

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique

Ecole Nationale Supérieure de Management
Koléa



وزارة التعليم العالي و البحث العلمي
المدرسة الوطنية العليا للمناجمت
القلية

Projet de fin d'études

En vue de l'obtention d'un diplôme de Master en Management

Spécialité : Management Financier des Entreprises.

Modélisation prédictive en assurance-crédit à l'exportation

Evaluation de l'acheteur étranger

Cas : Compagnie Algérienne d'Assurance et de Garantie des Exportations (CAGEX)

Élaboré par :

ATTIA Mounira

Encadré par :

Pr. ZEROUTI Messaoud

Année universitaire 2025 – 2026

RESUME

L'assurance-crédit à l'exportation permet de protéger les exportateurs contre le risque de non-paiement des acheteurs étrangers. Dans ce cadre, l'évaluation du risque acheteur constitue un enjeu important pour les organismes d'assurance-crédit, notamment face aux limites des méthodes traditionnelles d'analyse.

Cette recherche a pour objectif d'étudier l'apport du Machine Learning dans l'évaluation prédictive du risque de défaut des acheteurs étrangers au sein de la CAGEX. Pour cela, une approche quantitative a été adoptée à partir d'une base de données semi-synthétique regroupant des variables financières, macroéconomiques et transactionnelles. Trois modèles ont été comparés : la régression logistique, XGBoost et LightGBM.

Les résultats montrent que les modèles de Machine Learning, particulièrement XGBoost et LightGBM, offrent de meilleures performances prédictives que la régression logistique. L'étude met également en évidence l'importance des variables financières et comportementales dans l'évaluation du risque acheteur.

En conclusion, le Machine Learning représente un outil pertinent pour améliorer l'évaluation du risque dans l'assurance-crédit à l'exportation et renforcer la qualité des décisions de couverture au sein de la CAGEX.

Mots Clés : Assurance-crédit à l'exportation, Risque de non-paiement, Acheteur étranger, Machine Learning, régression logistique, Algorithme XGBoost, Algorithme LightGBM, CAGEX.

ABSTRACT

Export credit insurance protects exporters against the risk of non-payment by foreign buyers. In this context, buyer risk assessment represents a major challenge for export credit insurance institutions, especially given the limitations of traditional analysis methods.

This research aims to study the contribution of Machine Learning to the predictive assessment of foreign buyers' default risk within CAGEX. To achieve this, a quantitative approach was adopted based on a semi-synthetic database including financial, macroeconomic, and transactional variables. Three models were compared: Logistic Regression, Boost, and LightGBM.

The results show that Machine Learning models, particularly XGBoost and LightGBM, provide better predictive performance than Logistic Regression. The study also highlights the importance of financial and behavioral variables in buyer risk assessment.

In conclusion, Machine Learning represents a relevant tool for improving risk assessment in export credit insurance and strengthening the quality of coverage decisions within CAGEX.

Key Words: Export credit insurance, Risk of non-payment, Foreign buyers, Machine Learning, Logistic Regression, Algorithm XGBoost, Algorithm LightGBM, CAGEX.

ملخص

يسمح التأمين على الصادرات بحماية المصدرين من خطر عدم سداد المشتريين الأجانب. وفي هذا الإطار، يُعدّ تقييم مخاطر المشتري من القضايا الأساسية بالنسبة لمؤسسات التأمين على الصادرات، خاصة في ظل محدودية أساليب التحليل التقليدية.

تهدف هذه الدراسة إلى إبراز مساهمة التعلم الآلي في التقييم التنبؤي لخطر تعثر المشتريين الأجانب داخل الشركة الجزائرية للتأمين وضمن الصادرات (CAGEX). ولتحقيق ذلك، تم اعتماد منهج كمي بالاعتماد على قاعدة بيانات شبه اصطناعية تضم متغيرات مالية واقتصادية ومعاملاتية. كما تمت مقارنة ثلاثة نماذج وهي: الانحدار اللوجستي، وXGBoost، وLightGBM.

أظهرت النتائج أن نماذج التعلم الآلي، خاصة XGBoost وLightGBM، توفر أداءً تنبؤياً أفضل مقارنة بالانحدار اللوجستي. كما أبرزت الدراسة أهمية المتغيرات المالية والسلوكية في تقييم مخاطر المشتري.

وفي الختام، يُعتبر التعلم الآلة أداة فعّالة لتحسين تقييم المخاطر في التأمين على الصادرات وتعزيز جودة قرارات التغطية داخل CAGEX.

الكلمات المفتاحية: تأمين على الصادرات، خطر عدم السداد، مشتريين أجانب، تعلم الآلة، انحدار لوجستي، خوارزمية XGBoost، خوارزمية LightGBM، الشركة الجزائرية للتأمين وضمن الصادرات (CAGEX).

REMERCIEMENT

Tout d'abord, je tiens à remercier Dieu, le Tout-Puissant, de m'avoir accordé la santé, le courage et la volonté nécessaires pour mener à bien ce travail.

Je tiens à adresser mes plus tendres remerciements à ma famille, mon pilier le plus solide. À ma mère, mon père : votre amour inconditionnel, vos sacrifices, vos encouragements et votre confiance en moi, même dans les moments de doute, ont été ma plus grande force. Merci d'être toujours présents, d'être mon refuge et mon moteur.

Je souhaite exprimer ma profonde gratitude à mon encadrant interne, **Pr. Zerouti Messaoud**, pour ses conseils avisés, sa disponibilité et sa bienveillance tout au long de ce projet. Ses orientations ont grandement contribué à la réussite de ce mémoire.

Un immense merci à mes tuteurs de stage, **M. Kashi Hakim** et **Mme Bensalem Amel**. Votre disponibilité, votre patience, vos précieux retours et votre soutien constant ont été déterminants. Vous m'avez guidé avec générosité, rigueur et humanité, et je vous en suis profondément reconnaissante.

À mes chers amies, Tiago et Rania, merci du fond du cœur. Vos messages motivants, votre bonne humeur, vos éclats de rire et votre énergie positive ont illuminé cette période parfois éprouvante. Vous m'avez rappelé que je n'étais jamais seule.

Enfin, je remercie chaleureusement toutes les personnes qui, de près ou de loin, ont contribué à la réalisation de ce projet de fin d'études. Chaque conseil, chaque geste et chaque parole bienveillante ont laissé leur empreinte dans ce travail.

Ce mémoire n'est pas seulement le fruit de mon effort, il est aussi le vôtre.

Du fond du cœur, merci.

TABLE DES MATIERES

RESUME	I
ABSTRACT	II
ملخص.....	III
REMERCIEMENT	IV
LISTE DES FIGURES.....	VII
LISTE DES TABLEAUX	VIII
LISTE DES ABREVIATION	IX
INTRODUCTION GENERALE	1
CHAPITRE I : CADRE THEORIQUE.....	7
I.Revue de littérature :.....	8
II.LE CADRE CONCEPTUEL	14
1. Delimitation theorique de l'assurance-credit à l'export	14
2. Le concept d'acheteur étranger et son évaluation dans l'assurance-crédit :	20
3. Le machine learning comme outil d'amélioration de l'évaluation de l'acheteur étranger... ..	28
4. L'explicabilité comme condition d'acceptation des modèles prédictifs	32
CHAPITRE II : CADRE METHODOLOGIQUE ET ORGANISATIONNEL	41
I. Cadre méthodologique :	42
II. Contexte de la recherche	47
1. L'assurance-crédit à l'exportation, réglementation et adoption des modèles avancés :.....	47
2. Présentation de la CAGEX	50
CHAPITRE III : RESULTATS ET DISCUSSION	56
I. Présentation des résultats :	57
1. Analyse descriptive de la base :	57
2. Dictionnaire des variables :	58
3. Préparation des données et choix méthodologiques :	61
4. Interprétation détaillée des résultats par modèle :.....	64
5. Test des hypothèses :	85
II. Discussion des résultats :.....	94
CONCLUSION GENERALE.....	98
Bibliographie.....	93

ANNEXE Unique Organigramme de la CAGEX97

LISTE DES FIGURES

Figure 1: Schéma représentatif du mécanisme de l'assurance-crédit à l'exportation	16
Figure 2: Représente une Cartographie de Classification des pays.....	18
Figure 3: Représente Les partenaires de la CAGEX	55
Figure 4: Représente la Distribution de la variable cible	61
Figure 5: le résultat du modèle de régression logistique simple.....	65
Figure 6: représente la courbe ROC-AUC de régression logistique	66
Figure 7: représente le résultat du modèle de régression logistique optimisé	67
Figure 8: représente la matrice de confusion du modèle de régression logistique optimisé	68
Figure 9: Représente la courbe ROC-AUC de la régression logistique optimisée.....	69
Figure 10: Représente la courbe ROC-AUC de la régression logistique optimisée.....	70
Figure 11: Représente la matrice de confusion de la régression logistique +SMOTE.....	71
Figure 12: Représente la courbe ROC-AUC de la régression logistique +SMOTE	71
Figure 13: Représente les résultats de XGBoost +SMOTE	72
Figure 14: Représente la matrice de confusion de XGBoost +SMOTE.....	73
Figure 15: Représente la courbe Precision-recall de XGBoost +SMOTE	74
Figure 16: Représente le résultat de XGBoost +SMOTE avec changement de seuil.....	75
Figure 17: Représente la matrice de confusion avant et après le changement de seuil.....	76
Figure 18 : Représente la matrice de confusion avant et après le changement de seuil.....	76
Figure 19: Représente la courbe precision-recall	77
Figure 20: Représente les graphes après le tuning	79
Figure 21: Représente les variables les plus importantes dans le modèle XGBoost.....	80
Figure 22: Représente les résultats obtenus du modèle LightGBM	82
Figure 23: Représente la matrice de confusion avant et après l'optimisation en LightGBM	82
Figure 24: Représente la courbe ROC-AUC en LightGBM.....	83
Figure 25: Représente les variables importantes en LightGBM.....	84

