Master in Artificial Intelligence & Management

Cycle Grande Ecole

SynthoScore

Unified Credit Evaluation Algorithm

IAS-M2JV-DA1

Année Scolaire

2022 - 2023

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon projet et qui m'ont aidé lors de la rédaction de ce mémoire.

Je tiens également à adresser mes remerciements à mes camarades de classe et à mes amis qui ont partagé ce parcours avec moi. Les discussions enrichissantes et les échanges fructueux ont largement contribué à façonner mes idées et à m'inspirer dans cette recherche.

Mes remerciements vont également à ma famille en Tunisie, Lucie Schneller Lorenzoni, ma tante à Toulouse et Clone dont le soutien inébranlable et les encouragements constants ont été une source de motivation essentielle.

Nous atterrissons bientôt, merci à vous tous pour avoir fait partie de ce voyage académique et pour avoir contribué à son succès.

Sommaire

Résumé

Nous sommes actuellement confrontés à des temps économiquement difficiles, marqués par une inflation persistante qui affecte notre pouvoir d'achat. Cette situation a rendu de plus en plus complexe la réalisation de nos projets ou même la simple survie financière sans avoir recours à des emprunts bancaires.

Cependant, cette nécessité de recourir aux prêts bancaires présente également des risques pour les institutions financières. Il n'y a aucune garantie absolue que les emprunteurs remboursent intégralement les sommes empruntées en espèces, ce qui pourrait engendrer des problèmes financiers importants pour les banques.

C'est pourquoi les institutions bancaires ont développé diverses techniques et méthodes pour évaluer la solvabilité de leurs clients. Parmi ces méthodes figurent l'Approche Expert, le score Bayésien et la régression logistique, chacun ayant ses avantages et ses inconvénients.

Cependant, il est apparu que ces modèles présentent des limitations et des risques inhérents. Pour remédier à cela, une approche novatrice a été entreprise. Elle repose sur l'utilisation de données actualisées, de scores pondérés calculés méticuleusement, et d'un modèle de régression logistique. Cette fusion de trois éléments distincts vise à créer un modèle "SynthoScore".

Ce modèle a pour objectif d'améliorer la précision et la fiabilité de l'évaluation des risques de crédit en prenant en compte les avantages de ces différentes approches. Il incarne une tentative d'offrir aux institutions financières un outil plus puissant et précis pour prendre des décisions éclairées en matière de prêts et de crédits, tout en minimisant les risques liés à la solvabilité des emprunteurs.

Abstract

We are currently living in challenging economic times characterized by persistent inflation that affects our purchasing power. This situation has made it increasingly difficult to realize our goals or even maintain financial stability without resorting to bank loans.

However, this necessity of turning to bank loans also poses risks to financial institutions. There is no absolute guarantee that borrowers will fully repay the borrowed amounts in cash, which could lead to significant financial issues for banks.

This is why banking institutions have developed various techniques and methods to assess the creditworthiness of their clients. These methods include the Expert Approach, Bayesian scoring, and logistic regression, each with its own advantages and disadvantages.

However, it has become apparent that these models have limitations and inherent risks. To address this, an innovative approach has been undertaken. It relies on the use of updated data, meticulously calculated weighted scores, and a logistic regression model. This fusion of three distinct elements aims to create a synthetic model called "SynthoScore."

This model aims to enhance the accuracy and reliability of credit risk assessment by leveraging the strengths of these different approaches. It represents an endeavor to provide financial institutions with a more robust and precise tool for making well-informed decisions regarding loans and credit, while minimizing the risks associated with borrower solvency.

Introduction

Le secteur bancaire joue un rôle central dans l'économie mondiale en facilitant l'allocation efficace des ressources financières et en favorisant la croissance économique. Cependant, ce rôle ne vient pas sans son lot de risques. L'un des risques les plus significatifs auxquels les institutions financières sont confrontées est le risque de crédit. Ce risque découle de la possibilité que les emprunteurs ne parviennent pas à rembourser leurs dettes conformément aux accords convenus, entraînant des pertes financières pour les prêteurs.

L'évaluation des risques de crédit, en tant que processus fondamental au sein des institutions financières, revêt une importance stratégique et opérationnelle incontestable. Son objectif principal est d'examiner la capacité des emprunteurs potentiels et existants à honorer leurs obligations financières, garantissant ainsi une prise de décision éclairée lors de l'octroi de prêts et de crédits.

Cette évaluation n'est pas seulement une mesure préventive pour éviter les défauts de paiement, mais aussi un moyen de maintenir un équilibre entre les objectifs commerciaux et les risques financiers. Au cœur de ce processus se trouve la nécessité d'évaluer la solvabilité d'un emprunteur. Les institutions financières doivent non seulement identifier les emprunteurs dignes de confiance, mais également quantifier le niveau de risque associé à chaque prêt.

Cela implique d'analyser de manière approfondie les informations financières, les antécédents de crédit, les revenus, l'emploi et d'autres facteurs pertinents pour établir un portrait complet du profil de l'emprunteur. Les décisions d'octroi de crédit sont prises dans un environnement économique en constante évolution, où l'incertitude et les fluctuations peuvent impacter la capacité de remboursement des emprunteurs. Une mauvaise évaluation des risques de crédit peut avoir des répercussions graves.

Si les prêteurs ne parviennent pas à évaluer correctement le risque, ils pourraient être exposés à des pertes financières considérables résultant de défauts de paiement.

De plus, ces pertes pourraient éventuellement affecter la stabilité financière de l'institution et entraîner une perte de confiance de la part des investisseurs, des déposants et même des régulateurs.

Avant toute évaluation du risque de crédit, il y a lieu de procéder à son identification.

L'identification du risque est une opération ou un ensemble d'opérations qui permet d'identifier un risque en le décrivant et en présentant ses principales caractéristiques.

Le but de l'identification des risques est d'identifier les problèmes potentiels avant qu'ils ne deviennent de vrais problèmes et d'intégrer ces informations dans le processus d'évaluation.

Dans un environnement économique en constante évolution, caractérisé par des fluctuations, des incertitudes et une complexité croissante des transactions financières, l'évaluation des risques de crédit revêt une importance capitale pour les institutions financières. Face à la variabilité des marchés, aux changements de comportement des emprunteurs et aux facteurs macroéconomiques changeants, il est impératif de déployer des approches sophistiquées et des outils avant-gardistes pour appréhender ces dynamiques complexes.

C'est précisément à ce point que l'intelligence artificielle (IA) et l'analyse des données se révèlent être des catalyseurs essentiels. Les technologies émergentes dans ces domaines ouvrent de nouvelles perspectives pour élever l'évaluation des risques de crédit à un niveau supérieur de précision et de fiabilité. La volumétrie croissante de données financières disponibles offre un potentiel énorme pour une prise de décision plus informée et plus pointue.

Les institutions financières sont confrontées à des ensembles de données de plus en plus vastes et variés. L'IA permet d'exploiter ces données de manière efficiente et exhaustive. En utilisant des algorithmes sophistiqués d'apprentissage automatique, il est possible de détecter des tendances et des relations subtiles qui échappent souvent aux méthodes traditionnelles. Ces modèles peuvent identifier des schémas complexes dans les comportements de remboursement, les transactions financières et les antécédents de crédit, permettant ainsi de prévoir plus précisément les risques potentiels et c’est ce qu’on va découvrir dans les prochaines parties.

L'analyse des données associée à l'IA permet également de modéliser les risques d'une manière plus précise et personnalisée. Plutôt que de se fier uniquement à des critères généraux, les modèles d'IA peuvent prendre en compte des caractéristiques spécifiques à chaque emprunteur, ce qui entraîne des évaluations plus pertinentes et des stratégies de gestion des risques mieux adaptées.

Revue de littérature et bonnes pratiques

Les cotes de crédit ont pour but de mesurer la solvabilité des entreprises. Elles représentent l'opinion d'une agence de notation qui évalue la force fondamentale de crédit d'un émetteur et sa capacité à satisfaire pleinement et ponctuellement ses obligations de dette (Gonzalez & Descamps-Julien, 2004).

Selon Krahnen (2001), c'est une pratique courante pour les banques d'approuver des prêts et du crédit à l'aide de modèles internes de type « notation de crédit fantôme ».

Ceci découle de l'accord de Bâle II (2004) qui dit que, pour le risque de crédit, les banques peuvent employer différents mécanismes d'évaluation :

- La méthode dite « standard » consiste à utiliser des systèmes de notation fournis par des organismes externes (agences de notation).
- Les méthodes plus sophistiquées (méthodes IRB) avec la méthode dite IRB-fondation et celle dite IRB-avancée impliquent des méthodologies internes et propres à l'établissement financier d'évaluation de côtes ou de notes, afin de peser le risque relatif du crédit.

Les cotes de crédit sont principalement influencées par trois catégories de déterminants :

Dans la première catégorie, on retrouve les ratios financiers et les informations financières. Selon Ederington (1985), ces facteurs approximent les facteurs spécifiques de l'entreprise comme le levier, la liquidité et la taille de celle-ci.

La deuxième catégorie comprend les mécanismes de gouvernance de la firme, qui selon Bhojraj & Sengupta (2003), inclut des facteurs comme la structure de propriété et l'indépendance du conseil d'administration.

La dernière catégorie de déterminants de la cote de crédit est constituée des facteurs macroéconomiques.

Ces derniers peuvent être, par exemple, selon Amato & Furfine (2004), la croissance de PIB. D'autres facteurs comme l'inflation, les taux d'intérêt, etc. peuvent également être considérés dans cette catégorie.

En pratique, il y a plusieurs méthodes statistiques utilisées afin de tester si les facteurs compris dans les catégories ci-haut ont réellement un pouvoir explicatif sur la qualité de crédit.

Matthies (2013) relate que les ratios financiers sont essentiels pour déterminer la qualité de crédit d'une entreprise.

Les deux autres catégories peuvent être vues comme de l'information complémentaire à ces derniers. Dans Matthies (2013), il est dit que les méthodes statistiques d'estimation de la qualité de crédit peuvent être séparées en deux catégories, les méthodes ordonnées et les méthodes non-ordonnées.

La première catégorie comprend la méthode des moindres carrés ordinaires (MCO) et les modèles probit ordonnés.

Cette catégorie de modèle repose sur l'hypothèse principale que les cotes de crédit sont classées. La deuxième catégorie, quant à elle, comprend des méthodes comme les modèles probit non-ordonnés et les analyses discriminantes linéaires.

L'hypothèse que les cotes ne sont pas classées est faite dans ces modèles. Dans une étude empirique, Ederington (1985) trouve que les modèles probit ordonnés surpassent les modèles de type MCO.

Évaluation des risques de crédit dans le secteur bancaire

Le risque de crédit est très important dans le contexte de décisions financières prises par les institutions bancaires et les prêteurs.

- Concept de risque de crédit et son rôle dans les décisions financières

Le risque de crédit est analogue aux pertes potentielles que les banques encourent lorsqu'elles prêtent de l'argent aux agents économiques. Si le débiteur, qu'il s'agisse d'un particulier ou d'une entreprise, ne rembourse pas la dette à temps, la banque perdra une partie des fonds empruntés à tempérament. L'importance du risque de crédit est déterminée par trois paramètres principaux :

- ➢ le montant de la créance,
- ➢ la probabilité de défaut,
- ➢ la proportion de non-recouvrement en cas de défaut du débiteur. 1

En d'autres termes, c'est la possibilité que l'emprunteur fasse défaut sur son obligation de remboursement envers le prêteur. Cette notion fondamentale joue un rôle central dans la prise de décisions financières et a un impact significatif sur la stabilité et la rentabilité des institutions bancaires et des prêteurs.

Lorsqu'un emprunteur n'est pas en mesure de rembourser son prêt, cela peut engendrer des enjeux financiers complexes et des conséquences qui s'étendent bien au-delà de l'emprunteur et du prêteur immédiat.

Conséquences pour les prêteurs :

Si un prêteur accorde un prêt sans évaluer correctement le risque de crédit, il court le risque de ne pas être remboursé. Cela peut entraîner des pertes financières directes pour le prêteur, en particulier dans le cas de prêts importants. Une mauvaise évaluation du risque

de crédit peut également affecter la rentabilité globale de l'institution financière et sa capacité à fonctionner de manière optimale.

