modèle le plus affirmé dans ce processus de notation est le réseau de neurones artificiels (87,1%), suivi du modèle des K voisins les plus proches (80,3%) et finalement, la régression logistique (62,7%). Pour des précisions respectives de 81,03%, 74,24% et 81,39%.

La procédure de prétraitement des données à travers la méthode de réduction des dimensionnalité ACP ainsi que la codification des variables et la suppression des valeurs manquantes a eu un impact manifeste sur la performance des modèles, cela nous a permis de poursuivre l'étude comparative avec des modèles optimisés.

De cette étude, nous avons déduit que l'approche neuronale et la régression logistique sont deux méthodes complémentaires. Tandis que la régression logistique nous permet de sélectionner l'ensemble de variables, le réseau de neurones permet une modélisation optimale. Ces informations peuvent aider les gouvernements, les investisseurs, les gestionnaires et d'autres parties prenantes dans la prise des décisions économiques. D'un point de vue Marketing, l'intelligence artificielle permettra aux banques d'identifier les préférences de leurs clients. Ainsi, elle renforcera la relation Banque-

¹ Figure N° 1 : Evolution des crédits d'exploitation CPA.

Client et améliorera « l'expérience client » en proposant des produits et des solutions de manière proactive.

5.1. Problèmes ouverts et travaux futurs

A travers cette étude, nous avons implémenter trois modèles d'apprentissage supervisé dans un contexte de classification des clients de la banque en classes binaire : défaillant et non défaillant. Les résultats nous ont amené à la conclusion de la primauté du modèle de réseau de neurones dans la bonne classification des créanciers. Néanmoins, cette étude ne s'est pas réalisée sans embuches, ces difficultés peuvent constituer des problématiques à de nouvelles études :

Dans un premier lieu, l'ensemble de donnée peut être étendu en tenant compte d'un plus grand nombre et d'une multitude de variables explicatives. Nous pouvons, par exemple, introduire des variables macroéconomiques tel que l'inflation. Les modèles pourraient également étudier l'influences des chocs économiques, tels que la pandémie mondiale COVID-19 sur la notation des crédits d'exploitation.

Des facteurs plus subjectifs doivent être pris en compte dans les critères de notation, comme la réputation, la longévité de la relation avec la banque, l'entente avec le dirigeant de la banque, ...

Par ailleurs, la méthode de réduction de dimensionnalité (ACP) est appliquée au modèle RL. Il en résulte comme limite, la perte de 41% de l'information. Ça peut être la raison pour laquelle les résultats du modèle de régression logistique sont les moins efficace.

D'autre part, la codification des variable, le traitement des valeurs manquantes et des valeurs aberrantes, le traitement du déséquilibre de classe et bien d'autres méthodes de prétraitement de données peuvent être introduites ou optimiser dans le modèle, cette étape peut complètement changer les résultats et améliorer les modèles.

Finalement, on propose l'implémentation d'autres modèles supervisés et non supervisés pour de futures comparaisons.

Références bibliographiques

ADDON Peter et al. : « *Credit risk analysis using machine and deep learning models* », Maison des sciences économiques, 2018.

GLIZ Abdelkader et TOUATI-TLIBA Mohamed : « *Estimation du risque de crédit et qualité de l'information comptable en Algérie* », Les cahiers du cread n°98-99, 2012.

KENNEDY Kenneth : « *Credit scoring using machine learning* », Doctoral thesis, Technological University Dublin, 2013.

MATOUSSI Hamadi et ABDELMOULA Aida Krichène : « *La prévision du risque de défaut dans les banques tunisiennes : Analyse comparative entre les méthodes linéaires classiques et les méthodes de l'intelligence artificielle : les réseaux de neurones artificiels* », Crises et nouvelles problématiques de la valeur, 2010.

MILANA Carlo et ASHTA Arvind : « *Artificial intelligence techniques in finance and financial markets : A survey of the literature* », Strategic change, 2021.