LISTE DES TABLEAUX

Tableau 1: Représente les principaux ratios sélectionnés par Beaver	22
Tableau 2 : Représente Les ratios choisis par Altman.....	23
Tableau 3: Représente Les indicateurs choisis par Ohlson	24
Tableau 4: Représente les produits fournis par la CAGEX.....	53
Tableau 5: Représente le contenu général de la base de données	58
Tableau 6:Représente le dictionnaire des principales variables de la base	59
Tableau 7:Représente les variables tenues pour la création de modèle.....	62
Tableau 8: Représente le découpage de la base de données.....	63
Tableau 9:Représente Comparaison des modèles prédictifs pour la détection des défauts des acheteurs étrangers	90

LISTE DES ABREVIATION

AI	Artificial Intelligence (Intelligence Artificielle)
ATIDI	African Trade & Investment Development Insurance
AUC	Area Under the Curve (Aire sous la courbe)
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
CAGEX	Compagnie Algérienne d'Assurance et de Garantie des Exportations
CHIN	Change in Net Income (Variation du résultat net entre deux ans)
CLCA	Current Liabilities / Current Assets (Passif circulant / Actif circulant)
CNIL	Commission Nationale de l'Informatique et des Libertés
COFACE	Compagnie Française d'Assurance pour le Commerce Extérieur
CRAM	Country Risk Assessment Model (Modèle d'évaluation du risque pays de l'OCDE)
DNB	De Nederlandsche Bank
DSCR	Debt Service Coverage Ratio (Taux de couverture des dettes)
EBA	European Banking Authority (Autorité Bancaire Européenne)
EBIT	Earnings Before Interest and Taxes (Résultat avant intérêts et impôts)
EBITDA	Earnings Before Interest, Taxes, Depreciation and Amortization
ECA	Export Credit Agency (Agence de crédit à l'exportation)
ECIC	Export Credit Insurance Corporation of South Africa
EIOPA	European Insurance and Occupational Pensions Authority
ENEG	Situation Nette Négative (variable binaire : 1 si dettes > Actifs)
EU	European Union (Union Européenne)
EXIM	Export-Import Bank of the United States
FIS	Fuzzy Inference System (Système d'inférence floue)
FUTL	Funds from Operations / Total Liabilities (Flux de fonds / Total des dettes)
IA	Intelligence Artificielle
IFC	International Finance Corporation

INTWO	Variable binaire : 1 si l'entreprise enregistre des pertes pendant deux années consécutives
IRB	Internal Ratings-Based Approach (Approche fondée sur les notations internes)
LIME	Local Interpretable Model-Agnostic Explanations
MDPI	Multidisciplinary Digital Publishing Institute
MIGA	Multilateral Investment Guarantee Agency
ML	Machine Learning (Apprentissage automatique)
NEXI	Nippon Export and Investment Insurance
NITA	Net Income / Total Assets (Résultat net / Total actif)
OCDE	Organisation de Coopération et de Développement Économiques
OECD	Organisation for Economic Co-operation and Development (voir OCDE)
OMC	Organisation Mondiale du Commerce
RGPD	Règlement Général sur la Protection des Données
ROA	Return on Assets (Rentabilité des actifs)
ROC	Receiver Operating Characteristic (Courbe caractéristique de fonctionnement du récepteur)
ROE	Return on Equity (Rentabilité des capitaux propres)
SHAP	Shapley Additive exPlanations
SIZE	Taille de l'entreprise : Log (Total actif / Indice des prix)
SMOTE	Synthetic Minority Over-sampling Technique
SPA	Société Par Actions
TLTA	Total Liabilities / Total Assets (Total des dettes / Total actif)
UE	Union Européenne
UK	United Kingdom (Royaume-Uni)
WCTA	Working Capital / Total Assets (Fonds de roulement / Total actif)
XAI	Explainable Artificial Intelligence (Intelligence Artificielle Explicable)
ZLECAf	Zone de Libre-Échange Continentale Africaine

INTRODUCTION GENERALE

Contexte et problématique

Le commerce international se caractérise, de plus en plus, par une expansion considérable des échanges transfrontaliers, accompagnée d'une complexification croissante des risques auxquels sont exposés les opérateurs économiques. Dans ce contexte, le risque de non-paiement de l'acheteur étranger constitue l'une des menaces les plus redoutées par les entreprises exportatrices. En effet, lorsqu'un exportateur accorde un délai de paiement à son partenaire commercial étranger, il s'expose à une incertitude majeure sur la solvabilité de ce dernier, aggravée par la distance géographique, les différences institutionnelles et les asymétries d'information inhérentes aux transactions internationales (Agarwal, et al., 2023).

C'est précisément pour répondre à ce besoin de sécurisation des échanges que l'assurance-crédit à l'exportation a émergé comme instrument incontournable de la politique commerciale internationale. En permettant aux exportateurs de se prémunir contre les risques commerciaux et politiques liés au non-paiement, elle joue un rôle fondamental de facilitateur du commerce extérieur, tout en contribuant à la compétitivité des économies sur les marchés mondiaux (Kramar, 1979). En Algérie, cet instrument est engendré par la Compagnie Algérienne d'Assurance et de Garantie des Exportations (CAGEX), un établissement public chargé de soutenir et de sécuriser les exportations nationales hors hydrocarbures.

Traditionnellement, l'évaluation du risque de défaut des acheteurs étrangers reposait sur des méthodes d'analyse financière classiques : ratios de solvabilité, ratios de liquidité, taux de couverture des dettes : combinées à l'expertise des analystes crédit et à des modèles de scoring statistiques inspirés des travaux pionniers de (Beaver, 1966), (Altman, 1968) et (Ohlson, 1980). Si ces approches ont fait leurs preuves sur les marchés domestiques, elles font face à des limites structurelles lorsqu'elles sont appliquées à des acheteurs étrangers : incomplétude des informations disponibles, diversité des contextes institutionnels, déséquilibre des bases de données et incapacité à intégrer des variables macroéconomiques ou comportementales de manière dynamique (Yazdi, Hanne, & Wang, 2019).

Face à ces limites, les développements récents de l'intelligence artificielle et, plus spécifiquement, du Machine Learning (ML) ouvrent de nouvelles perspectives pour l'évaluation prédictive du risque de crédit. Des algorithmes tels que XGBoost ou LightGBM, issus de la famille des de gradient boosting, ont démontré des capacités prédictives supérieures aux modèles économétriques traditionnels dans de nombreux domaines

financiers (Noriega & Rivera, 2023). Leur capacité à traiter des bases de données volumineuses, à capturer des relations non linéaires entre les variables et à intégrer simultanément des indicateurs financiers, macroéconomiques et transactionnels en fait des outils particulièrement adaptés à la problématique de l'évaluation de l'acheteur étranger en assurance-crédit à l'exportation.

Toutefois, l'adoption de ces modèles dans un domaine aussi réglementé et sensible que l'assurance-crédit soulève des enjeux cruciaux d'explicabilité et de transparence. Le cadre réglementaire international notamment le Règlement Général sur la Protection des Données (RGPD), les recommandations de Bâle II et les lignes directrices de l'Autorité Bancaire Européenne (EBA) impose en effet que les décisions automatisées soient compréhensibles et auditables par les parties prenantes concernées (Bussmann, Giudici, Marinelli, & Papenbrock, 2020).

C'est dans cette articulation entre modélisation prédictive avancée, évaluation du risque de l'acheteur étranger et contraintes d'explicabilité que s'inscrit la présente recherche. Partant du constat qu'aucune étude empirique à ce jour n'a proposé une démarche intégrée combinant Machine Learning et exigences d'explicabilité dans le cadre spécifique de l'assurance-crédit à l'exportation. En Algérie, cette recherche entend combler ce vide scientifique. La problématique centrale qui guide notre travail est dès lors la suivante :

Dans quelle mesure la modélisation prédictive par Machine Learning, peut-elle améliorer l'évaluation du risque de défaut de l'acheteur étranger dans le cadre de l'assurance-crédit à l'exportation, et comment garantir l'explicabilité des modèles retenus dans un contexte réglementaire et opérationnel tel que celui de la CAGEX ?

De cette question centrale de recherche , découlent des questions subsidiaires qui se présentent comme suit :

Q1 : Les indicateurs financiers de l'acheteur étranger (ratios de solvabilité, liquidité, rentabilité) constituent-ils des prédictifs significatifs du risque de défaut dans un modèle de Machine Learning appliqué à l'assurance-crédit à l'exportation ?

Q2 : Les variables macroéconomiques et le risque pays exercent-ils une influence significative sur la probabilité de défaut de l'acheteur étranger, au-delà des indicateurs financiers individuels ?

Q3 : Les données transactionnelles et comportementales (historique de paiement, nature de l'opération) améliorent-elles le pouvoir prédictif des modèles retenus ?