Stabilité du Secteur Financier :

Dans le cas où de nombreux emprunteurs feraient défaut en même temps (par exemple, en raison d'une crise économique), cela pourrait avoir un effet domino sur l'ensemble du secteur financier. Cela a été particulièrement évident lors de la crise financière mondiale de 2008, où les défaillances de remboursement de prêts hypothécaires à risque ont entraîné des perturbations majeures dans le système financier mondial.

La crise de 2008 a débuté avec les difficultés rencontrées par les ménages états-uniens à faible revenu pour rembourser les crédits qui leur avaient été consentis pour l'achat de leur logement.

Ces prêts s'adressaient à des emprunteurs qui ne disposaient pas de garanties suffisantes pour bénéficier de taux d'intérêt préférentiels (« *prime rate* ») et ne pouvaient bénéficier que de taux d'intérêt moins favorables (« *subprime* »).

Il est assez aisé de comprendre ce qui en a résulté.

La Banque Centrale américaine a progressivement augmenté les taux d'intérêt de 1 % en 2004 à plus de 5 % en 2006, reflétant les tendances de l'inflation et de la croissance aux États-Unis. Le coût financier de l'emprunt a considérablement augmenté. De plus en plus de ménages ne pouvaient plus y faire face.

Les prix des maisons ont finalement chuté dans tout le pays.

Résultat : les maisons valaient moins que les prêts qui les garantissaient. Une avalanche de défauts de paiement et de reventes de maisons hypothéquées a accéléré la baisse des prix des maisons. Par conséquent, les pertes du côté des prêteurs se sont également accumulées. Les agences de crédit professionnelles ont été les premières à rencontrer des difficultés.

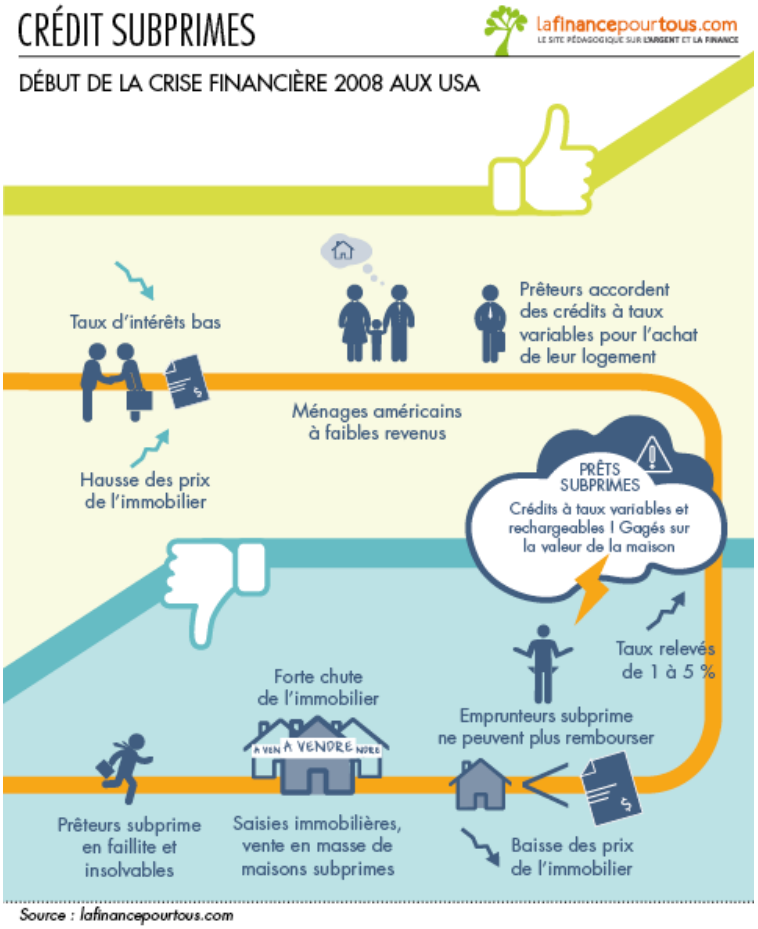


Figure 1 : Début de la crise financières aux Etats-Unis en 2008

Impact sur les emprunteurs :

Les emprunteurs dont le risque de crédit est élevé peuvent se voir refuser des prêts ou se voir attribuer des taux d'intérêt plus élevés pour compenser le risque pris par le prêteur. Cela peut rendre l'accès au crédit plus difficile ou plus coûteux, ce qui peut avoir un impact sur leur capacité à réaliser des projets importants tels que l'achat d'une maison ou la création d'une entreprise.

Confiance des investisseurs :

Les investisseurs qui détiennent des obligations ou des titres de créance émis par des emprunteurs font face à des risques similaires. Si les emprunteurs ne parviennent pas à rembourser leurs obligations, cela peut entraîner des pertes pour les investisseurs et ébranler leur confiance dans le marché financier.

Les facteurs Contribuant au Risque sont des éléments-clés qui influencent la probabilité qu'un emprunteur ne rembourse pas son prêt :

Historique de crédit :

L'historique de crédit d'un emprunteur est l'un des facteurs les plus importants. Cela inclut les antécédents de prêt de l'emprunteur, y compris les antécédents de prêt, de carte de crédit et d'autres remboursements de dettes financières. De bons antécédents de crédit avec des paiements ponctuels et une utilisation responsable du crédit réduisent généralement le risque perçu.

Stabilité financière :

Les prêteurs évaluent la stabilité financière d'un emprunteur, y compris sa capacité à générer un revenu stable et à maintenir des finances personnelles stables. Les emprunteurs disposant d'une source de revenus stable et régulière sont susceptibles de continuer à rembourser leurs prêts et sont donc considérés comme moins risqués.

Revenu :

Le revenu de l'emprunteur est un indicateur important de sa capacité à rembourser son prêt. Le prêteur vérifiera si les revenus sont suffisants pour couvrir les mensualités du prêt et les autres dépenses. Un ratio dette/revenu élevé peut accroître le risque de défaut.

Emploi :

La stabilité et le statut d'emploi de l'emprunteur sont pris en compte. Les emprunteurs permanents à long terme sont considérés comme moins risqués que les emprunteurs temporaires ou temporaires.

Solde de la dette :

Les prêteurs évaluent le fardeau de la dette de l'emprunteur, y compris le solde du prêt, la dette de carte de crédit et d'autres obligations financières. Une dette existante importante peut affecter la capacité de l'emprunteur à rembourser le nouveau prêt.

Ratio prêt-valeur (LTV) :

Ce ratio compare le montant du prêt demandé à la valeur de l'actif en garantie (comme une maison ou une voiture). Un LTV élevé indique un risque accru pour le prêteur car la valeur de la garantie peut ne pas couvrir le montant du prêt en cas de défaut.

Pointage de crédit :

Les agences de notation attribuent un pointage de crédit en fonction des informations financières de l'emprunteur. Une bonne cote de crédit indique une bonne situation financière et réduit le risque perçu.

Type de prêt :

Le type de prêt lui-même peut affecter le risque. Par exemple, les prêts non garantis (comme les cartes de crédit) sont généralement plus risqués que les prêts garantis (comme les hypothèques).

Conditions économiques :

Les conditions économiques actuelles et projetées peuvent également jouer un rôle. Le risque de défaut peut augmenter en période de ralentissement économique.

- Les méthodes traditionnelles de notation de crédit et ses limites

La notation de crédit constitue le fondement du secteur financier et joue un rôle important dans les décisions relatives au crédit et aux prêts. Cela comprend l'évaluation systématique du risque de défaillance associé aux emprunteurs, qu'il s'agisse de particuliers, d'entreprises ou d'organismes gouvernementaux. Cette évaluation permet aux institutions financières de prendre des décisions de prêt éclairées en évaluant la capacité d'un emprunteur à respecter ses obligations financières.

L'histoire de la notation de crédit remonte au XIX^{e} siècle, lorsque les premières agences de notation de crédit ont été créées aux États-Unis. À cette époque, ces agences avaient pour principale mission de fournir des informations aux investisseurs sur la solvabilité des entreprises émettant des obligations. L'objectif était d'aider les investisseurs à prendre des décisions éclairées en évaluant le risque associé aux titres de créance.

Lorsque les premières agences de notation ont émergé vers 1920, leurs métiers étaient peu différents de ceux des analystes financiers.

Depuis les années 1930, en réponse à la crise de 1929, leur marché s'est concentré sur l'analyse de la qualité de crédit des émetteurs. Les notations de crédit reflètent actuellement l'opinion de l'agence basée sur une analyse financière et opérationnelle. Elle repose sur l'analyse de facteurs quantitatifs et qualitatifs liés à la situation actuelle et prévisible de l'entreprise (Ferri et Liu, 2005). L'activité principale des agences de notation de crédit est d'exprimer des opinions sur la qualité de crédit et la capacité des émetteurs ou des instruments financiers à faire face à leurs obligations financières (Cantor et Packer, 1994 ; Paget-Blanc et Painvin, 2007). Les agences de notation facilitent ainsi l'évaluation de la solvabilité d'un émetteur ou d'un emprunteur. Les notations attribuées éliminent la possibilité d'un manque d'informations ou la possibilité que les participants au marché libre effectuent eux-mêmes des analyses complexes.

Ces premières méthodologies de notation reposent principalement sur des évaluations subjectives par des professionnels de la finance. Les agences gouvernementales ont recueilli des informations financières et opérationnelles sur les entreprises et leur ont attribué des notes qualitatives qui reflètent leur perception du risque de défaut. Bien que cette approche soit nouvelle à l'époque, elle était limitée par sa nature subjective et sa dépendance aux connaissances et au jugement individuels.

Au fil du temps, avec l'augmentation des données financières disponibles et l'avènement des statistiques, les méthodes d'évaluation du crédit ont évolué vers des modèles plus quantitatifs. L'introduction de techniques statistiques telles que la régression logistique a permis d'identifier les variables financières les plus pertinentes pour prédire le risque de non-paiement. Il en est résulté une certaine standardisation du processus de notation, améliorant ainsi la comparabilité des notations des différents emprunteurs.

Les trois grandes institutions – Moody’s, Standard & Poor’s et Fitch – ont été fondées à des époques différentes et ont survécu et prospéré à leurs niveaux respectifs, suffisamment pour rester l'une des trois composantes d'une structure oligopolistique viable qui disposait de grands atouts. Standard & Poor’s et Moody’s représentent 80 % du marché en chiffre d'affaires, Fitch Ratings 14 % et les autres agences les 6 % restants.

Moody’s est une entreprise essentielle sur les marchés des capitaux.

En 1913, la base de données des entreprises notées est élargie pour mieux inclure les entreprises industrielles et les services publics.

Moody’s Investors Service est fondée l'année suivante et, en 1914, note les obligations des grandes villes états-uniennes.

La crise de 1929 ne ralentit pas le développement de l'entreprise, elle continue à coter des obligations d'entreprises, mais les taux d'intérêt sur celles-ci augmentent fortement, entraînant une réduction du capital.

Puis, dans les années 1970, Moody’s a élargi ses activités de notation pour inclure divers types de billets de trésorerie et de dépôts bancaires impliquant des émetteurs et des

investisseurs.

La société a décidé de faire payer les entreprises notées pour leurs prestations, la rémunération ne pouvant plus se limiter à une simple inscription dans un annuaire de notation, ce qui peut parfois conduire à des situations ambiguës.

Standard & Poor's est créée en 1941, résultant de la fusion entre Standard Statistics et Poor's Publishing Company. Cependant, les racines de l'entreprise remontent encore plus loin. Dès 1860, Henry Varnum Poor a jeté les bases de ce qui deviendrait plus tard Standard & Poor's. Il a créé une société qui compilait des informations financières sur les entreprises de navigation fluviale et de chemins de fer états-uniens. Cette initiative visait à fournir aux investisseurs, notamment aux épargnants européens, des informations crédibles sur les opportunités d'investissement, étant donné qu'ils ne pouvaient pas facilement évaluer les entreprises à distance.

Henry Varnum Poor a lancé cette initiative avec pour devise « *the investor's right to know* », soulignant ainsi l'importance de fournir aux investisseurs des informations précises pour prendre des décisions éclairées. Il a créé la société H.V. & H.W. Poor Co. et a publié le *Manual of the Railroads of the United States* en 1868, qui a rapidement gagné en popularité. Cet ouvrage fournit des informations essentielles sur les investissements dans le secteur ferroviaire, offrant ainsi un moyen fiable d'évaluation des opportunités.