MHLANGA David : « *Financial inclusion in emerging economies : The application of machine learning and artificial intelligence in credit risk assessment* », International journal of financial studies, 2021.

NAHMIAS Laurent : « *Impact économique des défaillances d'entreprise* », Bulletin de la banque de France n°137, 2005.

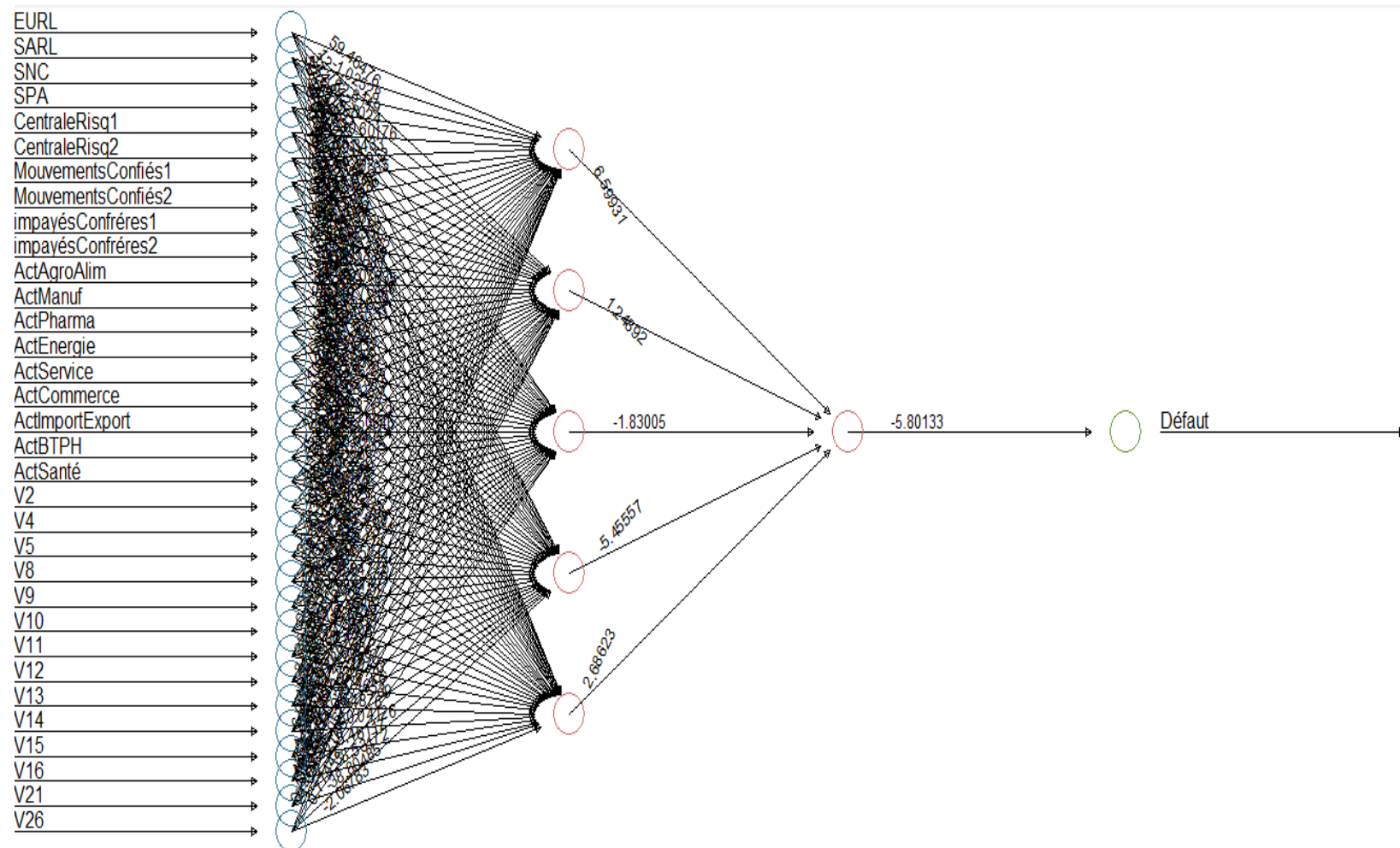
Annexe N° 1 : Présentation des variables quantitatives