Q4 : Parmi les algorithmes testés régression logistique, XGBoost et LightGBM ,lequel offre le meilleur compromis entre performance prédictive et explicabilité dans le contexte de la CAGEX ?

Afin de mieux cadrer ce travail, des réponses provisoires faciliteraient substantiellement le travail de recherche. Il s'agit des hypothèses suivantes :

H1 : Les indicateurs financiers influencent directement l'évaluation de défaut des acheteurs étrangers.

H2 : Les données macroéconomiques et le risque pays influencent significativement l'évaluation de défaut des acheteurs étrangers.

H3 : Les données transactionnelles contribuent significativement à la prédiction du défaut acheteur.

H4 : Les modèles de machine Learning, notamment XGBoost et LightGBM, présentent une meilleure performance prédictive que la régression logistique.

H5 : L'utilisation des méthodes d'explicabilité permet de mieux comprendre les facteurs déterminants du défaut acheteur.

Objectifs de la recherche

Cette recherche poursuit un double objectif, à la fois théorique et pratique.

Sur le plan théorique, elle vise à combler un vide scientifique en proposant une démarche intégrée qui articule modélisation prédictive par Machine Learning et exigences d'explicabilité dans le cadre spécifique de l'assurance-crédit à l'exportation. Plus précisément, elle cherche à évaluer comparativement les performances de trois algorithmes régression logistique, XGBoost et LightGBM appliqués à la prédiction du risque de défaut des acheteurs étrangers, et à identifier les variables les plus déterminantes dans la décision de couverture.

Sur le plan pratique, elle ambitionne de proposer à la CAGEX un cadre opérationnel d'aide à la décision fondé sur des modèles interprétables, susceptibles d'améliorer la qualité de l'évaluation des acheteurs étrangers tout en répondant aux contraintes de transparence et de traçabilité qui s'imposent aux institutions financières publiques.

Intérêt de la recherche

L'intérêt de cette recherche se situe à plusieurs niveaux complémentaires.

Sur le plan scientifique, elle contribue à enrichir la littérature sur l'évaluation du risque de crédit international en transposant des méthodologies issues du Machine Learning dans un secteur peu exploré : l'assurance-crédit à l'exportation. Elle répond ainsi à un besoin identifié dans la littérature, celui d'études empiriques combinant performance prédictive et explicabilité dans des contextes réglementaires spécifiques.

Sur le plan institutionnel et managérial, elle s'intéresse à une institution stratégique pour l'économie algérienne la CAGEX dont la mission de promotion des exportations hors hydrocarbures est au cœur des priorités de diversification économique nationale. Améliorer la qualité de l'évaluation des acheteurs étrangers pourrait directement contribuer à réduire les sinistres, optimiser la tarification des risques et renforcer la confiance des exportateurs algériens dans les marchés internationaux.

Sur le plan méthodologique, l'originalité de cette étude réside dans l'utilisation d'une base de données semi synthétique issue de la CAGEX, permettant de confronter les algorithmes de Machine Learning à un contexte opérationnel authentique, caractérisé par le déséquilibre des classes, la diversité des profils d'acheteurs et la complexité des variables en jeu.

Méthodologie adoptée

Pour répondre à notre problématique, nous avons adopté une approche hypothético-déductive s'appuyant sur une méthodologie quantitative. Notre démarche repose sur l'analyse d'une base de données semi synthétique, composée de dossiers d'acheteurs étrangers traités, contenant des informations financières, macroéconomiques et transactionnelles.

Trois algorithmes de classification ont été testés et comparés : la régression logistique, utilisée comme modèle de référence en raison de son interprétabilité reconnue, XGBoost et LightGBM, représentant des méthodes de gradient boosting reconnues pour leur haute performance prédictive. Le déséquilibre de la variable cible a été traité par la technique

SMOTE, et les modèles ont été évalués sur la base de métriques adaptées : AUC-ROC, F1-score, precision et recall, avec un ajustement du seuil de décision pour optimiser la détection des acheteurs à risque. L'explicabilité globale des modèles retenus a ensuite été analysée à travers l'importance des variables, afin de vérifier leur compatibilité avec les exigences réglementaires et opérationnelles de la CAGEX.

Structure du mémoire

Afin de répondre de manière systématique à notre problématique, ce mémoire est structuré en trois chapitres complémentaires.

Le premier chapitre est consacré au cadre théorique. Il présente une revue de littérature approfondie portant sur les fondements de l'assurance-crédit à l'exportation, les méthodes traditionnelles et avancées d'évaluation du risque des acheteurs étrangers, la contribution du Machine Learning à la modélisation prédictive, ainsi que les enjeux d'explicabilité et de conformité réglementaire. Il développe ensuite le cadre conceptuel qui structure notre modèle de recherche.

Le deuxième chapitre expose le cadre méthodologique et organisationnel de la recherche. Il détaille notre posture épistémologique, l'approche retenue, la construction et la préparation de la base de données, les choix méthodologiques en matière de modélisation, et présente la CAGEX en tant qu'institution d'accueil et terrain d'investigation.

Le troisième chapitre présente les résultats empiriques obtenus et en discute les implications. Il rend compte de l'analyse descriptive de la base de données, des performances comparatives des trois algorithmes testés, de l'interprétation des variables les plus importantes, et de la pertinence des modèles retenus au regard des contraintes opérationnelles et réglementaires de la CAGEX.

Le mémoire s'achève par une conclusion générale qui synthétise les principaux résultats enseignements de la recherche, répond aux hypothèses formulées, soulève les limites inhérentes à ce travail et ouvre des perspectives pour de futures recherches dans le domaine.

CHAPITRE I
CADRE THÉORIQUE

Introduction :

Ce chapitre représente une étape primordiale pour répondre à la question : comment le Machine Learning peut évaluer l'acheteur étranger dans l'opération de l'assurance-crédit. En établissant les fondements théoriques de notre réflexion, il permet de contextualiser notre recherche dans son cadre scientifique et de justifier ainsi la pertinence de notre réflexion. Il est divisé en deux parties. La première porte sur la revue de littérature, où nous examinerons les recherches antérieures liées à notre sujet. La seconde est consacrée au cadre conceptuel, dans laquelle nous définirons les idées fondamentales, présenterons le modèle de recherche et énoncerons les hypothèses à vérifier ultérieurement dans la partie empirique.

I. Revue de littérature :

Pour les exportateurs, le risque de non-paiement de l'acheteur étranger est une problématique centrale dans le commerce international. L'assurance-crédit à l'exportation joue donc un rôle crucial, car elle permet de sécuriser les opérations caractérisées par l'incertitude sur la solvabilité du débiteur, le risque pays et les conditions économiques mondiales. L'évaluation du risque, jusqu'à présent principalement fondée sur des méthodes traditionnelles d'analyse financière et sur l'expertise, tend désormais à s'orienter vers des approches plus sophistiquées, comme l'apprentissage automatique. Il faut néanmoins concilier la quête de performance avec les exigences d'explicabilité et de transparence dans un domaine aussi sensible. Le but de cette revue de littérature est d'analyser les fondements de l'assurance-crédit à l'exportation, les modèles d'évaluation du risque commercial des acheteurs étrangers, la contribution de la modélisation prédictive ainsi que les enjeux associés aux approches explicables.

1. Fondements théoriques et institutionnels de l'assurance-crédit à l'export :

La naissance de l'assurance-crédit à l'export est due au problème central du commerce international : l'incertitude sur la solvabilité de l'acheteur étranger et le risque de non-paiement.

(Petrová, Krügerová, & Kozieł, 2020) rappellent que l'identification des risques au sein de l'entreprise est une condition essentielle d'un management efficace des risques. Précisément, le risque territorial peut être décomposé en : risque politique, risque lié à la politique commerciale, ainsi que le risque monétaire et de change, ces risques étant fortement liés les uns aux autres. C'est pourquoi leur gestion doit être pensée de façon globale. Le risque

politique renvoie à la guerre, aux troubles internes, à l'expropriation ou à la confiscation des biens, des événements susceptibles de minimiser ou d'interrompre les relations économiques, d'empêcher la circulation des marchandises ou de rendre impossible le recouvrement des créances. À cela s'ajoute le risque de politique commerciale, qui prend la forme d'embargos à l'importation, de droits de douane, de taxes, de restrictions sanitaires ou de réglementations techniques. Ces éléments montrent que l'évaluation du risque à l'export ne peut jamais être réduite à une simple lecture financière de l'acheteur : elle doit aussi intégrer la situation du pays destinataire. C'est pour cette raison que les compagnies d'assurance-crédit internationales élaborent généralement leurs propres cartes de risque pays, afin de proposer des solutions de couverture adaptées aux spécificités de chaque marché.

Selon (Agarwal, et al., 2023), les frictions informationnelles constituent un élément central du commerce international car elles reflètent les problèmes auxquels fait face un exportateur qui ne dispose pas d'une information complète, fiable et vérifiable sur son acheteur étranger. Pratiquement, l'entreprise qui exporte ignore souvent si son client dispose réellement des fonds pour payer, si son comportement est conforme à ses engagements, ou si la situation économique et institutionnelle de son pays risque de perturber le bon déroulement de la transaction. Par ailleurs, l'acheteur peut obtenir la marchandise sans verser le prix convenu, ce qui diminue directement les profits attendus par l'exportateur et peut le dissuader d'entrer sur certains marchés étrangers. Les frictions informationnelles sont donc un problème majeur : elles constituent un véritable obstacle au commerce extérieur, au même titre que les coûts de transit ou que les barrières tarifaires.