Au fil des années, d'autres entreprises ont rejoint le paysage, comme le Standard Statistics Bureau en 1906, qui a publié des fiches d'information sur les entreprises ferroviaires. En 1923, la société a gagné en notoriété en étudiant en détail 253 sociétés états-uniennes chaque semaine, démontrant ainsi son sérieux et son professionnalisme.

Un moment marquant dans l'histoire de Standard & Poor's s'est produit en 1929, lorsque l'entreprise a prévu avec précision la crise boursière imminente et a conseillé à ses clients de liquider leurs actifs financiers. Cette prévision précoce a renforcé la crédibilité de l'entreprise et a marqué une étape importante dans son rôle de conseiller financier.

Au fil du temps, Standard & Poor's a continué à évoluer, intégrant des technologies telles que l'ordinateur IBM à cartes perforées en 1946 pour améliorer l'efficacité et la sécurité de ses travaux. En 1957, l'entreprise a créé l'indice S & P 500, un indice majeur du marché

boursier états-unien.

Standard & Poor's a continué à se développer mondialement, ouvrant des agences à Londres, Tokyo et ailleurs. Elle a élargi sa gamme d'indices avec le MidCap 400, le SmallCap 600 et d'autres. En 2010, l'entreprise était présente dans plus de vingt pays et notait un large éventail d'entreprises, d'établissements financiers et de collectivités locales.

Avec son engagement envers la transparence financière et son influence mondiale, Standard & Poor's est devenue une référence incontournable dans l'évaluation des risques et des opportunités économiques.

L'agence Fitch a débuté en tant qu'éditeur d'informations financières, la Fitch Publishing Company, fondée en 1913 à New York par John Knowles Fitch. Cependant, elle s'est rapidement engagée dans l'activité de notation, introduisant en 1924 l'échelle bien connue de notes allant de « AAA » à « D ». Fitch a été parmi les sept premières agences de notation à être reconnues comme des NRSRO par la SEC en 1975. Son ascension s'est poursuivie avec des performances remarquables après une augmentation de capital en 1989 et des efforts pour fournir des recherches originales et des explications pédagogiques pour des produits financiers de plus en plus complexes.

La fusion en 1997 avec IBCA Limited, une agence de notation basée à Londres, a élargi la portée mondiale de Fitch. L'agence a étendu son analyse aux grandes banques, institutions financières majeures et fonds souverains, des domaines souvent délicats à évaluer. Le transfert à la société holding française Fimalac S.A. a suivi cette fusion. Fitch a ensuite acquis Duff and Phelps en 2000, et Thomson BankWatch, renforçant sa position dans la notation bancaire.

Fitch a continué à croître et à diversifier ses activités, acquérant Algorithmics en 2005 pour enrichir ses méthodes d'analyse des risques. La création de Fitch Solutions en 2008, axée sur la formation avancée en gestion des titres à revenu fixe et gestion des risques de crédit, a également renforcé son portefeuille.

Avec une stratégie ambitieuse, Fitch a constamment élargi son influence et sa présence mondiale. En 2010, l'agence notait un grand nombre d'entités, dont 3 500 banques, 1 400 compagnies d'assurances, 2 000 émissions d'entreprises, 300 émissions d'États et d'autorités

territoriales, 46 000 émissions municipales et 6 500 émissions de financement structurel aux États-Unis.

Agences de notation	Chiffre d'affaires 2010	Résultat 2010	Ratio R/CA	Effectif 2010
Standard & Poor's	1 695 M$	762,4 M$	45 %	8 500
Moody's	2 032 M$	507,8 M$	25 %	4 500
Fitch Rating	609 M$	125,6 M$	21 %	2 997

Tableau 1 : Principaux chiffres-clés des agences de notation

Au fil du temps, avec l'augmentation du volume de données financières disponibles et l'émergence des statistiques, les méthodes de notation de crédit ont évolué vers des modèles plus quantitatifs. L'introduction de techniques statistiques telles que la régression logistique a permis d'identifier les variables financières les plus pertinentes pour prédire le risque de non-remboursement. Cela a conduit à une certaine standardisation dans le processus de notation, améliorant ainsi la comparabilité des évaluations entre différentes entités emprunteuses.

Cependant, malgré ces avancées, les méthodes traditionnelles de notation de crédit ont montré leurs limites, notamment en termes de précision dans la prévision des défauts et de leur incapacité à capturer des modèles de risque plus complexes et dynamiques. Ces défis ont ouvert la voie à l'intégration de technologies plus avancées, telles que l'intelligence artificielle et l'apprentissage automatique, pour relever les défis actuels de l'évaluation du risque de crédit.

Les approches traditionnelles telles que l'Approche Expert, le Score Bayésien, la Régression Logistique et le Score Altman Z ont longtemps été les piliers de l'évaluation des risques de crédit dans le secteur bancaire. Cependant, ces méthodes sont sujettes à des limitations importantes qui ont motivé la recherche et l'adoption de nouvelles approches plus avancées et sophistiquées.

L’approche Expert :

C’est une méthode traditionnelle d'évaluation des risques de crédit qui repose sur l'expertise et le jugement des analystes financiers. Les analystes utilisent leurs connaissances et leur expérience pour évaluer la solvabilité d'un emprunteur en examinant une combinaison de facteurs financiers et non financiers. Ces facteurs peuvent inclure le profil financier de l'emprunteur, ses antécédents de crédit, sa stabilité financière, son industrie et d'autres informations pertinentes.

Bien que l'Approche Expert puisse fournir des évaluations personnalisées et tenir compte de facteurs spécifiques à chaque emprunteur, elle présente des inconvénients majeurs. La subjectivité est l'un des principaux défis de cette approche. Les évaluations dépendent du jugement individuel des analystes, ce qui peut entraîner des variations et des biais. De plus, les facteurs non financiers, tels que des préjugés personnels, peuvent influencer les décisions, affectant ainsi l'objectivité des évaluations. L'absence de cadre formalisé peut également rendre difficile la comparaison des évaluations entre différents analystes ou au fil du temps.

Le Score Bayésien :

C’est une méthode qui s'appuie sur les principes de la théorie des probabilités et les statistiques bayésiennes pour évaluer le risque de crédit. Cette approche suppose une distribution probabiliste pour les variables pertinentes et utilise des données historiques pour estimer les probabilités conditionnelles. En d'autres termes, le Score Bayésien calcule la probabilité de défaut d'un emprunteur en tenant compte de la relation entre les variables de crédit.

Cependant, le Score Bayésien a des limites. Il peut avoir du mal à capturer les interactions complexes entre les variables, en particulier lorsque les données historiques sont limitées. Cette méthode peut également être influencée par le choix des distributions probabilistes et des paramètres, ce qui peut introduire un certain degré de subjectivité. De plus, le Score Bayésien peut manquer de flexibilité pour modéliser des schémas non linéaires ou des relations complexes, ce qui peut réduire sa capacité à fournir des évaluations précises dans des scénarios complexes.

La Régression Logistique :

C’est une méthode statistique largement utilisée pour prédire une variable binaire, telle que la probabilité de défaut d'un emprunteur. Elle modélise la relation entre les variables indépendantes (telles que les caractéristiques financières de l'emprunteur) et la variable dépendante (le défaut ou non-défaut). La Régression Logistique est simple à comprendre et à mettre en œuvre, ce qui en fait une méthode attractive pour les institutions financières.

Lorsque nous voulons modéliser une variable à réponse binaire, la forme de la relation est souvent non linéaire. On recourt alors à une fonction non-linéaire, de type logistique par exemple, en pareils cas. Le principe de la régression logistique binaire est de considérer une variable à prévoir binaire (variable cible admettant uniquement deux modalités possibles) Y = {0,1} d'une part, et p variables explicatives notées X = (X1, X2, ..., Xj), continues, binaires ou qualitatives. L'objectif de la régression logistique est de modéliser l'espérance conditionnelle E(Y/X=x), par l'estimation d'une valeur moyenne de Y pour toute valeur de X. Pour une valeur Y valant 0 ou 1 (loi de Bernouilli), cette valeur moyenne est la probabilité que Y=1. On obtient donc :

$$E(Y/X=x) = Prob(Y=1/X=x)$$

En fait, en cherchant à expliquer la probabilité de réalisation de l'évènement Prob(Y=1/X=x), il nous faudrait une transformation de E(Y) qui étende l'intervalle de définition [0,1]. C'est le calcul des ratios de chance « *odds ratio* » qui permet d'envisager cette transformation. Ainsi, le quotient pi /(1-pi) est appelé « *odds* », et la fonction f(p)=ln (pi/1-pi) est appelée « *logit* ».

Le fonctionnement consiste à calculer des coefficients de régression de façon itérative. En d'autres termes, le programme informatique, à partir de certaines valeurs de départ pour Y0 et Y1, vérifiera si les log chances (*odd ratios*) estimés sont bien ajustés aux données, corrigera les coefficients, réexaminera le bon ajustement des valeurs estimées, jusqu'à ce qu'aucune correction ne puisse atteindre un meilleur résultat (Howell, 1998).

Cependant, la simplicité de la Régression Logistique peut également être sa limite. Cette méthode suppose une relation linéaire entre les variables indépendantes et la probabilité de défaut, ce qui peut ne pas être réaliste dans le contexte des données de crédit. Les schémas non linéaires et les interactions complexes entre les variables peuvent ne pas être correctement capturés, ce qui peut entraîner des évaluations biaisées ou inexactes. Des méthodes plus avancées sont nécessaires pour modéliser la complexité inhérente aux données de crédit.

Le Score Altman Z :

Il est développé par le professeur Edward Altman, est une méthode spécifiquement conçue pour évaluer la solvabilité des entreprises en utilisant une combinaison de ratios financiers. Il attribue des points à différents ratios financiers tels que la rentabilité, l'endettement, la liquidité, etc., et agrège ces points pour calculer un score global. Ce score est ensuite utilisé pour classer les entreprises en différentes zones de risque.

Bien que le Score Altman Z ait été largement utilisé pour évaluer la santé financière des entreprises, il peut ne pas être adapté aux jeunes entreprises en croissance. Les startups et les entreprises en croissance peuvent avoir des caractéristiques financières différentes de celles des entreprises établies, ce qui peut rendre le modèle moins précis pour ce type d'entreprises. De plus, le modèle ne prend pas toujours en compte des facteurs macroéconomiques ou des événements non financiers qui peuvent influencer la solvabilité d'une entreprise.

Intelligence artificielle et analyse prédictive

- Introduction à l'intelligence artificielle et ses applications dans la finance

L'introduction à l'intelligence artificielle (IA) dans le contexte de la finance ouvre un nouvel horizon de possibilités en matière d'évaluation des risques de crédit. L'IA, en tant que domaine de la science informatique, vise à doter les machines de la capacité de réaliser des tâches qui, normalement, nécessitent l'intelligence humaine. Dans le secteur financier, l'IA a émergé comme une force transformative grâce à sa capacité à analyser de vastes quantités de données complexes et à extraire des informations significatives pour prendre des décisions éclairées. Ces caractéristiques en font un choix naturel pour relever les défis inhérents à l'évaluation des risques de crédit, où la précision et la rapidité sont essentielles.

L'IA propose un éventail d'applications dans la finance, allant de l'automatisation des processus commerciaux à la prise de décisions d'investissement sophistiquées. En ce qui concerne l'évaluation des risques de crédit, l'IA offre plusieurs avantages significatifs par rapport aux méthodes traditionnelles.

Tout d'abord, elle permet de traiter des volumes massifs de données hétérogènes en un temps beaucoup plus court que les méthodes manuelles ou semi-automatisées. Cela comprend non seulement les données financières telles que les relevés de compte, les rapports financiers et les antécédents de crédit, mais aussi des données non structurées telles que les médias sociaux, les actualités et d'autres sources externes qui peuvent fournir des informations supplémentaires sur la santé financière d'un emprunteur.

Deuxièmement, l'IA est capable de détecter des tendances et des modèles complexes qui pourraient échapper aux méthodes traditionnelles. Elle peut identifier des relations non linéaires entre les variables, détecter des signaux faibles indiquant un risque potentiel et fournir des prévisions plus précises en fonction de l'évolution des données au fil du temps. Par exemple, dans le cas de l'évaluation des risques de crédit, les modèles d'IA peuvent

repérer des indicateurs subtils de détérioration de la santé financière d'un emprunteur, bien avant qu'ils ne deviennent évidents pour les analystes humains.