Ratio	Nom	Calcul	Code	Reference
Trésorerie	Valeur ajoutée	VA / CA	V1	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Fonds de roulement	(FR x 360) /CA	V2	CHOW Jacky C. K, 2017.BERRAIH Radia, 2020.
	Besoin de fonds de roulement	(BFR x 360) /CA	V3	BERRAIH Radia, 2020.
	Délai de règlement client	(Créances clients x 360) /CA	V4	ELHAMMMA Azzouz, 2011. BERRAIH Radia, 2020.
	Délai de règlement fournisseur	(Dettes fournisseurs x 360) /CA	V5	ELHAMMMA Azzouz, 2011. BERRAIH Radia, 2020.
Liquidité	Ratio de liquidité réduite	ACT-Stocks/DCT	V6	CHOW Jacky C.K, 2017. BERRAIH Radia, 2020.
	Disponibilité	Disponibilité (net) / Actif	V7	CHOW Jacky C. K, 2017. BERRAIH Radia, 2020.
	Ratio de liquidité immédiate	Disponibilité (net) / DCT	V8	ELHAMMMA Azzouz, 2011. BERRAIH Radia, 2020.
	Ratio de couverture	Charges financières / EBE	V9	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017.BERRAIH Radia, 2020.
	Dette à court terme	DCT/Total dette	V10	BERRAIH Radia, 2020.
	Rotation de stock	(Stock moyen de marchandises HT / coût d'achat des marchandises vendues HT) * 360	V11	CHOW Jacky C. K, 2017. BERRAIH Radia, 2020.
Rentabilité	ROA	Résultat net après impôt / Total actif	V12	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	ROE	Résultat net / Fonds propres	V13	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Marge brute d'autofinancement	CAF / CA	V14	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Taux d'intégration	Valeur Ajoutée d'exploitation / CA	V15	BERRAIH Radia, 2020.
	Rentabilité opérationnelle	EBE / CA	V16	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.

Solvabilité	Marge opérationnelle	Résultat opérationnel / CA	V17	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Taux de marge financière	Résultat financier / CA	V18	CHOW Jacky C.K, 2017. BERRAIH Radia, 2020.
	Rentabilité commerciale avant impôt	Résultat net avant impôts / CA	V19	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Rentabilité commerciale après impôt	Résultat net après impôts / CA	V20	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
Structure	Solvabilité générale	Total dettes / Total Actif	V21	LOTFI Siham et MESK Hicham ,2020. BERRAIH Radia, 2020.
	Levier financier	Dettes financières / Fonds propres	V22	BERRAIH Radia, 2020.
	Autonomie financière	Fonds propres / Passif Non courant	V23	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Ratio d'adéquation des fonds propres	Fonds propres / Total actif	V24	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.
	Capacité de remboursement	Dettes financières / Résultat net	V25	BERRAIH Radia, 2020.
	Capacité de remboursement	DLT / CAF	V26	KHEMAKHEM Sihem et BOUJELBENE Younes, 2017. BERRAIH Radia, 2020.

Source : Elaboré par les auteurs.