Dans ce contexte, recourir à un financement du commerce ou aux garanties à l'exportation s'avère indispensable. En théorie, lorsque les acteurs privés (banques, institutions financières privées) jugent une opération internationale trop risquée, les garanties publiques à l'exportation prennent le relais pour sécuriser l'exportateur. Ces dernières permettent de couvrir le risque de non-paiement et de réduire les difficultés liées à l'information, renforçant ainsi la sécurité des opérations à l'export et soutenant le commerce extérieur. Dans plusieurs nations, ce soutien aux exportations repose sur des établissements publics d'assurance-crédit, l'assurance-crédit à l'export constituant l'un des instruments les plus importants de ce dispositif, ce qui explique que les pays fortement tournés vers l'exportation disposent en général d'agences de crédit à l'exportation soutenues par l'État. Ces études montrent par ailleurs que les organismes publics de garantie à l'exportation peuvent avoir un effet positif sur la dynamique commerciale, en mettant en évidence une relation positive entre la

disponibilité de l'assurance-crédit publique et le niveau des exportations : ces organismes ne jouent donc pas seulement un rôle de protection, ils participent aussi à la stimulation effective du commerce international.

2. Modèles traditionnels et hybrides d'évaluation du risque des acheteurs étrangers :

(Yazdi, Hanne, & Wang, 2019) montrent que l'évaluation du risque repose, dans un premier temps, sur des méthodes traditionnelles. Les agences de crédit à l'export (ECA) évaluent les acheteurs étrangers en combinant des indicateurs financiers classiques tels que le ratio dette/capitaux propres, le ratio dette/EBITDA, la couverture des intérêts, la rentabilité, la liquidité et la rotation de l'actif, avec des critères non financiers comme la structure de management, la gouvernance, l'historique de paiement, la transparence du reporting, la position concurrentielle de l'entreprise et le risque-pays. Toutefois, les auteurs soulignent les limites de cette approche lorsqu'elle est appliquée seule : dans le domaine de l'assurance-crédit à l'export, de nombreuses variables sont qualitatives, imprécises ou exprimées de manière linguistique, ce qui rend les décisions difficiles en s'appuyant uniquement sur des méthodes quantitatives strictes. En réponse à ces limites, ils proposent un modèle hybride fondé sur un système d'inférence floue (Fuzzy Inference System), conçu pour traiter les environnements incertains. Ce modèle marque ainsi une évolution de l'évaluation traditionnelle, centrée sur les ratios et le jugement analytique, vers une approche plus avancée qui intègre la logique floue afin de mieux apprécier la probabilité de défaut des acheteurs étrangers et d'appuyer la décision de recommandation de crédit.

(Petrová, Krügerová, & Kozieł, 2020) S'inscrivent dans la continuité de cette évolution en abordant la dimension macroéconomique et institutionnelle de l'évaluation du risque. Ils expliquent que l'OCDE recourt au modèle Country Risk Assessment Model (CRAM) pour évaluer le risque-pays à partir de données macroéconomiques et financières. Selon leurs travaux, l'évaluation finale intègre plusieurs composantes complémentaires : le risque-pays, la qualité de crédit du débiteur, la durée du risque, le niveau de couverture ainsi que les techniques d'atténuation du risque.

3. Modélisation prédictive par apprentissage automatique en assurance-crédit à l'export :

(Bärtland & Krummaker, 2020) Montrent que les méthodes avancées de machine learning peuvent apporter une véritable valeur ajoutée dans un secteur où les sinistres sont peu

fréquents mais parfois lourds lorsqu'ils surviennent. Les auteurs comparent plusieurs approches d'apprentissage automatique : arbres de décision, forêts aléatoires, réseaux de neurones et réseaux de neurones probabilistes, sur la base de données de Burns & Wilcox, afin d'évaluer leur aptitude à prédire la survenue et la gravité des sinistres. Les résultats indiquent que ces modèles parviennent de manière satisfaisante à déterminer si un sinistre va se produire, et à en estimer l'ordre de grandeur, bien que la prédiction précise du ratio de sinistre demeure difficile. Parmi les méthodes testées, la Random Forest se distingue comme la plus performante et la plus stable. Cette étude suggère ainsi que le machine learning ne se substitue pas à l'analyse traditionnelle, mais la complète en apportant une capacité de prévision plus fine dans un environnement marqué par l'incertitude.

(Uddin, Chi, & Al Janabi, 2022) soulignent que les modèles hybrides constituent une piste prometteuse pour améliorer l'évaluation du risque de crédit. Les auteurs rappellent que les approches classiques, notamment la régression logistique, conservent une place importante grâce à leur transparence et à leur large reconnaissance dans la pratique. Ils montrent néanmoins que des méthodes plus récentes, telles que la Random Forest et le Gradient Boosting, offrent souvent une meilleure capacité prédictive, et que leur combinaison permet d'obtenir des résultats plus robustes. Cette hybridation améliore surtout la qualité de la classification sur des jeux de données de faible dimension, tandis que les gains apparaissent plus limités lorsque les données sont plus complexes ou de grande dimension. Les auteurs en tirent une conclusion importante : la performance d'un modèle ne dépend pas uniquement de sa sophistication technique, mais aussi de son adéquation avec la structure des données disponibles.

(Noriega & Rivera, 2023) Mettent en évidence que les modèles de type *Boosted Ensemble*, tels que XGBoost et LightGBM, figurent parmi les plus étudiés et les plus performants, surpassant plusieurs approches plus classiques. Ils précisent que l'évaluation de ces modèles repose généralement sur des indicateurs tels que l'AUC, l'exactitude (accuracy), le rappel (recall), la précision et le score F1. Cependant, cette montée en performance s'accompagne de plusieurs limites importantes : le déséquilibre des données, la difficulté de sélectionner les variables pertinentes, et surtout le manque d'explicabilité de certains modèles complexes. Appliqués au domaine de l'assurance-crédit à l'export, les auteurs suggèrent que le machine learning peut améliorer la qualité de la prédiction du risque, à condition de construire une base de données solide, de sélectionner les variables adéquates et d'assurer une interprétation rigoureuse des résultats.

(Liao, Jiao, & Zhang, 2025) Prolongent cette réflexion en proposant un cadre fondé sur LightGBM, enrichi par les méthodes d'explicabilité SHAP (Shapley Additive Explanations) et LIME (Local Interpretable Model-agnostic Explanations), afin d'identifier les variables les plus déterminantes et d'expliquer les décisions du modèle à la fois de manière globale et locale. Les auteurs soulignent également que l'évaluation du risque ne doit pas viser uniquement la meilleure précision statistique, mais doit aussi tenir compte du coût asymétrique des erreurs de classification — notamment le cas où une entreprise à haut risque serait incorrectement classée comme peu risquée, ce qui représente une erreur aux conséquences particulièrement graves dans le contexte de l'assurance-crédit.

4. Approches explicables (XAI) et considérations réglementaires :

Pour Ethics guidelines for trustworthy AI du groupe d'experts de la Commission européenne. Les principes évoqués qui sont directement liés à l'XAI : la supervision humaine, la transparence et la responsabilité. Cela signifie que les décisions doivent rester sous contrôle humain, que le fonctionnement de l'IA doit pouvoir être expliqué selon le type d'acteur concerné, et que les systèmes doivent être auditables et attribuables en termes de responsabilité. (European Commission, 2019)

(Bussmann, Giudici, Marinelli, & Papenbrock, 2020) Commencent par une idée simple : les modèles d'intelligence artificielle de type « boîte noire » ne sont pas appropriés aux services financiers réglementés. L'explication de l'intelligence artificielle devient nécessaire dès qu'une décision automatisée doit être comprise par les personnes concernées, en particulier dans la gestion du risque de crédit. L'explicabilité est donc la capacité de rendre clairs les facteurs qui expliquent une décision produite par un modèle. Il mentionne notamment le Financial Stability Board, qui considère que le manque d'interprétabilité et d'auditabilité des méthodes d'IA et de ML peut devenir un risque à l'échelle macro-financière. En rappelant aussi que dans certains cas, le droit lui-même impose un certain degré d'explicabilité.

(Bussmann, Giudici, Marinelli, & Papenbrock, 2020) Accordent une place importante au RGPD (Règlement Général sur la Protection des Données), selon cette réglementation européenne, lorsqu'une décision automatisée est utilisée, la personne concernée doit pouvoir recevoir des informations utiles sur la logique du traitement, sur sa portée et ses conséquences prévues. Cela veut dire que l'explicabilité n'est pas seulement une exigence technique ou éthique, elle devient aussi une exigence juridique.

Selon (de Lange, Melsom, Vennerød, & Westgaard, 2022) les modèles de scoring bancaire doivent être explicables. En effet, les autorités financières exigent que les décisions prises à partir des données soient compréhensibles et justifiables. Les modèles d'IA avancés sont souvent considérés comme des boîtes noires, il est difficile de comprendre comment les systèmes d'intelligence artificielle fonctionnent et comment ils parviennent à leurs prédictions. Le cadre réglementaire européen, notamment Bâle II et le Règlement Général sur la Protection des Données RGPD, impose aux banques européennes une transparence dans leurs modèles de décision automatisée. Ces banques doivent donc appliquer certaines règles pour expliciter leur manière de prendre leurs décisions. L'EBA, ou Autorité bancaire européenne, a relevé trois problèmes majeurs posés par les modèles d'intelligence artificielle complexes. D'abord, il est difficile d'expliquer les résultats obtenus par ces modèles. Ensuite, les équipes de direction ont du mal à bien comprendre le fonctionnement de ces modèles. Enfin, il est difficile de justifier les résultats devant les organismes de contrôle. On finalise par L'Union européenne est en train de créer des règles encore plus strictes pour les systèmes d'intelligence artificielle.