Troisièmement, l'IA peut aider à réduire les biais humains et à augmenter l'objectivité dans le processus d'évaluation. Contrairement aux méthodes traditionnelles qui peuvent être influencées par des préjugés inconscients ou des jugements individuels, les modèles d'IA se fondent sur des données et des algorithmes, ce qui réduit la subjectivité et assure une certaine cohérence dans le processus d'évaluation.

Enfin, l'IA permet également une personnalisation accrue dans l'évaluation des risques de crédit. En utilisant des modèles d'apprentissage automatique, les institutions financières peuvent créer des modèles de crédit adaptés à des segments spécifiques de clients, en prenant en compte des caractéristiques individuelles telles que le secteur d'activité, la taille de l'entreprise, la géographie, etc.

- Analyse prédictive et apprentissage automatique pour l'évaluation des risques de crédit

L'analyse prédictive et l'apprentissage automatique jouent un rôle crucial dans l'évolution de l'évaluation des risques de crédit dans le secteur financier. Ces techniques modernes permettent d'aller au-delà des méthodes traditionnelles en exploitant la puissance des données et de l'informatique pour créer des modèles de prévision plus précis et sophistiqués. Lorsqu'appliquées à l'évaluation des risques de crédit, l'analyse prédictive et l'apprentissage automatique apportent une approche plus dynamique et personnalisée à la prise de décision.

L'analyse prédictive consiste à utiliser des données historiques pour identifier des schémas, des tendances et des relations entre les variables. Dans le contexte de l'évaluation des risques de crédit, cela implique l'analyse de données passées sur les emprunteurs et les emprunts, ainsi que les résultats associés tels que les remboursements ou les défauts. L'objectif est d'identifier des caractéristiques spécifiques et des combinaisons de variables qui sont fortement corrélées avec la probabilité de défaut de paiement. Ces modèles peuvent ensuite être utilisés pour prédire le risque de crédit associé à de nouveaux emprunteurs.

L'apprentissage automatique, une sous-catégorie de l'IA, s'appuie sur l'analyse prédictive en utilisant des algorithmes pour entraîner des modèles sur des données historiques et leur permettre de s'améliorer avec le temps. Les algorithmes d'apprentissage automatique peuvent être classés en deux grandes catégories : supervisés et non supervisés. Dans le contexte de l'évaluation des risques de crédit, les modèles supervisés sont les plus couramment utilisés.

Les modèles supervisés nécessitent des données d'entraînement étiquetées, c'est-à-dire des exemples d'emprunteurs avec des étiquettes indiquant s'ils ont remboursé ou fait défaut sur leur prêt. Parmi les techniques d'apprentissage automatique populaires pour l'évaluation des risques de crédit, on peut citer :

I. Les arbres de décision : ils sont utilisés pour diviser les données en plusieurs groupes homogènes en fonction des caractéristiques des emprunteurs. Ces arbres peuvent être interprétés visuellement, ce qui aide à expliquer les décisions prises par le modèle.

II. Les forêts aléatoires : ce sont des ensembles d'arbres de décision qui travaillent ensemble pour améliorer la précision et réduire le surajustement. Ils sont particulièrement utiles pour gérer des ensembles de données complexes avec de nombreuses variables.

III. Les réseaux de neurones artificiels : inspirés par le fonctionnement du cerveau, ces modèles sont capables de capturer des relations non linéaires complexes entre les variables. Ils sont efficaces pour l'analyse de grandes quantités de données, mais peuvent nécessiter une plus grande puissance de calcul pour l'entraînement.

IV. Les machines à vecteurs de support (SVM) : ces modèles trouvent une frontière de décision qui maximise la marge entre les exemples de différentes classes, ce qui les rend efficaces pour les problèmes de classification binaire.

Ces techniques permettent de créer des modèles plus précis et adaptés à des cas spécifiques d'évaluation des risques de crédit. Cependant, il est important de noter que l'apprentissage automatique n'est pas exempt de défis. La sélection de variables pertinentes, le traitement des données manquantes, la gestion des biais et la garantie de l'interprétabilité des modèles sont autant de considérations essentielles pour garantir l'efficacité et l'exactitude des prédictions.

Collecte et prétraitement des données

- Types de données pertinentes pour l'évaluation des risques de crédit :

Le processus d'évaluation des risques de crédit repose en grande partie sur la qualité et la pertinence des données utilisées pour l'analyse. Dans cette section, nous discuterons des types de données pertinentes pour l'évaluation du risque de crédit et nous fournirons des informations sur le jeu de données utilisé pour cette recherche.

Pour mener à bien cette recherche sur l'évaluation des risques de crédit, il était essentiel de disposer d'un jeu de données fiable et représentatif. Dans cette optique, plusieurs sources de données ont été examinées, notamment des bases de données gouvernementales, des données financières publiques, et d'autres ressources.

Après une recherche approfondie, le jeu de données "Bank_loan_data" de Kaggle a été identifié comme une source de confiance. Ce jeu de données a suscité un intérêt considérable au sein de la communauté des sciences des données, comme en témoignent ses statistiques impressionnantes. Avec un taux de téléchargement par rapport aux vues de 0,13 % et plus de 30 000 vues, il est clair que cette ressource est hautement appréciée par les professionnels et les chercheurs.

L'utilisation de données provenant de cette source fiable garantit la qualité et la validité de notre recherche sur l'évaluation des risques de crédit. Les informations extraites de ce jeu de données ont servi de base solide pour la construction de notre modèle d'évaluation des risques de crédit, contribuant ainsi à la crédibilité et à la pertinence de cette étude.

L'utilisation de données provenant de cette source fiable garantit la qualité et la validité de notre recherche sur l'évaluation des risques de crédit. Les informations extraites de ce jeu de données ont servi de base solide pour la construction de notre modèle d'évaluation des risques de crédit, contribuant ainsi à la crédibilité et à la pertinence de cette étude.

Kaggle fournit un environnement collaboratif où les chercheurs peuvent partager leurs compétences, échanger des idées et collaborer sur des projets de données. Cela a été particulièrement bénéfique pour notre recherche, car il nous a permis de bénéficier des

contributions d'une communauté de scientifiques des données expérimentés.

1. Informations sur le Client

Customer_ID (Identifiant Client) : Un identifiant unique pour chaque client, qui aide à suivre et à gérer les comptes des clients.

2. Historique Financier

- ➔ *Status_Checking_Acc* (État du Compte Courant) : Ce champ peut indiquer l'historique de crédit du client, en particulier l'état de son compte courant. Il pourrait inclure des catégories telles que … < 0 USD, 0 <= … < 10000, no checking account, etc.
- ➔ *Savings_Acc* (Compte Épargne) : Similaire à l'état du compte courant, ce champ peut fournir des informations sur l'état du compte épargne du client.
- ➔ *Credit_Amount* (Montant du Crédit) : Le montant du crédit demandé ou accordé au client.
- ➔ *Inst_Rt_Income* (Taux d'Échéance par Rapport au Revenu) : Le taux d'échéance en pourcentage du revenu disponible.
- ➔ Years_At_Present_Employment (Années à l'Emploi Actuel) : Ces données peuvent offrir des détails sur le nombre d'années pendant lesquelles le client est employé à son emploi actuel.
- ➔ *Marital_Status_Gender* (État Civil et Genre) : Des informations sur l'état civil et le genre du client, qui pourraient être indicatives de sa stabilité financière.
- ➔ *Other_Debtors_Guarantors* (Autres Débiteurs et Garants) : Des informations sur d'autres débiteurs ou garants impliqués dans la transaction de crédit.
- ➔ *Current_Address_Yrs* (Années à l'Adresse Actuelle) : Le nombre d'années pendant lesquelles le client réside à son adresse actuelle, ce qui pourrait être un facteur dans l'évaluation de la stabilité.
- ➔ *Property (Propriété)* : Des détails sur la propriété du client, qui pourrait servir de garantie pour le prêt.
- ➔ *Age* (Âge) : L'âge du client, qui pourrait être un facteur important dans l'évaluation du risque de crédit.
- ➔ *Other_Inst_Plans* (Autres Plans d'Échéance) : Des informations sur d'autres plans d'échéance que le client pourrait avoir, ce qui pourrait affecter sa capacité à rembourser.

3. Informations sur le Prêt

- ➔ *Duration_in_Months* (Durée en Mois) : La durée du prêt en mois.
- ➔ *Purposre_Credit_Taken* (Raison du Crédit) : La raison pour laquelle le crédit est demandé, telle que 'radio/tv', 'éducation', 'meubles/équipement', etc.
- ➔ *Housing* (Logement) : Des détails sur la situation de logement du client, qui pourrait être liée à sa stabilité financière.
- ➔ *Num_CC* (Nombre de Cartes de Crédit) : Le nombre de cartes de crédit détenues par le client.
- ➔ *Job* (Emploi) : Des informations sur l'emploi ou le statut d'emploi du client.
- ➔ *Dependents* (Personnes à Charge) : Le nombre de personnes à charge du client, ce qui pourrait affecter ses obligations financières.
- ➔ *Telephone* (Téléphone) : Indique si le client dispose d'une connexion téléphonique (oui/non).
- ➔ *Foreign_Worker* (Travailleur Étranger) : Indique si le client est un travailleur étranger (oui/non).

4. Historique de Crédit

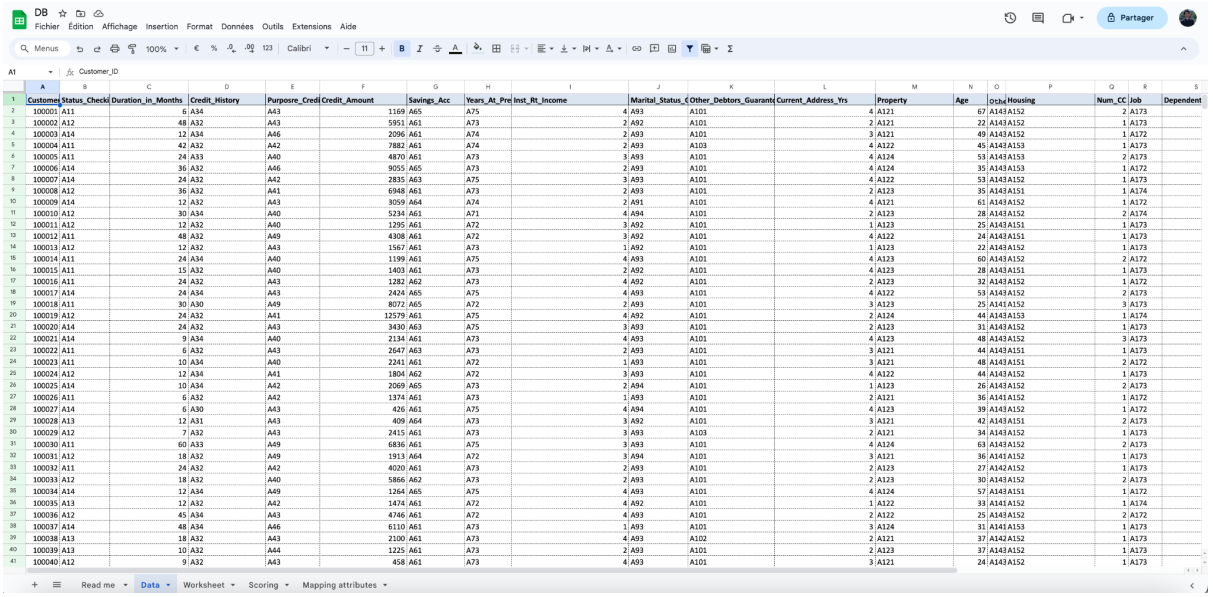
- ➔ *Credit_History* (Historique de Crédit) : Fournit des informations historiques sur le crédit du client, notamment s'il a remboursé les prêts précédents ou connu des défauts de paiement.