Annexe N° 2 : Présentation du modèle ANN



Annexe N° 3 : Contribution des variables explicatives dans les composantes principales

```
> head(var$contrib, 14)
```

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
FEnJrsCA	26.25962876	6.597722077	0.02682376	0.43677654	6.869107e-01
DélaiClient	7.29000905	6.571189517	11.61960525	0.04670262	4.111389e+00
DélaiFournisseur	26.34110256	4.028311870	0.13448921	1.40557575	3.615772e+00
Disponibilité(net)/DCT	1.44629915	15.006222464	4.50563206	20.06430798	4.574188e-01
CF/EBE	1.05046760	10.653230590	1.37497698	7.14502603	1.270422e+01
DCT/Total D	0.28644184	4.470466719	11.74339979	9.89657907	1.871850e+01
RotationStocks	0.17195819	2.064825487	2.28429192	1.57727913	2.029191e+00
ROA	5.10254582	22.918648225	1.79330910	0.76076718	1.082655e+00
ROE	0.14932282	3.527222213	0.92455692	34.71921283	5.748034e+00
CAF/CA	0.17117605	9.499053363	9.45654686	4.16708596	2.106882e+01
VAExpIt/CA	0.65579621	2.881109395	35.40733411	2.34682911	5.344418e+00
EBE/CA	24.38209941	0.001138943	11.00742554	0.01680576	4.222751e-05
TotalD/TotalActif	6.67702012	10.628334165	7.73632453	16.99970946	3.367656e-01
DLT/CAF	0.01613241	1.152524971	1.98528398	0.41734257	2.409587e+01

```
> |
```

Annexe N° 4 : Table récapitulative du modèle RL

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.74326	0.65562	-2.659	0.00784	**
FormeJuridique2	-0.75794	0.48758	-1.554	0.12007	
FormeJuridique3	-0.79952	1.17137	-0.683	0.49489	
FormeJuridique4	1.07130	1.27045	0.843	0.39909	
centraleRisques1	0.13874	0.49187	0.282	0.77790	
ImpayésConfrères1	0.02410	0.49628	0.049	0.96127	
MouvementsCconfiés0	2.77929	0.43034	6.458	1.06e-10	***
Activité2	-0.43181	0.62661	-0.689	0.49075	
Activité3	-0.11211	0.82442	-0.136	0.89183	
Activité4	-0.99666	1.54476	-0.645	0.51881	
Activité5	-0.96212	0.94484	-1.018	0.30854	
Activité6	-17.75651	1246.77668	-0.014	0.98864	
Activité7	1.86782	1.01164	1.846	0.06484	.
Activité8	0.57032	1.99325	0.286	0.77478	
Activité9	-16.55458	6522.63864	-0.003	0.99797	
PC1	-0.07389	0.17845	-0.414	0.67885	
PC2	0.02487	0.18549	0.134	0.89332	
PC3	0.36998	0.22761	1.626	0.10405	
PC4	-0.55106	0.27143	-2.030	0.04234	*
PC5	0.23463	0.19994	1.173	0.24061	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 254.20 on 227 degrees of freedom
Residual deviance: 161.59 on 208 degrees of freedom
(1 observation deleted due to missingness)
AIC: 201.59

Annexe N° 5 : Présentation des configurations des modèles

Codage	Méthode	Nombre de variable explicatives	Vecteur			
			ACP	NA	ASC	Précision
M1	KNN	31	-	Supprimer	77%	90%
M2	KNN	19	-	Supprimer	80,3%	74,24%
M3	KNN	5	5	Supprimer	71%	76%
M4	ANN	31	-	Supprimer	61%	66%
M5	ANN	19	-	Supprimer	87,1%	81,39%
M6	ANN	5	5	Supprimer	55%	73%
				Ne converge	Ne converge	Ne converge
M7	RL	31	-	pas	pas	pas
M8	RL	31	-	Garder	61%	78%
M9	RL	19	-	Garder	62%	68%
M10	RL	19	-	Supprimer	55,5%	76%
M11	RL	5	5	Garder	61%	84%
M12	RL	5	5	Garder	66%	85%
M13	RL	5	5	Supprimer	62,7%	81,03%

Source : Elaboré par les auteurs à l'aide du logiciel R.