On conclure par l'étude de (Černevičienė & Kabašinskas, 2024) qui montre d'abord que dans la finance, l'XAI devient nécessaire parce que les modèles d'IA sont complexes et donc plus difficiles à interpréter précisément les modèles de type boîte noire, alors dans des domaines sensibles comme le crédit, la fraude, la faillite ou l'optimisation de portefeuille, une forte précision ne suffit pas à elle seule, il faut aussi comprendre comment la décision est produite. Ils rappellent également que, dans la finance, l'XAI est liée à des enjeux de confiance, responsabilité, contrôle humain et conformité.

L'XAI en finance n'est pas limité à une seule méthode. Il existe plusieurs approches fréquentes sont l'importance des variables, SHAP, les méthodes fondées sur des règles, ainsi que des Framework combinés qui utilisent plusieurs techniques à la fois. Les domaines où l'XAI est plus utilisé sont surtout le crédit management, la prédiction des prix boursiers et la détection de fraude.

En conclusion, cette revue montre que l'assurance-crédit à l'export se fonde sur des bases théoriques et institutionnelles solides, et que l'évaluation du risque des acheteurs étrangers a progressivement évolué des approches traditionnelles vers des modèles plus avancés, notamment ceux fondés sur la machine Learning. Cependant, même si plusieurs recherches ont étudié le risque de crédit, les modèles prédictifs ou encore l'explicabilité des décisions

automatisées, peu d'études ont réellement réuni ces dimensions dans le cadre spécifique de l'assurance-crédit à l'export. Cette limite représente l'intérêt de notre mémoire. En cherchant à mieux comprendre comment le machine Learning peut contribuer à l'évaluation des acheteurs étrangers, tout en respectant les exigences de transparence et de compréhension des résultats, notre travail ambitionne d'ouvrir une réflexion utile aussi bien pour la recherche que pour les pratiques professionnelles dans ce domaine.

II. LE CADRE CONCEPTUEL

Ce cadre conceptuel constitue la deuxième partie de ce chapitre, car il permet de structurer les principaux concepts liés à notre sujet de recherche. Après avoir présenté, dans la revue de littérature, les fondements de l'assurance-crédit à l'export, les méthodes d'évaluation du risque de l'acheteur étranger, l'apport du machine Learning et les exigences d'explicabilité, il devient nécessaire d'organiser ces éléments dans une logique cohérente. Il vise ainsi à définir les notions essentielles de l'étude, à préciser les relations entre elles et à préparer la construction du modèle de recherche ainsi que des hypothèses.

1. Délimitation théorique de l'assurance-crédit à l'export

Avant d'analyser les modèles prédictifs appliqués à l'évaluation du risque de défaut des acheteurs étrangers, il convient de poser les bases conceptuelles de l'instrument au cœur de cette étude, en définissant ce qu'est l'assurance-crédit à l'export, en retraçant les conditions historiques de son émergence et en comprenant pourquoi elle s'est progressivement imposée comme un dispositif incontournable dans la sécurisation des transactions commerciales internationales.

1.1. Définition et origine :

L'assurance-crédit à l'export est un instrument de protection qui permet à une entreprise exportatrice de se protéger contre le risque de non-paiement de ses créances commerciales par un acheteur étranger. Elle intervient surtout lorsque l'exportateur accorde un délai de paiement à son client, c'est-à-dire lorsqu'il accepte de livrer les biens ou les services avant d'être payé. Dans ce cas, l'entreprise supporte un risque important : l'acheteur peut devenir insolvable, retarder le paiement ou ne pas régler sa dette pour des raisons commerciales, politiques ou monétaires. Selon (Colombier, 1958), cette assurance est devenue nécessaire car le vendeur, ou bien sa banque, peut difficilement apprécier la solvabilité d'un client situé à l'étranger, et parce que le recouvrement d'une créance internationale est plus complexe que sur le marché national.

L'origine de l'assurance-crédit s'explique principalement par l'évolution du crédit commercial. (Tarde, 1929) Montre que cette forme d'assurance est née de deux mouvements complémentaires : le premier est la progression du crédit dans l'espace, dans le temps et en volume ; le deuxième est la diffusion progressive de l'idée d'assurance appliquée aux risques commerciaux. Autrement dit, plus les entreprises vendaient à crédit, plus elles avaient besoin d'un outil permettant de limiter les pertes en cas de défaut de paiement. (Tarde, 1929) Parle ainsi d'une double origine : une origine bancaire, car les banques ne pouvaient pas supporter seules tous les risques liés au crédit, et une origine assurantielle, car l'assurance permettait de répartir les risques grâce à la mutualisation et à la réassurance. Cette origine devient encore plus importante dans le cas des exportations. Les opérations internationales sont plus incertaines que les ventes domestiques, car elles impliquent des acheteurs éloignés, des systèmes juridiques différents, des délais plus longs et parfois des risques politiques. C'est pourquoi (Tarde, 1929) souligne que le besoin de crédit et d'assurance s'est manifesté très tôt dans le domaine de l'exportation, notamment après la Première Guerre mondiale, ce qui a favorisé l'apparition d'une aide gouvernementale à l'exportation. (Kramar, 1979) Précise également que l'assurance-crédit à l'exportation s'est développée à la fin du XIX^e siècle au Royaume-Uni, avant de prendre plus d'importance après la Première Guerre mondiale, durant les années 1930, puis avec la reprise économique de l'Europe occidentale et du Japon dans les années 1950.

Initialement, l'assurance des risques à l'exportation relevait surtout de l'initiative privée. Cependant, les limites de l'assurance privée sont rapidement apparues, notamment lorsque les risques dépassaient le simple risque commercial. (Colombier, 1958) Explique que les assureurs privés pouvaient couvrir certains risques ordinaires, comme l'insolvabilité de l'acheteur privé sur des opérations limitées dans leur montant et leur durée. Par contre, les risques politiques, catastrophiques, monétaires ou liés aux acheteurs publics nécessitaient souvent l'intervention de l'État, car ils étaient plus difficiles à évaluer et à mutualiser. Cette évolution explique pourquoi l'assurance-crédit à l'export est progressivement devenue un mécanisme lié aux politiques publiques de soutien aux exportations.

(Kramar, 1979) Distingue que sur le plan pratique, l'assurance-crédit à l'export couvre deux grandes catégories de risques. La première correspond aux **risques commerciaux**, tel que l'insolvabilité de l'acheteur, le défaut prolongé de paiement ou bien le refus injustifié d'accepter les marchandises. La seconde correspond aux **risques non commerciaux**, souvent appelés risques politiques : guerre, troubles, catastrophes naturelles, non-transfert

des devises, annulation d'une licence d'importation, nouvelles restrictions administratives ou encore risque de change. (Berne Union, n.d.)Présente aussi l'assurance-crédit comme un outil permettant de soulager la trésorerie de l'exportateur lorsque ses clients deviennent insolvable ou ne paient pas leurs factures à cause de risques commerciaux ou politiques.

Ainsi, l'assurance-crédit à l'export peut être définie comme un instrument de gestion du risque qui protège l'exportateur contre le non-paiement de ses créances étrangères, tout en facilitant le financement, la confiance commerciale et le développement des exportations. Elle occupe donc une place particulière entre l'assurance, la banque et la politique commerciale. Dans cette logique, (OECD Export credits, 2026) précise que les crédits à l'exportation officiellement soutenus sont fournis par des agences de crédit à l'exportation, appelées ECA, afin d'aider les exportateurs nationaux à concurrencer les entreprises étrangères sur les marchés internationaux. Ce soutien peut prendre la forme d'une assurance, d'une garantie ou d'un financement.

En récapitulant, l'assurance-crédit à l'export est née du besoin de sécuriser les ventes à crédit dans un contexte international plein de l'incertitude. Son développement historique montre qu'elle répond à un double objectif : protéger l'entreprise exportatrice contre le non-paiement et soutenir le commerce extérieur. Alors elle permet à l'exportateur de vendre à l'étranger toute en sécurité, d'accorder des délais de paiement plus attractifs et de préserver sa trésorerie en cas de défaillance de l'acheteur.

Figure 1: Schéma représentatif du mécanisme de l'assurance-crédit à l'exportation



Source : Site officiel de Berne Union (Berne Union, n.d.)