5. Résultat du Prêt

- ➔ *Default_On_Payment* (Défaut de Paiement) : La variable cible indiquant si le client a fait défaut de paiement (oui/non).
- ➔

Ce jeu de données, "Bank_loan_data," est une ressource précieuse pour l'étude du risque de crédit. Il contient une gamme variée d'attributs pouvant être utilisés pour construire des modèles prédictifs d'évaluation du risque de crédit. Avec un taux de téléchargement par rapport aux vues de 0,13 % et plus de 30 000 vues sur Kaggle, il est évident que ce jeu de

données a suscité un intérêt important et une confiance de la part de la communauté des sciences des données.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Customer	Status_Check	Duration_in_Months	Credit_History	Purposre_Credi	Credit_Amount	Savings_Acc	Years_At_Pre	Inst_Rt_Income	Marital_Status_(	Other_Debtors_Guarant	Current_Address_Yrs	Property	Age	Other	Housing	Num_CC	Job	Dependent
2	100001	A11	6	A34	A43	1169	A65	A75	4	A93	A101	4	A121	67	A143	A152	2	A173	
3	100002	A12	48	A32	A43	5951	A61	A73	2	A92	A101	2	A121	22	A143	A152	1	A173	
4	100003	A14	12	A34	A46	2096	A61	A74	2	A93	A101	3	A121	49	A143	A152	1	A172	
5	100004	A11	42	A32	A42	7882	A61	A74	2	A93	A103	4	A122	45	A143	A153	1	A173	
6	100005	A11	24	A33	A40	4870	A61	A73	3	A93	A101	4	A124	53	A143	A153	2	A173	
7	100006	A14	36	A32	A46	9055	A65	A73	2	A93	A101	4	A124	35	A143	A153	1	A172	
8	100007	A14	24	A32	A42	2835	A63	A75	3	A93	A101	4	A122	53	A143	A152	1	A173	
9	100008	A12	36	A32	A41	6948	A61	A73	2	A93	A101	2	A123	35	A143	A151	1	A174	
10	100009	A14	12	A32	A43	3059	A64	A74	2	A91	A101	4	A121	61	A143	A152	1	A172	
11	100010	A12	30	A34	A40	5234	A61	A71	4	A94	A101	2	A123	28	A143	A152	2	A174	
12	100011	A12	12	A32	A40	1295	A61	A72	3	A92	A101	1	A123	25	A143	A151	1	A173	
13	100012	A11	48	A32	A49	4308	A61	A72	3	A92	A101	4	A122	24	A143	A151	1	A173	
14	100013	A12	12	A32	A43	1567	A61	A73	1	A92	A101	1	A123	22	A143	A152	1	A173	
15	100014	A11	24	A34	A40	1199	A61	A75	4	A93	A101	4	A123	60	A143	A152	2	A172	
16	100015	A11	15	A32	A40	1403	A61	A73	2	A92	A101	4	A123	28	A143	A151	1	A173	
17	100016	A11	24	A32	A43	1282	A62	A73	4	A92	A101	2	A123	32	A143	A152	1	A172	
18	100017	A14	24	A34	A43	2424	A65	A75	4	A93	A101	4	A122	53	A143	A152	2	A173	
19	100018	A11	30	A30	A49	8072	A65	A72	2	A93	A101	3	A123	25	A141	A152	3	A173	
20	100019	A12	24	A32	A41	12579	A61	A75	4	A92	A101	2	A124	44	A143	A153	1	A174	
21	100020	A14	24	A32	A43	3430	A63	A75	3	A93	A101	2	A123	31	A143	A152	1	A173	
22	100021	A14	9	A34	A40	2134	A61	A73	4	A93	A101	4	A123	48	A143	A152	3	A173	
23	100022	A11	6	A32	A43	2647	A63	A73	2	A93	A101	3	A121	44	A143	A151	1	A173	
24	100023	A11	10	A34	A40	2241	A61	A72	1	A93	A101	3	A121	48	A143	A151	2	A172	
25	100024	A12	12	A34	A41	1804	A62	A72	3	A93	A101	4	A122	44	A143	A152	1	A173	
26	100025	A14	10	A34	A42	2069	A65	A73	2	A94	A101	1	A123	26	A143	A152	2	A173	
27	100026	A11	6	A32	A42	1374	A61	A73	1	A93	A101	2	A121	36	A141	A152	1	A172	
28	100027	A14	6	A30	A43	426	A61	A75	4	A94	A101	4	A123	39	A143	A152	1	A172	
29	100028	A13	12	A31	A43	409	A64	A73	3	A92	A101	3	A121	42	A143	A151	2	A173	
30	100029	A12	7	A32	A43	2415	A61	A73	3	A93	A103	2	A121	34	A143	A152	1	A173	
31	100030	A11	60	A33	A49	6836	A61	A75	3	A93	A101	4	A124	63	A143	A152	2	A173	
32	100031	A12	18	A32	A49	1913	A64	A72	3	A94	A101	3	A121	36	A141	A152	1	A173	
33	100032	A11	24	A32	A42	4020	A61	A73	2	A93	A101	2	A123	27	A142	A152	1	A173	
34	100033	A12	18	A32	A40	5866	A62	A73	2	A93	A101	2	A123	30	A143	A152	2	A173	
35	100034	A14	12	A34	A49	1264	A65	A75	4	A93	A101	4	A124	57	A143	A151	1	A172	
36	100035	A13	12	A32	A42	1474	A61	A72	4	A92	A101	1	A122	33	A141	A152	1	A174	
37	100036	A12	45	A34	A43	4746	A61	A72	4	A93	A101	2	A122	25	A143	A152	2	A172	
38	100037	A14	48	A34	A46	6110	A61	A73	1	A93	A101	3	A124	31	A141	A153	1	A173	
39	100038	A13	18	A32	A43	2100	A61	A73	4	A93	A102	2	A121	37	A142	A152	1	A173	
40	100039	A13	10	A32	A44	1225	A61	A73	2	A93	A101	2	A123	37	A143	A152	1	A173	
41	100040	A12	9	A32	A43	458	A61	A73	4	A93	A101	3	A121	24	A143	A152	1	A173	

Figure 2 : La base de données

- Défis liés à la collecte et à la qualité des données

La collecte de données pour cette étude sur l'évaluation des risques de crédit a été une étape cruciale. Le jeu de données utilisé, provenant de Credit One Bank, contenait un total de 5000 instances avec 20 attributs différents. Parmi ces attributs, 7 étaient de nature numérique, tandis que 13 étaient catégoriques.

Les données catégoriques, bien qu'essentielles pour une évaluation complète du risque de crédit, présentent souvent des défis uniques en termes de traitement et d'analyse. Dans notre cas, ces catégories étaient associées à des libellés au lieu de valeurs numériques directes, ce qui nécessitait une transformation pour les intégrer dans notre modèle.

Pour surmonter ce défi, nous avons opté pour une approche de mapping. Cette opération a été réalisée en utilisant des outils standard tels qu'Excel, où des formules comme "arrayformula" et "vlookup" ont été mises en œuvre. Cette méthodologie a permis de convertir efficacement les catégories en valeurs numériques, ce qui était nécessaire pour l'entraînement de notre modèle d'évaluation des risques de crédit.

De plus, une attention particulière a été portée à la qualité des données. Cela inclut la détection et la gestion des valeurs manquantes, des valeurs aberrantes, ainsi que la validation de la cohérence des données. Cette étape est cruciale pour garantir la robustesse et la fiabilité de notre modèle.

En fin de compte, la combinaison d'une approche de mapping soigneusement mise en œuvre pour les données catégoriques et d'un processus de contrôle de qualité rigoureux nous a permis de disposer d'un jeu de données complet et cohérent pour la construction de notre modèle.

- Techniques de prétraitement des données pour appliquer les scores pondérés :

Pour parvenir à une évaluation précise des risques de crédit, nous avons mis en place des scores pondérés qui ont été calculés en fonction de plusieurs caractéristiques pertinentes du client. Ces scores ont été développés en utilisant des approches spécifiques pour chaque caractéristique, et ils ont ensuite été combinés pour produire un score global qui détermine le niveau de risque de crédit associé à chaque client. Voici comment nous avons élaboré cette approche :

Scores Individuels pour les Caractéristiques Clés : Nous avons établi des scores individuels pour des caractéristiques clés, comme le statut du compte, l'historique de crédit, le montant du crédit, l'épargne, la stabilité de l'emploi, la situation des autres débiteurs ou garants, la durée de résidence actuelle, le type d'emploi et la propriété. Chacun de ces scores a été calculé en fonction des plages de valeurs spécifiques à chaque caractéristique. Par exemple, pour le montant du crédit, nous avons défini des seuils tels que < 0 USD correspond à un score de 1, no checking account à 0, 0 <= … < 10000 à 2, et … >= 10000 USD à 3.

Pondération des Scores Individuels : Après avoir obtenu les scores individuels pour chaque caractéristique, nous avons procédé à une étape cruciale de pondération. Cette pondération reflète l'importance relative de chaque caractéristique dans l'évaluation globale du risque de crédit. Nos poids attribués étaient basés sur une analyse rigoureuse et ont été déterminés comme suit :

Status score : 20%

Credit history score : 20%

Credit Amount score : 13%

Savings score : 10%

Years At Present Employment score : 10%

Other Debtors Guarantors score : 7%

Current Address Years score : 10%

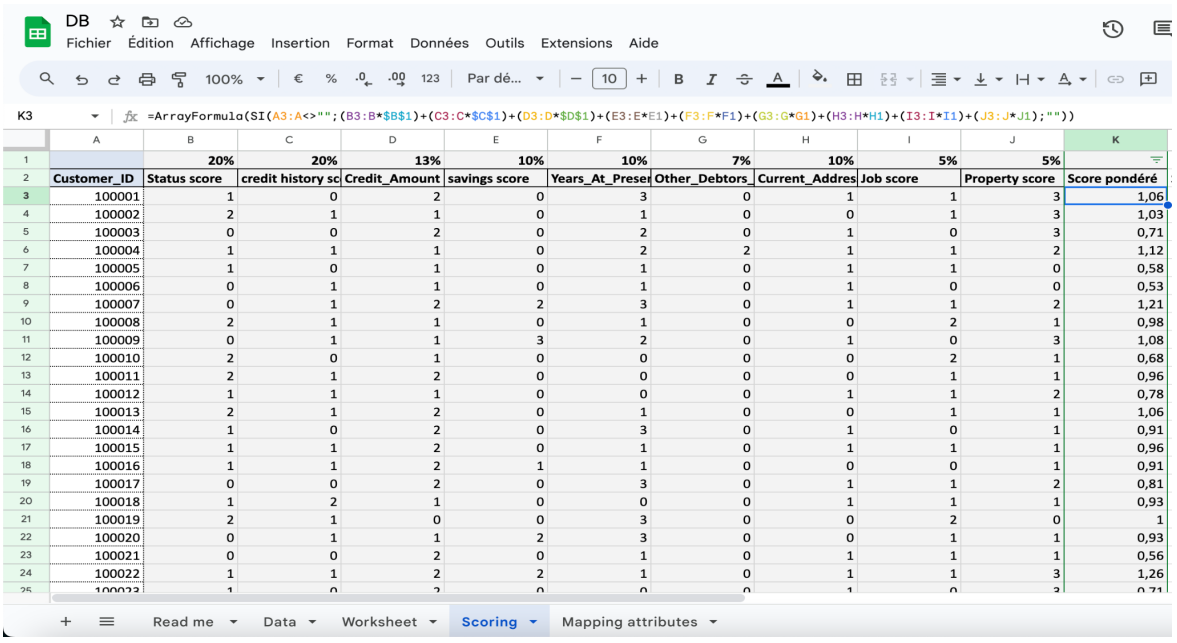
Job score : 5% / Property score : 5%

Calcul du Score Pondéré Global : Enfin, nous avons combiné les scores pondérés individuels pour chaque caractéristique en calculant un score global pour chaque client. Ce score global représente le niveau de risque de crédit pour ce client spécifique. Il est obtenu en sommant les produits des scores individuels par leurs poids correspondants. Par exemple, si un client obtient un score de 2 pour le montant du crédit (avec une pondération de 13%), leur contribution à la note globale sera de 0.26 (2 * 13%)

Interprétation des Scores Pondérés : Les scores pondérés obtenus sont essentiels pour évaluer la probabilité de défaut de crédit. Plus le score est élevé, plus le risque de défaut n'est pas important. Ils servent de base pour la prise de décision sur l'octroi ou le refus de crédit à un client donné, et ils aident également à définir les conditions du prêt, notamment les taux d'intérêt et les limites de crédit.

Mise à Jour et Réajustement : Il est important de noter que ces scores pondérés ne sont pas statiques. Ils peuvent être mis à jour périodiquement pour refléter les changements de situation des clients ou de l'économie en général. Cette approche permet une évaluation continue et précise des risques de crédit.

L'utilisation de scores pondérés basés sur des caractéristiques spécifiques et une pondération réfléchie permet d'obtenir une évaluation robuste et individualisée du risque de crédit pour chaque client.