1.2. Les risques couverts en assurance-crédit à l'export :

L'assurance-crédit à l'export ne couvre pas uniquement le risque financier lié à l'acheteur étranger. Elle vise plus largement à protéger l'exportateur, ou parfois la banque qui finance l'opération, contre plusieurs formes de risques pouvant empêcher le paiement d'une créance internationale. La somme des risques couverts est :

a) Le risque commercial : (Allianz Trade , 2023)

Le risque commercial est un terme vaste, qui englobe tous les éléments et facteurs susceptibles de réduire les bénéfices ou de menacer l'activité d'une entreprise. Les événements et circonstances qui constituent un risque commercial peuvent être intérieurs ou extérieurs à l'entreprise. Il regroupe les risques suivants :

- Risque Economique : tout changement au sein de l'économie dont les conséquences peuvent affecter les ventes, les revenus ou les bénéfices d'une entreprise.
- Risque lié a la conformité : concerne particulièrement les industries et les secteurs fortement réglementés. En effet, toute entreprise se doit de respecter les lois et les règlements. Ne pas les respecter expose à des amendes, voire des sanctions pouvant aller jusqu'à l'arrêt de l'activité.
- Risque financier : désigne tout élément interne ou externe qui, lorsqu'il survient, affecte directement le résultat financier d'une entreprise.
- Risque de réputation : La réputation d'une entreprise est un élément clé. Si celle-ci atteinte, les répercussions peuvent être terribles, ce risque est particulièrement élevé notamment à cause des réseaux sociaux qui permettent à tout le monde d'exprimer son opinion et de la relayer au plus grand nombre.
- Risque opérationnel : Il s'agit d'un risque interne, il peut être la conséquence d'une inadéquation ou de la défaillance.

b) Le risque politique : (COFACE FOR TRADE , 2024)

Le risque politique peut être d'ordre politique, social, sécuritaire ou climatique. Il peut aussi s'agir d'une décision politique qui empêche le bon déroulement des opérations de commerce tel que : blocage d'un transfert d'avoir ou la confiscation d'investissements ou de biens. Il est majoré par ces 4 risques :

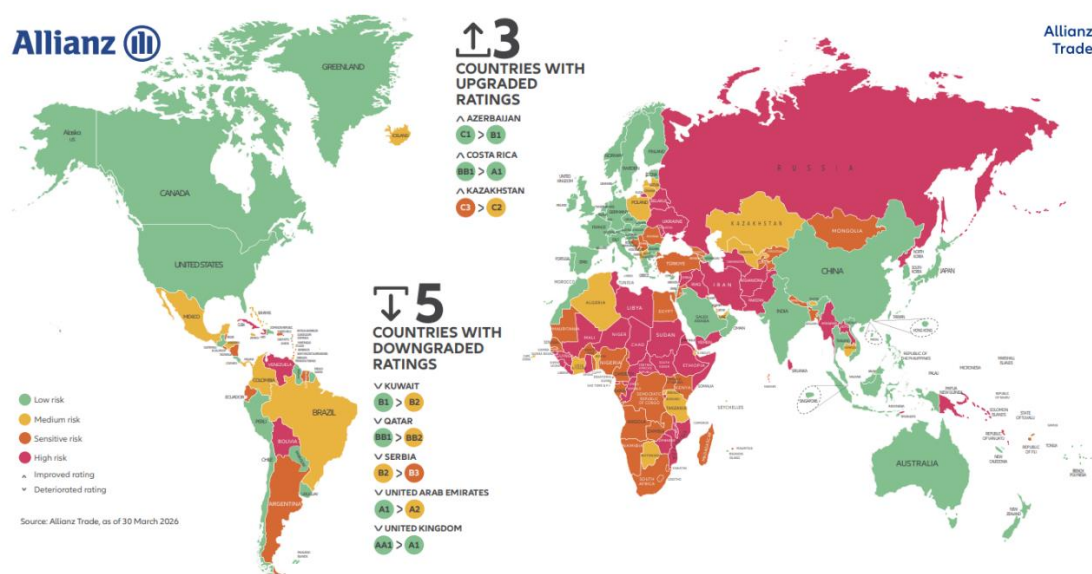
- Ingérence gouvernementale : En cas de crise politique ou de changement de gouvernement, le pouvoir en place peut décider de confisquer les investissements des entreprises, ou elle peut perdre ses actifs, ses investissements.
- Guerres, Emeutes, et Attaques terroristes : les activités d'une entreprise peuvent être entravées de plusieurs manières comme la Destruction ou fermeture des locaux et endommagement ou perte d'actifs. Cet incident engendre de sérieux problèmes de

trésorerie pour les entreprises, avec le risque de se retrouver en situation de défaut de paiement vis-à-vis de leurs partenaires commerciaux.

- Embargos et sanctions : Lorsqu'une entreprise exerce une activité d'importation ou d'exportation avec un pays sous embargo ou sanctions, ses échanges commerciaux pourraient être rapidement freinés et entraînant des pertes importantes.
- Risque climatique : les entreprises sont principalement vulnérables aux risques physiques des événements climatiques et au risque de transition lié aux décisions prises par les gouvernements.

Alors le risque politique est une notion plus englobante qui intègre des dimensions macroéconomiques, institutionnelles et géopolitiques comme le niveau d'endettement extérieur, instabilité de la balance des paiements, faiblesse des réserves de change, instabilité politique structurelle, qualité des institutions juridiques et de la gouvernance, selon (OECD, n.d.) produit annuellement une classification des risques pays en huit catégories (0 à 7), qui sert de référence internationale pour les assureurs de crédit à l'exportation.

Figure 2: Représente une Cartographie de Classification des pays



Source : Site Officiel de Allianz Trade (AllianzTrade, 2026)

1.3. Les acteurs institutionnels de l'assurance-crédit à l'export :

L'assurance-crédit à l'export s'appuie sur un écosystème institutionnel constitué d'acteurs qui interviennent à différents niveaux : couvre le risque, accompagne les exportateurs et régule le marché. Donc elle ne peut pas être comprise uniquement comme un contrat entre

un exportateur et un assureur. Il s'agit alors d'un dispositif organisé autour des acteurs publics, privés, bancaires et multilatéraux, dont un objectif commun est de sécuriser les transactions internationales et de faciliter l'accès au financement du commerce extérieur.

D'abord, les **organismes publics**, appelés **Export Credit Agencies** ou **ECA**, occupent une place centrale dans cet écosystème. Selon (OECD, n.d.), les crédits à l'exportation officiellement soutenus sont encadrés par l'Arrangement sur les crédits à l'exportation bénéficiant d'un soutien public, qui vise à organiser les conditions de concurrence entre pays exportateurs. Ces agences peuvent intervenir directement ou avec l'appui de l'État afin de couvrir les risques liés aux exportations, notamment lorsque les risques commerciaux ou politiques sont jugés trop élevés pour être assumés uniquement par le marché privé. Ces agences publiques font plusieurs fonctions. Elles peuvent accorder des garanties, proposer une assurance-crédit, soutenir le financement bancaire ou encore couvrir des risques politiques et commerciaux liés au non-paiement de l'acheteur étranger, par exemple (UK Export Finance, 2025) offre des garanties, des assurances et des solutions de financement afin que l'exportation britannique soit viable ne bloque par manque de financement ou d'assurance. D'autre côté (Bpifrance, n.d.) se présente comme l'agence française de crédit à l'exportation et propose des solutions d'assurance avec la garantie directe de l'État français.

Aux États-Unis, (EXIM BANK, n.d.) constitue l'agence officielle de crédit à l'exportation. Elle soutient les exportations américaines à travers plusieurs instruments, notamment l'assurance-crédit export, les garanties de fonds de roulement, les garanties de prêts et les prêts directs.

Ces Agences montrent que l'assurance ce n'est pas son rôle unique, elles participent aussi au financement du commerce extérieur et à la compétitivité des entreprises nationales sur les marchés internationaux.

Ensuite, les **assureurs privés** ont également une place importante dans le marché de l'assurance-crédit. Des compagnies comme **Coface**, **Allianz Trade** ou **Atradius** proposent des solutions de couverture contre le risque de non-paiement des acheteurs, que ce soit sur le marché domestique ou à l'export. Leur rôle ne se limite pas à indemniser l'entreprise en cas d'impayé ; ils fournissent aussi des informations commerciales, des notations, une surveillance du risque client et des limites de crédit recommandées. Ainsi, ils contribuent directement à l'évaluation de l'acheteur étranger et à la prévention du risque avant même la survenance du sinistre. (Atradius, n.d.)

À côté des ECA et des assureurs privés, il se trouve aussi des **institutions multilatérales d'appui au commerce et à l'investissement** tel que, la **MIGA (Multilateral Investment Guarantee Agency)** (World Bank Group MIGA , n.d.) membre du Groupe de la Banque mondiale, offre des garanties contre les risques politiques pour encourager les investissements dans les pays en développement. L'**IFC (International Finance Corporation, n.d.)**, par son Global Trade Finance Program, propose des garanties partielles ou totales aux banques pour couvrir certains risques de paiement dans les marchés émergents. Ces institutions renforcent la sécurité des opérations internationales, mais elles ne remplacent pas les assureurs-crédit, surtout dans les pays où les risques économiques, politiques ou bancaires sont plus élevés.

Ces institutions financières forment un système complémentaire chacun a sa mission, les agences publiques ont pour le but de soutenir les politiques de l'exportation des Etats et couvrir principalement les risques politiques. Les assureurs privés basent principalement et par expertise en analyse de risque client, et on finalise par les institutions multilatérales renforcent la sécurité des transactions dans un environnement plus risqué.

2. Le concept d'acheteur étranger et son évaluation dans l'assurance-crédit :

Avant d'examiner les méthodes et les dimensions qui structurent son évaluation, il convient de définir avec précision ce que recouvre la notion d'acheteur étranger dans le cadre de l'assurance-crédit à l'export, en tant qu'entité économique dont le comportement de paiement constitue la source principale du risque que l'assureur cherche à anticiper et à mesurer.

2.1. Définition de l'acheteur étranger comme unité d'analyse :

Dans le cadre de l'assurance-crédit à l'exportation, l'acheteur étranger est le client situé dans un autre pays que le fournisseur, d'où l'entreprise vend des biens et services avec une date d'échéance. Alors il est le débiteur commercial international. Le risque principal pour cet acheteur est le non-paiement de la facture. (Yang, Bakshi, & Chen, 2020) Définit credit Buyer comme suit : c'est le débiteur qui bénéficie d'un crédit commercial accordé par le fournisseur grâce à l'assurance-crédit commerciale elle est un outil de gestion du risque utilisé par les fournisseurs pour se protéger contre le défaut de paiement des acheteurs bénéficiant d'un crédit commercial. Cette idée montre que l'acheteur n'est pas seulement un client, mais aussi une source potentielle de risque financier pour l'entreprise vendeuse.

Par contre (Wang, Zhao, & Peng, 2017) trouvent que cet acheteur est une unité d'analyse centrale pour l'assureur-crédit. Avant d'accorder une couverture, l'assureur doit évaluer sa

capacité à payer, sa stabilité financière et son comportement de paiement. Montrant que le crédit commercial crée un risque de défaut pour le fournisseur, surtout lorsque l'acheteur possède davantage d'informations sur sa propre situation financière que le vendeur.

Sur le plan pratique, l'acheteur étranger est donc une entreprise ou une organisation importatrice qui présente un risque de contrepartie pour l'exportateur. Ce risque correspond à la possibilité que l'acheteur ne respecte pas son obligation de paiement. Dans les modèles classiques de gestion du risque de crédit, cette logique rejoint l'évaluation de la solvabilité du débiteur, c'est-à-dire sa capacité financière et comportementale à rembourser ses dettes. Le Comité de Bâle rappelle que les systèmes de notation du risque de crédit doivent notamment définir le niveau de risque de défaut associé aux emprunteurs classés dans chaque catégorie de risque. (Basel Committee on Banking, 2022)

2.2. Les dimensions d'évaluation de l'acheteur étranger :

L'évaluation de l'acheteur étranger en assurance-crédit à l'export ne se limite jamais à une seule lecture financière. Elle repose sur une analyse multidimensionnelle, car le risque de non-paiement dépend à la fois de la situation interne de l'acheteur, de son comportement commercial, de la relation avec l'exportateur et de son environnement économique et politique. Ce principe rejoint les pratiques des assureurs-crédit, qui évaluent le risque acheteur avant la vente, surveillent l'évolution du comportement du client et utilisent cette analyse pour fixer des limites de crédit. L'assurance-crédit sert aussi à évaluer le risque acheteur, suivre les changements de comportement des clients et aider à décider des limites de crédit. (Atradius, 2025)

a) La dimension financière

La première dimension d'évaluation est la **dimension financière**. Elle consiste à analyser la capacité de l'acheteur étranger à honorer ses engagements de paiement à partir de ses états financiers. Dans cette perspective, l'assureur-crédit s'intéresse principalement à la solvabilité, à la liquidité, à la rentabilité, à la structure d'endettement et à la capacité de remboursement. Ces éléments permettent de mesurer si l'entreprise acheteuse dispose de ressources suffisantes pour payer ses fournisseurs à l'échéance. (Beaver, 1966) a montré que certains ratios comptables permettent de distinguer assez tôt les entreprises saines des entreprises proches de la faillite. Il a testé plusieurs ratios, puis a retenu surtout six (06) ratios représentatifs pour prédire la défaillance financière.

Tableau 1: Représente les principaux ratios sélectionnés par Beaver

Ratio	Formule	Interprétabilité
Couverture de la dette par les flux de trésorerie	Cash-flow/Dette totale	Mesure la capacité de l'entreprise à rembourser ses dette grâce aux flux générés par son activité.
Rentabilité économique	Résultat net/Total actif	Mesure la rentabilité économique de l'entreprise. Une rentabilité faible ou négative peut annoncer des difficultés.
D'Endettement	Dette totale / Total actif	Mesure le niveau d'endettement. Plus ce ratio est élevé, plus le risque financier augmente.
De Fonds de roulement sur Total actif	Fonds de roulement/ Total actif	Évalue la marge de sécurité financière à court terme. Un fonds de roulement faible indique une fragilité de liquidité.
Ratio de liquidité générale	Actif circulant/Passif circulant	Mesure la capacité à payer les dettes à court terme avec les actifs à court terme.
Intervalle sans crédit	Actif liquide/ Dépenses journalières moyennes	Indique combien de jours l'entreprise peut continuer à fonctionner sans accès à un nouveau crédit.

Source : élaboré par nous-mêmes à partir de (Beaver, 1966)

Après Beaver, (Altman, 1968) a proposé le modèle **Z-score**, fondé sur l'**analyse discriminante multiple** (technique statistique qui vise à décrire, expliquer et prédire l'appartenance à des groupes prédéfinis d'un ensemble d'observations à partir d'une série de variables prédictives). Contrairement à Beaver, qui étudiait surtout les ratios séparément, Altman combine plusieurs ratios dans une seule formule pour mieux prédire la faillite.

Tableau 2 : Représente Les ratios choisis par Altman

Variabes	Ratio	Signification
X1	Fonds de roulement/ total actif	Mesure la liquidité et la capacité de l'entreprise à couvrir ses besoins à court terme
X2	Bénéfices non distribués/ Total actif	Mesure l'accumulation des profits passés et la solidité financière historique
X3	Résultat avant impôt/ Total actif	Mesure la rentabilité économique de l'entreprise
X4	Valeur du marché des Capitaux propres/Valeur comptables des dettes	Mesure la solvabilité et capacité de l'entreprise à absorber ses dettes
X5	Chiffre d'affaires/Total actif	Mesure l'efficacité de l'utilisation des actifs (Rotation des actifs)

Source : élaboré par nous-mêmes à partir de (Altman, 1968)

D'où vient la formule suivante pour calculer le **Z-score** :

$$Z = 1,2X1 + 1,4X2 + 3,3X3 + 0,6X4 + 1,0X5$$

Il a mis l'interprétation suivante :

- **Z > 2,99** : Zone saine avec un faible risque
- **1,81 < Z < 2,99** : Zone grise : situation incertaine
- **Z < 1,81** : Zone de détresse : risque élevé de faillite

On finalise par (Ohlson, 1980) a proposé une approche différente de celle d'Altman. Au lieu d'utiliser l'analyse discriminante, il utilise un **modèle logit**, c'est une approche probabiliste qui estime la **probabilité de faillite** d'une entreprise à partir de variables comptables et financières.

Tableau 3: Représente Les indicateurs choisis par Ohlson

Variable	Indicateur	Signification
Taille d'entreprise (SIZE)	Log (Total actif/Indice des prix)	Les grandes entreprises sont généralement considérées comme moins exposées au risque de faillite que les petites
Endettement (TLTA)	Total des dettes/ Total actif	Plus l'endettement est élevé , plus la fragilité financière augmente
Liquidité (WCTA)	Fond de roulement/Total actif	Mesure la capacité de l'entreprise à couvrir ses besoins à court terme
Liquidité à court terme (CLCA)	Passif circulant/Actif circulant	Indique la pression des dettes à court terme sur l'actif disponible
Situation nette négative (ENEG)	Variable binaire	*1 si dettes > Actifs * 0 sinon
Rentabilité (NITA)	Résultat net/Total actif	Une rentabilité faible signale une détérioration de la santé financière
Capacité de remboursement (FUTL)	Flux de Fond/Total des dettes	Mesure la capacité de l'entreprise à rembourser ses dettes grâce aux flux générés.
Pertes répétées (INTWO)	Variable binaire	*1 si l'entreprise enregistre des pertes pendant deux années consécutives. * 0 sinon
Variation du résultat net (CHIN)	Variation du résultat net entre 02 ans	Permet d'observer si la performance financière s'améliore où se détériore.

Source : élaboré par nous-mêmes à partir de (Ohlson, 1980)

On Calcule le score de Ohlson

$$Oscore = -1,32 - 0,407SIZE + 6,03TLTA - 1,43WCTA + 0,0757CLCA - 1,72OENEG - 2,37NITA - 1,83FUTL + 0,285INTWO - 0,521CHIN$$

Après cette calcul on trouve la probabilité de faillite $P(\text{faillite}) = \frac{e^{Oscore}}{1 + e^{Oscore}}$.

Ces méthodes permettent à l'assureur à évaluer l'acheteur étranger cote financier, au premier lieu les ratios principaux sont : ratio de liquidité générale, ratio de liquidité réduite, dette / capitaux propres, dette / total actif, rentabilité des actifs, rentabilité des capitaux propres, couverture des charges financières et cash-flow / dette. Ces indicateurs financiers peuvent donner une première image sur l'entreprise étrangère.

b) La dimension comportementale et qualitative :

La deuxième dimension concerne les éléments **comportementaux et qualitatifs**. Même lorsqu'un acheteur présente des états financiers acceptables, il peut représenter un risque s'il a un mauvais comportement et historique. (Grunert, Norden, & Weber, 2005) Montrent que les informations non financières apportent une valeur ajoutée réelle dans la prédiction du défaut. La combinaison des facteurs financiers et non financiers améliore la prédiction des événements de défaut par rapport à l'utilisation d'un seul type d'information.

- Qualité du management : compétence des dirigeants, sérieux de la gestion, capacité à diriger l'entreprise, fiabilité de l'équipe dirigeante.
- Position de l'entreprise sur son marché : place concurrentielle, solidité commerciale, perspectives sectorielles, capacité à maintenir son activité.
- Note non financière globale, construite à partir des éléments qualitatifs comme le management et la position sur le marché.