	A	B	C	D	E	F	G	H	I	J	K
1		20%	20%	13%	10%	10%	7%	10%	5%	5%	
2	Customer_ID	Status score	credit history sc	Credit_Amount	savings score	Years_At_Prese	Other_Debtors_	Current_Addres	Job score	Property score	Score pondéré
3	100001	1	0	2	0	3	0	1	1	3	1,06
4	100002	2	1	1	0	1	0	0	1	3	1,03
5	100003	0	0	2	0	2	0	1	0	3	0,71
6	100004	1	1	1	0	2	2	1	1	2	1,12
7	100005	1	0	1	0	1	0	1	1	0	0,58
8	100006	0	1	1	0	1	0	1	0	0	0,53
9	100007	0	1	2	2	3	0	1	1	2	1,21
10	100008	2	1	1	0	1	0	0	2	1	0,98
11	100009	0	1	1	3	2	0	1	0	3	1,08
12	100010	2	0	1	0	0	0	0	2	1	0,68
13	100011	2	1	2	0	0	0	0	1	1	0,96
14	100012	1	1	1	0	0	0	1	1	2	0,78
15	100013	2	1	2	0	1	0	0	1	1	1,06
16	100014	1	0	2	0	3	0	1	0	1	0,91
17	100015	1	1	2	0	1	0	1	1	1	0,96
18	100016	1	1	2	1	1	0	0	0	1	0,91
19	100017	0	0	2	0	3	0	1	1	2	0,81
20	100018	1	2	1	0	0	0	1	1	1	0,93
21	100019	2	1	0	0	3	0	0	2	0	1
22	100020	0	1	1	2	3	0	0	1	1	0,93
23	100021	0	0	2	0	1	0	1	1	1	0,56
24	100022	1	1	2	2	1	0	1	1	3	1,26

Figure 3 : Database scoring

- Techniques de prétraitement des données pour l'entraînement des modèles d'apprentissage automatique :

Dans cette section, nous explorerons en détail les techniques de prétraitement des données que nous avons utilisées pour préparer notre jeu de données en vue de l'entraînement de notre modèle d'apprentissage automatique. Cela comprendra également une brève introduction à l'apprentissage automatique, l'importance de certaines bibliothèques Python, et comment nous avons converti nos scores pondérés en valeurs binaires pour appliquer un modèle de régression logistique.

L'apprentissage automatique est une branche de l'intelligence artificielle qui permet aux ordinateurs d'apprendre et de s'améliorer à partir de l'expérience sans être explicitement programmés. Dans le contexte de notre étude, l'apprentissage automatique nous permet de construire un modèle prédictif basé sur des données historiques et des scores pondérés, ce qui est essentiel pour évaluer les risques de crédit

Nous avons utilisé plusieurs bibliothèques Python essentielles pour notre analyse et la création de notre modèle. Voici les principales bibliothèques et leur rôle :

pandas : Pour la manipulation des données et la création de DataFrames.

train_test_split de sklearn.model_selection : Pour diviser nos données en ensembles d'entraînement et de test.

LogisticRegression de sklearn.linear_model : Pour créer et entraîner notre modèle de régression logistique.

StandardScaler de sklearn.preprocessing : Pour la mise à l'échelle de nos caractéristiques.

accuracy_score et confusion_matrix de sklearn.metrics : Pour évaluer la performance de notre modèle.

LabelEncoder de sklearn.preprocessing : Pour convertir nos scores pondérés en valeurs binaires.

r2_score de sklearn.metrics : Pour calculer le coefficient de détermination.

seaborn, plotly.express, matplotlib.pyplot, matplotlib.gridspec : Pour créer des visualisations

et des graphiques pour une meilleure compréhension de nos données.

Nous avons commencé par diviser notre ensemble de données en deux parties : datascore et data. Datascore a été utilisé pour entraîner notre modèle, tandis que data a servi à créer des visualisations. Nous avons retiré les colonnes non nécessaires de chaque ensemble.

```
ntrée [46]:
##pour avoir les noms de colonnes
db.columns

Out[46]: Index(['Customer_ID', 'Status_Checking_Acc', 'Duration_in_Months',
       'Credit_History', 'Purposre_Credit_Taken', 'Credit_Amount',
       'Savings_Acc', 'Years_At_Present_Employment', 'Inst_Rt_Income',
       'Marital_Status_Gender', 'Other_Debtors_Guarantors',
       'Current_Address_Yrs', 'Property', 'Age', 'Other_Inst_Plans', 'Housing',
       'Num_CC', 'Job', 'Dependents', 'Telephone', 'Foreign_Worker',
       'Default_On_Payment', 'Customer_ID.1', 'Status score',
       'credit history score', 'Credit_Amount score', 'savings score',
       'Years_At_Present_Employment score', 'Other_Debtors_Guarantors score',
       'Current_Address_Yrs score', 'Job score', 'Property score',
       'Score pondéré', 'Score pondéré binaire'],
      dtype='object')

ntrée [72]:
##Diviser notre data
##db = datascore + data
##datascore pour entrainer notre model et data pour creer des visualisations

datascore =db.drop(['Customer_ID', 'Status_Checking_Acc', 'Duration_in_Months',
       'Credit_History', 'Purposre_Credit_Taken', 'Credit_Amount',
       'Savings_Acc', 'Years_At_Present_Employment', 'Inst_Rt_Income',
       'Marital_Status_Gender', 'Other_Debtors_Guarantors',
       'Current_Address_Yrs', 'Property', 'Age', 'Other_Inst_Plans', 'Housing',
       'Num_CC', 'Job', 'Dependents', 'Telephone', 'Foreign_Worker',
       'Default_On_Payment', 'Customer_ID.1'],axis=1)
data = db.drop(['Customer_ID','Customer_ID.1', 'Status score',
       'credit history score', 'Credit_Amount score', 'savings score',
       'Years_At_Present_Employment score', 'Other_Debtors_Guarantors score',
       'Current_Address_Yrs score', 'Job score', 'Property score',
       'Score pondéré', 'Score pondéré binaire'],axis=1)
```

Figure 4 : Data base diviser

```
##Quelques informations statistiques sur notre dataset
db.describe()
```

	Customer_ID	Duration_in_Months	Credit_Amount	Inst_Rt_Income	Current_Address_Yrs	Age	Num_CC	Dependents	Customer_ID.1	
count	5000.000000	5000.000000	5000.000000	5000.000000	5000.000000	5000.000000	5000.000000	5000.000000	5000.000000	5000.0
mean	102500.500000	20.903000	3271.258000	2.973000	2.845000	35.546000	1.407000	1.155000	102500.500000	1.0
std	1443.520003	12.053989	2821.607329	1.118267	1.103276	11.370917	0.577423	0.361941	1443.520003	0.9
min	100001.000000	4.000000	250.000000	1.000000	1.000000	19.000000	1.000000	1.000000	100001.000000	0.0
25%	101250.750000	12.000000	1365.500000	2.000000	2.000000	27.000000	1.000000	1.000000	101250.750000	0.0
50%	102500.500000	18.000000	2319.500000	3.000000	3.000000	33.000000	1.000000	1.000000	102500.500000	1.0
75%	103750.250000	24.000000	3972.250000	4.000000	4.000000	42.000000	2.000000	1.000000	103750.250000	2.0
max	105000.000000	72.000000	18424.000000	4.000000	4.000000	75.000000	4.000000	2.000000	105000.000000	3.0

Figure 5 : description de la Data base

```
Entrée [138]:
1 ##La somme de variables null
2 db.isnull().sum()

Out[138]:
Customer_ID                          0
Status_Checking_Acc                  0
Duration_in_Months                   0
Credit_History                       0
Purposre_Credit_Taken                0
Credit_Amount                        0
Savings_Acc                          0
Years_At_Present_Employment          0
Inst_Rt_Income                       0
Marital_Status_Gender                0
Other_Debtors_Guarantors             0
Current_Address_Yrs                  0
Property                             0
Age                                  0
Other_Inst_Plans                     0
Housing                              0
Num_CC                               0
Job                                  0
Dependents                           0
Telephone                            0
Foreign_Worker                       0
Default_On_Payment                   0
Customer_ID.1                        0
Status score                         0
credit history score                 0
Credit_Amount score                  0
savings score                        0
Years_At_Present_Employment score    0
Other_Debtors_Guarantors score       0
Current_Address_Yrs score            0
Job score                            0
Property score                       0
Score pondéré                        0
```

On peut constater que notre base de données a aucune valeur null, ce qui est bien pour entraîner notre model

Figure 6 : Data base informations

Conversion en Valeurs Binaires :

Pour appliquer notre modèle de régression logistique, nous avons converti les scores pondérés calculés précédemment en valeurs binaires. Ceci était essentiel pour la classification binaire du risque de crédit. Un score de 0 signifie qu'il y a un risque de crédit, tandis qu'un score de 1 signifie qu'il n'y a pas de risque de crédit.

L3 =ArrayFormula(SI(K3:K<>"";SI(K3:K>0,9;1;0);""))

	K	L	M	N
1				
2	**Score pondéré**	Score pondéré binaire		
3	1,06	1		
4	1,03	1		
5	0,71	0		
6	1,12	1		
7	0,58	0		
8	0,53	0		
9	1,21	1		
10	0,98	1		
11	1,08	1		
12	0,68	0		
13	0,96	1		
14	0,78	0		
15	1,06	1		
16	0,91	1		
17	0,96	1		
18	0,91	1		
19	0,81	0		
20	0,93	1		
21	1	1		
22	0,93	1		

+ ≡ Read me Data Worksheet Scoring

Figure 7 : Conversion en valeurs binaires

Nous avons réalisé une analyse de corrélation pour explorer les relations entre les variables de notre jeu de données. Les résultats de cette analyse ont mis en évidence des tendances intéressantes. En particulier, nous avons observé une corrélation significative entre deux variables spécifiques : le montant du crédit (Credit Amount) et la durée en mois (Duration in Months).

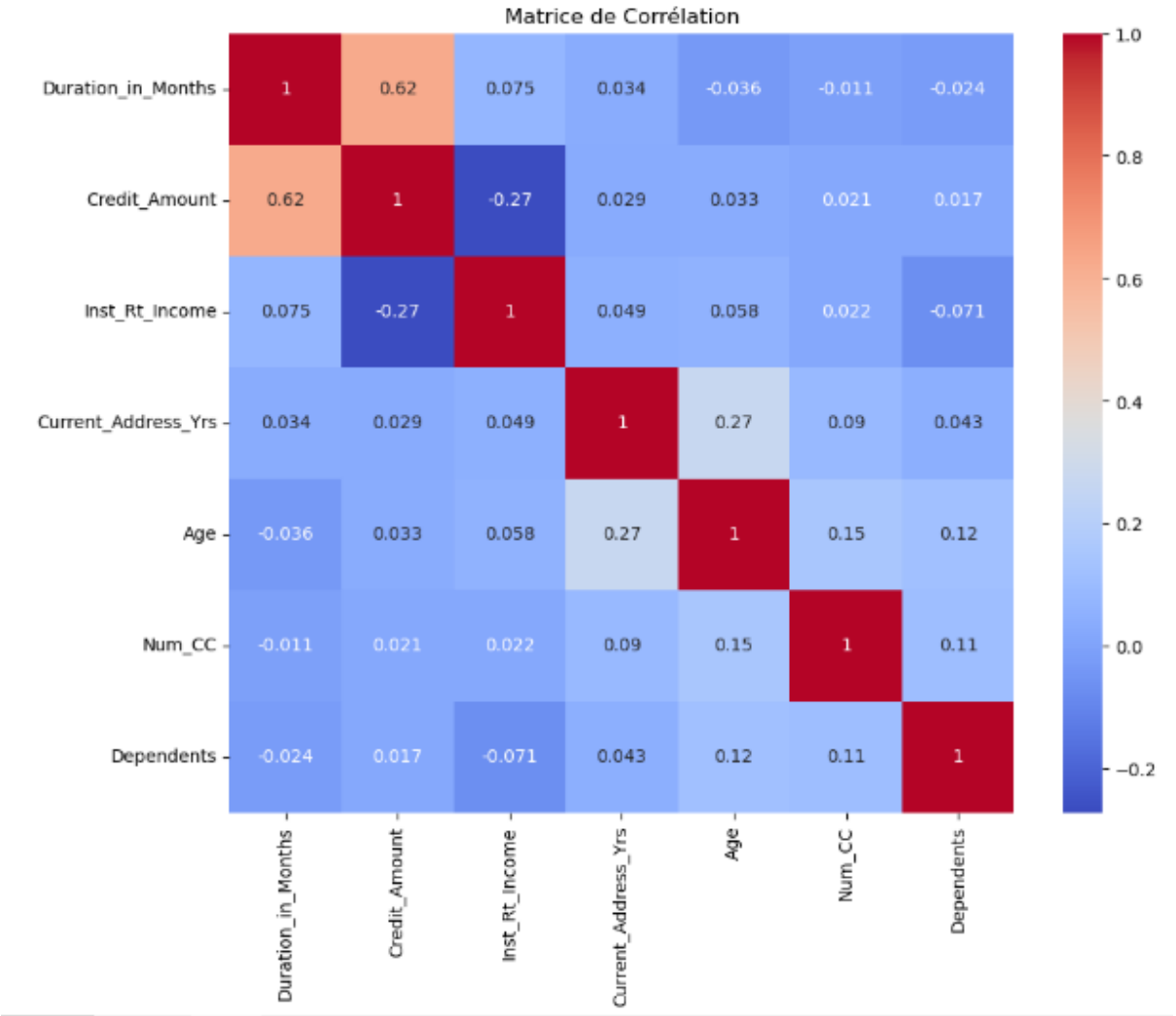


Figure 8 : Corrélation des variables

Plus précisément, il est apparu que le montant du crédit emprunté par un client et la durée de ce crédit en mois sont fortement corrélés positivement. Cela signifie que, en général, à mesure que le montant du crédit augmente, la durée du prêt a tendance à augmenter également. Cette observation suggère une relation importante entre ces deux variables, ce qui peut avoir des implications pour notre analyse des risques de crédit.

Cependant, il est à noter que pour la plupart des autres variables de notre ensemble de données, nous n'avons pas observé de corrélations significatives, ou même des corrélations négatives. Cela indique que ces variables peuvent être relativement indépendantes les unes des autres et du résultat que nous cherchons à prédire, à savoir le risque de crédit. Cette diversité de corrélations (ou de l'absence de corrélations) entre les variables souligne l'importance d'examiner chaque variable individuellement pour comprendre sa contribution à notre modèle de prédiction.

C'est pourquoi nous avons opté pour la création de visualisations graphiques, telles que des graphiques et des diagrammes, afin d'explorer plus en profondeur les relations entre les variables.

Par exemple, nous avons examiné la relation entre le montant du crédit et l'âge des emprunteurs. Nos visualisations ont révélé que la tranche d'âge la plus propice à la souscription de crédits se situe généralement entre 20 et 42 ans, avec des montants de crédit allant de 1000 à 6000 dollars.

```
### Créer de visualisation pour voir la Relation entre Montant du Crédit et Âge
fig = plt.figure(figsize=(36, 12))
spec = gridspec.GridSpec(2, 2, figure=fig)

ax1 = fig.add_subplot(spec[0, 0])
ax1.scatter(db["Credit_Amount"],db["Age"])
ax1.set_title("Relation entre Montant du Crédit et Âge")

```

```
Text(0.5, 1.0, 'Relation entre Montant du Crédit et Âge')
```

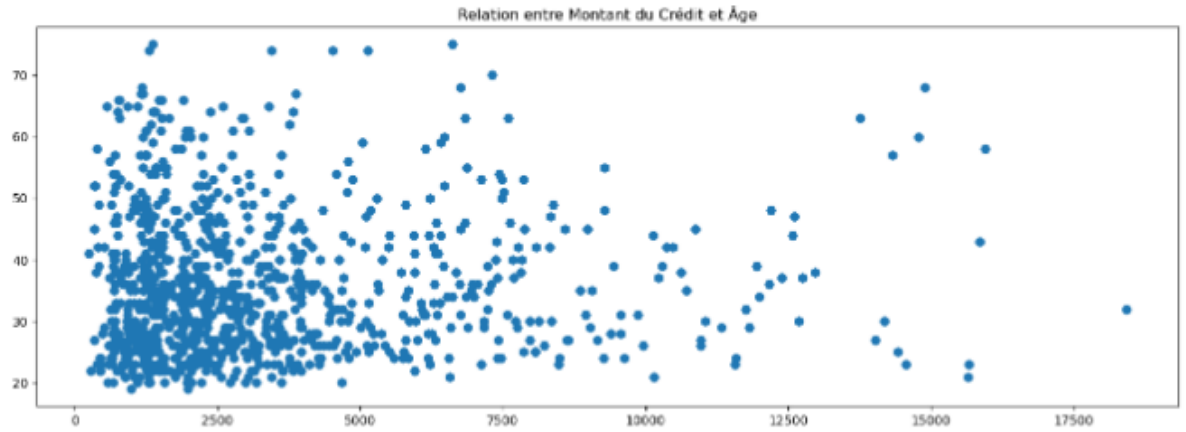


Figure 9 : Relation entre Montant de crédit et Âge

```
## Distribution de la Durée en Mois
fig = plt.figure(figsize=(36, 12))
spec = gridspec.GridSpec(2, 2, figure=fig)

ax2 = fig.add_subplot(spec[0, 1])
ax2.hist(db["Duration_in_Months"], bins=20, color="blue")
ax2.set_title("Distribution de la Durée en Mois")
```

```
Text(0.5, 1.0, 'Distribution de la Durée en Mois')
```

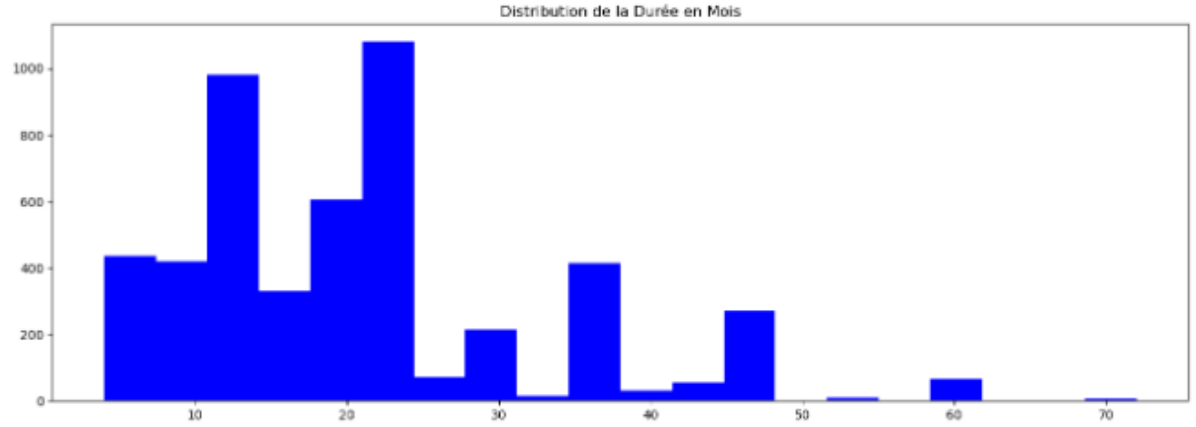


Figure 10 : Distribution de la durée en mois

```
##Montant du Crédit Moyen par Type d'Emploi et Statut Marital
pivot_table = db.pivot_table(index="Job", columns="Marital_Status_Gender", values="Credit_Amount", aggfunc="mean")
pivot_table.plot(kind="bar", cmap="Set2")
plt.title("Montant du Crédit Moyen par Type d'Emploi et Statut Marital")
plt.ylabel("Montant Moyen du Crédit")
plt.xlabel("Type d'Emploi")
plt.show()
```

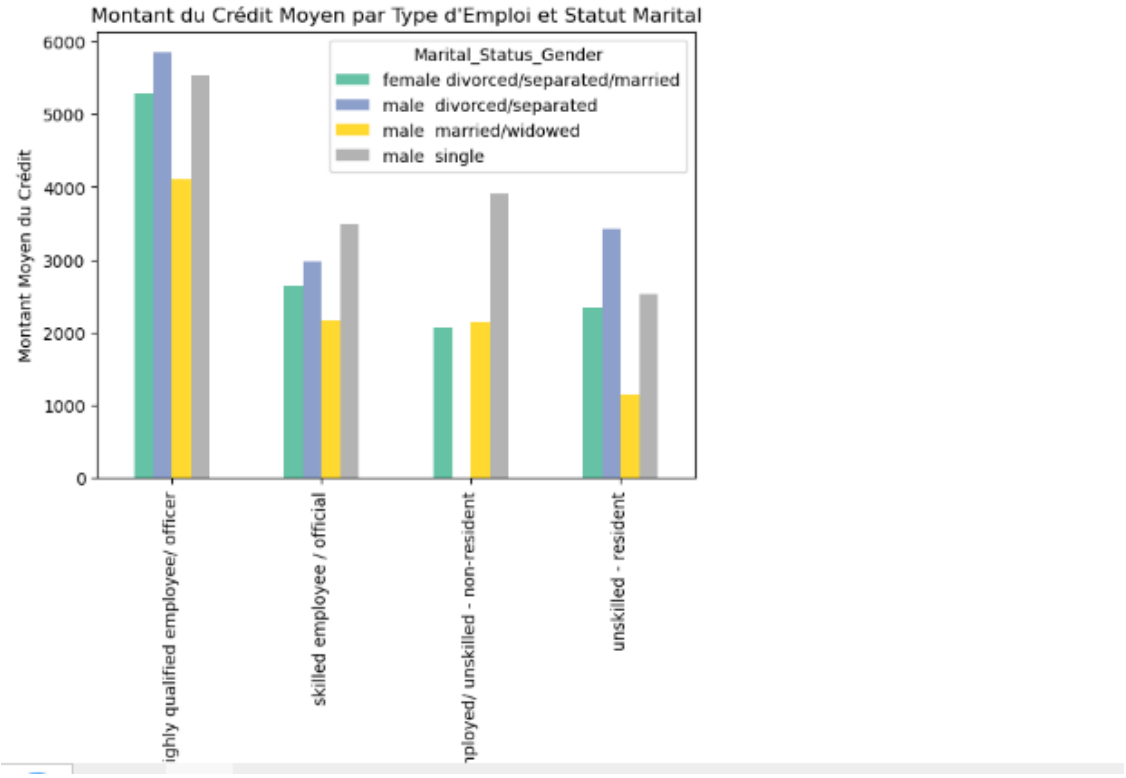


Figure 11 : Montant de crédit Moyen par type d'emploi et statut martial

De plus, nous avons également analysé la relation entre le montant moyen du crédit, le type d'emploi et le statut marital. Les résultats de cette analyse ont montré que les individus hautement qualifiés ont tendance à emprunter davantage, tandis que ceux ayant des compétences moins élevées sont moins enclins à souscrire des crédits. Il est intéressant de noter que, quelle que soit la catégorie professionnelle, ce sont généralement les hommes divorcés ou séparés qui présentent le plus fort taux d'emprunt.

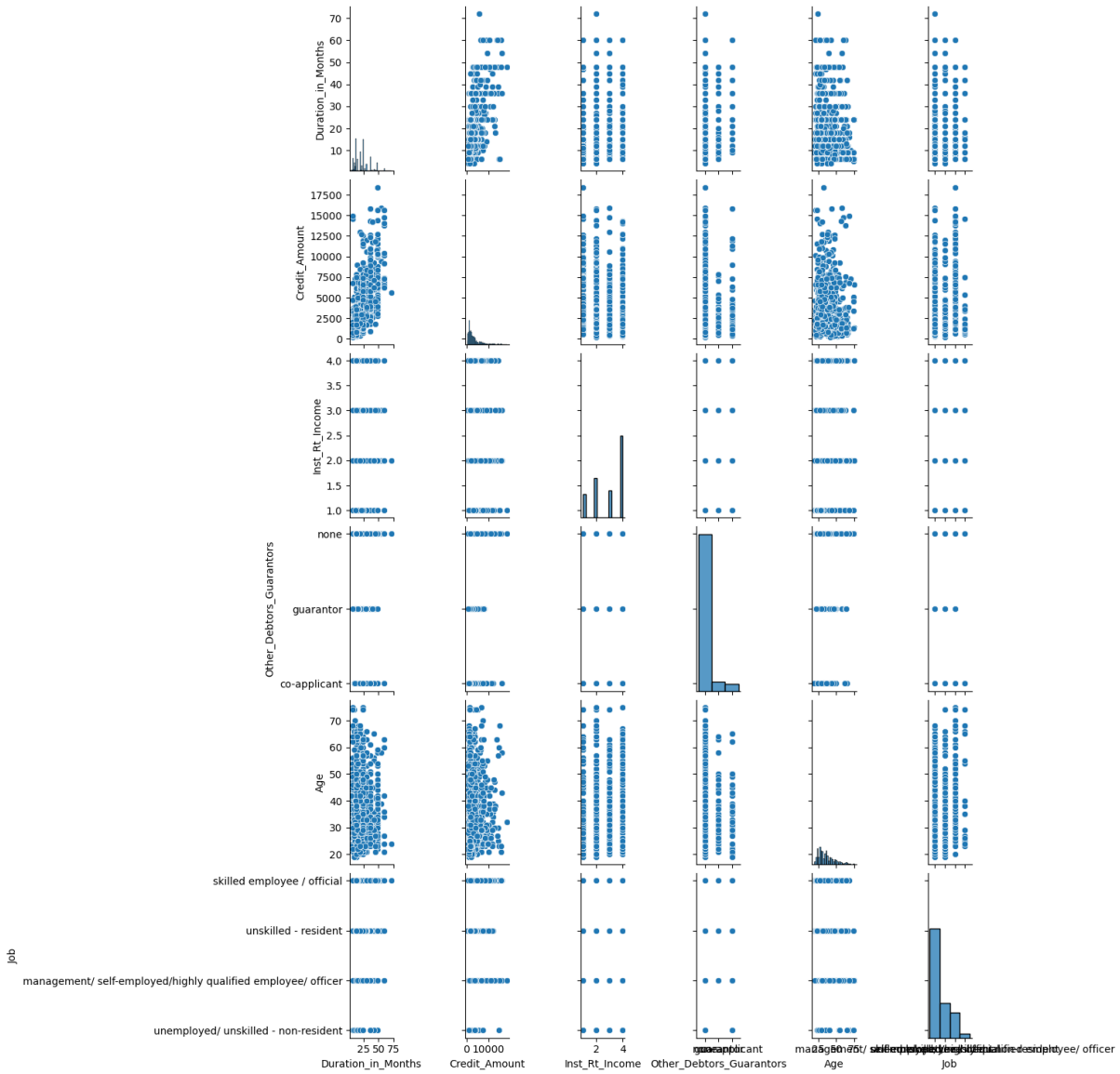


Figure 12 : Score pondéré binaire avec quelques variables

Nous avons mis en application notre modèle d'apprentissage automatique, en particulier la régression logistique. Pour cela, nous avons préparé nos données pour l'entraînement du modèle. Notre ensemble de données se compose de deux parties principales : les caractéristiques (X) et la variable cible (Y)

'Customer_ID' : Cet attribut est un identifiant client et n'est pas pertinent pour la prédiction, nous l'avons donc exclu.

Les scores des différentes caractéristiques : Nous avons utilisé les scores que nous avions calculés précédemment pour chaque client, y compris le score de statut, le score d'historique de crédit, le score du montant du crédit, le score d'épargne, le score d'ancienneté de l'emploi, le score d'autres débiteurs ou garants, le score d'adresse actuelle, le score d'emploi, le score de propriété, et enfin le score pondéré, qui représente la synthèse de tous ces scores.

La variable cible (Y) que nous cherchons à prédire est :

'Score pondéré binaire' : Cette variable est binaire, où 0 indique un risque de crédit et 1 indique l'absence de risque de crédit. Elle est cruciale pour notre objectif, car elle nous permet de déterminer si un client représente un risque de défaut de paiement.

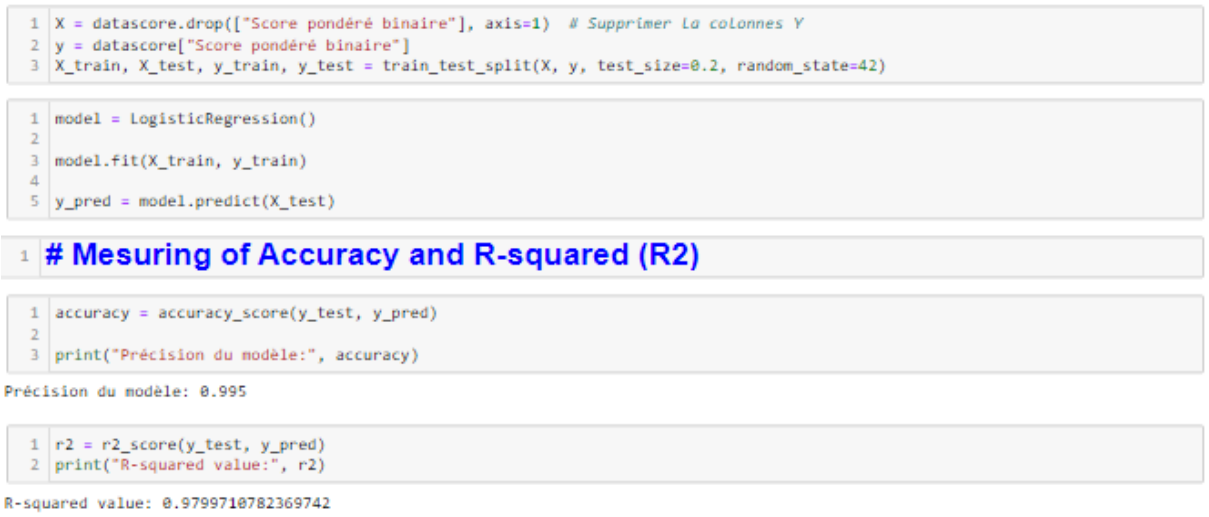


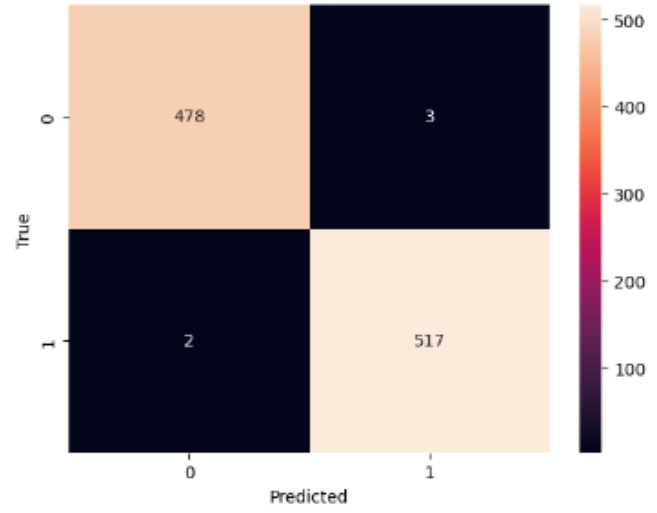
Figure 13 : Modèle de régression logistique

En ce qui concerne les résultats de notre modèle, nous avons obtenu une précision de modèle exceptionnellement élevée de 0,995, ce qui signifie que notre modèle a

correctement classé 99,5 % des échantillons de test. Cela suggère que notre modèle est très précis dans la prédiction de la solvabilité des clients.

Le R-carré (R-squared) est une mesure de la qualité de l'ajustement de notre modèle aux données. Une valeur de 1 indique un ajustement parfait, et une valeur proche de 1 indique que notre modèle explique la variance dans les données de manière significative. Dans notre cas, le R-carré est de 0,9799, ce qui est très élevé. Cela signifie que notre modèle explique près de 98 % de la variance dans les données, ce qui est un excellent résultat.

```
cm = confusion_matrix(y_test, y_pred)
sns.heatmap(cm, annot=True, fmt="d")
plt.xlabel('Predicted')
plt.ylabel('True')
plt.show()
```




```
La matrice de confusion est une table qui est souvent utilisée pour évaluer les performances d'un modèle de classification
(comme votre modèle de régression logistique). Elle permet de visualiser la manière dont le modèle classe les échantillons
par rapport à leurs véritables étiquettes.

         Prédit Positif   Prédit Négatif
Vrai Positif       VP             FP
Vrai Négatif       FN             VN

Vrais Positifs (VP) : Ce sont les cas où le modèle a correctement prédit la classe positive "risque de crédit" comme
positifs. => il n'ya pas de risque de crédit

Vrais Négatifs (VN) : Ce sont les cas où le modèle a correctement prédit la classe négative "risque de crédit" comme
négatifs. => il ya un risque de crédit

Faux Positifs (FP) : Ce sont les cas où le modèle a incorrectement prédit la classe négative comme positive
Faux Négatifs (FN) : Ce sont les cas où le modèle a incorrectement prédit la classe positive comme négative.
```

Figure 14 : Matrice de Confusion

La matrice de confusion est une table qui est souvent utilisée pour évaluer les performances d'un modèle de classification (comme votre modèle de régression logistique). Elle permet de visualiser la manière dont le modèle classe les échantillons par rapport à leurs véritables étiquettes.

	Prédit Positif	Prédit Négatif
Vrai Positif	VP : 478	FP : 3
Vrai Négatif	FN : 2	VN : 517

Tableau 2 : Matrice de confusion

Vrais Positifs (VP) : Ce sont les cas où le modèle a correctement prédit la classe positive "risque de crédit" comme positifs. => il n'ya pas de risque de crédit

Vrais Négatifs (VN) : Ce sont les cas où le modèle a correctement prédit la classe négative "risque de crédit" comme négatifs. => il ya un risque de crédit

Faux Positifs (FP) : Ce sont les cas où le modèle a incorrectement prédit la classe négative comme positive

Faux Négatifs (FN) : Ce sont les cas où le modèle a incorrectement prédit la classe positive comme négative.

Bibliographie

AMATO J. D. et FURFINE C. H., « Are credit ratings procyclical? », *Journal of Banking & Finance*, vol. 28, n° 11, 2004, p. 2641-2677.

BHORAJ S. et SENGUPTA P., « Effect of Corporate Governance on Bond Ratings and Yields: The Role of Institutional Investors and Outside Directors », *The Journal of Business*, n° 76, 2003, p. 455-475.

CANTOR R. et PARKER F., « The credit rating industry », *FRBNY Quarterly Review*, Summer-Fall, 1994, p. 1-26.

DEGOS J.-G., BEN HMIDEN O. et HENCHIRI J. E., « Les agences de notation financières. Naissance et évolution d'un oligopole controversé », *Revue française de gestion*, n° 227, 2012, p. 45-65.

EDERINGTON L. H., « Classification Models and Bond Ratings », *Financial Review*, vol. 20, n° 4, 1985, p. 237-262.

FERRI G. et LIU L. G., « Assessing the effort of rating agencies in emerging economies: Some emprirical evidence », *European Journal of Finance*, vol. 11, n° 3, 2005, p. 283-295.

GONZALEZ A. et DESCAMPS-JULIEN B., « Population and community variability in randomly fluctuating environments », *Oikos*, vol. 106, n° 1, 2004, p. 105-116.

HOWELL D., *Méthodes statistiques en sciences humaines*, Louvain-la-Neuve, De Boeck, 1998.

KRAHNEN J. et WEBER M., « Generally accepted rating principles: A primer », *Journal of Banking & Finance*, vol. 25, n° 1, 2001, p. 3-23.

MATTHIES Karsten, DAS Amit, ZIMMER Johannes *et al*., « Can equations of equilibrium predic all physical equilibria? A case study from Field Dislocation Mechanics », *Mathematics and Mechanics of Solids*, vol. 18, n° 8, 2013, p. 803-822.

N.S. J., MAKANY J. et GABSOUBO YIENEZOUNE C., « L'évaluation du risque de crédit des entreprises : cas de la banque congolaise de l'habitat », *Revue Congolaise de Gestion*, n° 17, 2013, p. 87-130.

PAGET-BLANC E. et PAINVIN N., *La notation financière : rôle des agences et méthodes de notation*, Malakoff, Dunod, 2007.

Sitographie

Anonyme, « Comment la crise de 2008 a-t-elle commencé ? », *La finance pour tous* [en ligne], janvier 2023 [consulté le 15/08/2023], disponible sur : https://www.lafinancepourtous.com/decryptages/crises-economiques/crise-des-subprimes/crise-financiere/comment-la-crise-de-2008-a-t-elle-commence/.

BOLUZE L., « Risque de crédit : définition, types et évaluation », *Capital* [en ligne], mars 2022 [consulté le 15/08/2023], disponible sur : https://www.capital.fr/economie-politique/risque-de-credit-definition-types-et-evaluation-1431056.

Our dataset : https://docs.google.com/spreadsheets/d/1hW0RDGBD63u14RYA57g4V50tKPH_4mI2ZxMvF1UhbBc/edit?usp=sharing

Liste des figures

Liste des tableaux