Par contre (Petersen & Rajan, 1997) proposent une autre perspective sur les critères qualitatifs, dont on peut retenir les éléments suivants :

- Ancienneté de l'entreprise : l'âge de l'entreprise est un indicateur de réputation et de qualité informationnelle. Une entreprise plus ancienne est généralement plus connue par ses fournisseurs et créanciers.
- Relation avec les fournisseurs : les fournisseurs disposent souvent d'informations privilégiées sur leurs clients grâce aux échanges commerciaux répétés.

- Comportement de paiement : le non-respect des délais ou l'incapacité à profiter des escomptes de paiement rapide peut signaler une détérioration de la situation de l'acheteur.
- Secteur d'activité : le risque varie selon la nature de l'activité.
- Localisation / environnement géographique : est-ce que l'entreprise se situe dans une zone métropolitaine ou non.
- Nature des biens échangés : Les fournisseurs peuvent mieux récupérer ou revendre certains biens en cas de non-paiement.

Généralement, un acheteur qui paie régulièrement ses factures, communique ses informations financières et entretient une bonne réputation auprès de ses partenaires commerciaux sera généralement considéré comme moins risqué. Par contre, un acheteur douteux, souvent en retard de paiement ou associé à des litiges commerciaux peut voir sa limite de crédit réduite ou refusée.

2.3. Les modèles traditionnels d'évaluation de l'acheteur :

L'évaluation traditionnelle du risque de l'acheteur repose sur un ensemble de méthodes financières, statistiques et expertes cette évaluation est essentielle, car elle permet à l'assureur d'estimer le niveau de risque présenté par un acheteur étranger avant d'accorder une limite de crédit ou une couverture. Ces méthodes traditionnelles constituent donc une base fondamentale dans l'analyse du risque, même si elles sont aujourd'hui complétées par des approches plus avancées.

a) Le scoring expert :

Le scoring expert désigne une méthode d'évaluation dans laquelle la décision repose principalement sur l'expérience, les connaissances et le jugement professionnel de l'analyste crédit. Dans cette approche, l'analyste attribue une appréciation ou une note à l'acheteur en s'appuyant sur plusieurs critères : la solvabilité, la liquidité, l'endettement, la rentabilité, l'historique de paiement, la qualité du management, la réputation commerciale et le contexte économique du pays. (SAUNDERS & ALLEN, 2002) expliquent que les approches traditionnelles du risque de crédit accordent une place importante à l'expertise des analystes, notamment lorsque l'information disponible est limitée ou difficile à interpréter.

Dans le cas de l'acheteur étranger, le recours au scoring expert est particulièrement important, car l'information financière peut être incomplète, ancienne ou difficile à vérifier.

Les différences entre les normes comptables, les pratiques commerciales et les niveaux de transparence selon les pays compliquent l'évaluation du risque. L'analyste doit donc compléter l'analyse chiffrée par des éléments qualitatifs, comme la réputation de l'entreprise, son appartenance à un groupe, son comportement de paiement ou la stabilité du pays dans lequel elle exerce son activité.

L'idée selon laquelle l'évaluation du risque de crédit ne doit pas se limiter à des indicateurs financiers, mais doit aussi intégrer des facteurs qualitatifs et comportementaux selon la logique de (Grunert, Norden, & Weber, 2005) Le scoring expert présente ainsi l'avantage d'intégrer des informations qui ne sont pas toujours facilement quantifiables. Par exemple, l'acheteur peut présenter des ratios financiers acceptables, mais être considéré comme risqué en raison de retards de paiement répétés, de litiges commerciaux ou d'un environnement politique instable. À l'inverse, une entreprise récente peut être jugée moins risquée si elle appartient à un groupe solide ou si elle dispose d'une bonne réputation commerciale.

Mais, cette méthode comporte certaines limites. Comme elle repose en grande partie sur le jugement humain, elle peut être influencée par l'expérience personnelle de l'analyste, sa perception du risque ou la qualité de l'information disponible. Deux analystes peuvent parfois interpréter différemment le même dossier. C'est pourquoi le scoring expert est souvent formalisé dans des grilles d'évaluation afin de réduire la subjectivité et de rendre la décision plus cohérente.

b) Les modèles de notation :

Selon (Grunert, Norden, & Weber, 2005) Les modèles de notation, ou systèmes de *rating*, sont des outils qui permettent de classer les acheteurs selon leur niveau de risque de crédit. Dans les systèmes internes de notation, la note globale résulte généralement d'une appréciation combinée de facteurs financiers et non financiers, ce qui permet de transformer plusieurs informations en une décision plus structurée. Dans le contexte de l'assurance-crédit à l'export, la notation de l'acheteur sert principalement à orienter plusieurs décisions : accepter ou refuser une couverture, fixer une limite de crédit, accorder une couverture partielle ou renforcer la surveillance de l'acheteur. Cette idée est confirmée par la pratique de (Coface , 2025), qui distingue la décision complète, la décision partielle et le refus après analyse de la solvabilité de l'acheteur.

Les modèles de notation reposent d'abord sur des critères quantitatifs, qui sont des données financières issues des états financiers de l'acheteur. Ces critères peuvent inclure le chiffre d'affaires, le résultat net, le total du bilan, l'endettement, la liquidité, la rentabilité et la solvabilité. L'utilisation de ces indicateurs est justifiée par les travaux classiques (Altman, 1968), qui montrent que les ratios financiers peuvent être combinés pour prédire la faillite d'une entreprise. (Ohlson, 1980) Confirme également l'importance des données comptables et financières dans l'estimation probabiliste du risque de défaillance.

Par contre, l'évaluation du risque ne peut pas être limitée aux seules données financières. Les critères qualitatifs, comme la qualité du management, la gouvernance, la réputation commerciale, l'historique de paiement et les perspectives du secteur, jouent aussi un rôle important dans l'évaluation du risque de crédit. Comme ils ont mentionné (Grunert, Norden, & Weber, 2005) que la combinaison des facteurs financiers et non financiers améliore la prédiction des défauts futurs par rapport à l'utilisation d'un seul type d'information.

Enfin, les modèles traditionnels de notation peuvent présenter certaines limites lorsque les relations entre les variables financières, comportementales et macroéconomiques deviennent complexes ou non linéaires. À ce niveau, la machine Learning apparaît comme une approche complémentaire intéressante. En effet, (Gambacorta, Huang, Qiu, & Wang, December 2019) montrent, à partir d'une étude menée sur les données d'une fintech chinoise, que les modèles fondés sur la machine Learning et les données non traditionnelles prédisent mieux les défauts et les pertes que les modèles traditionnels, notamment en période de choc économique ou réglementaire.

3. Le machine learning comme outil d'amélioration de l'évaluation de l'acheteur étranger

Avant d'examiner les apports et les limites du machine learning dans l'évaluation du risque de l'acheteur étranger, il convient de préciser ce que recouvre cette notion dans le contexte spécifique du risque de crédit, en distinguant ses fondements conceptuels des méthodes statistiques classiques auxquelles il est souvent comparé, et en montrant pourquoi ses caractéristiques le rendent particulièrement adapté à des problématiques où les données sont nombreuses, hétérogènes et les relations entre variables difficiles à modéliser de façon linéaire.

3.1. Définition du machine Learning dans le contexte du risque de crédit :

Le machine Learning, ou apprentissage automatique, est un ensemble de méthodes informatiques qui permettent à une machine d'apprendre à partir des données. Contrairement à une programmation classique, où les règles sont écrites directement par l'être humain, il permet au système de détecter lui-même des régularités dans les données afin de réaliser une prédiction, une classification ou une aide à la décision.

Selon (Mitchell, 1997), le machine Learning est un programme apprend à partir de l'expérience lorsqu'il améliore sa performance dans une tâche donnée grâce aux données observées. Cette définition est importante, car elle montre que le machine Learning repose sur trois éléments : une tâche, une expérience, et une mesure de performance.

Dans le même sens, (Bishop, 2002) explique que le machine Learning permet d'identifier des structures et des relations cachées dans les données. Il est donc particulièrement utile lorsque les phénomènes étudiés sont complexes, comme le risque de crédit, où plusieurs variables peuvent influencer la décision : situation financière, niveau d'endettement, rentabilité, historique de paiement, secteur d'activité ou risque pays.

Pour (Hastie, Tibshirani, & Friedman, 2009) le machine Learning est proche de l'apprentissage statistique. Il regroupe des méthodes capables de construire des modèles prédictifs à partir de données. Ces modèles peuvent être utilisés pour expliquer certaines relations, mais surtout pour prévoir un résultat futur, tel que la probabilité de défaut ou le niveau de risque de crédit.

Dans le domaine du risque de crédit, (Baesens, Rösch, & Scheule, 2016) montrent que les techniques d'analyse de données et de machine Learning sont de plus en plus utilisées pour le crédit scoring, la classification des clients et l'estimation du risque de défaut. Ces méthodes permettent d'exploiter des volumes importants de données et de prendre en compte des relations plus complexes que les modèles traditionnels.

3.2. Les apports du machine Learning à l'évaluation du risque de crédit :

Le machine Learning apporte une évolution importante à l'évaluation du risque de crédit, car il permet de passer d'une analyse principalement descriptive à une analyse plus prédictive. Dans les méthodes classiques, l'analyste financier observe les ratios, l'historique de paiement, la situation de trésorerie et l'endettement afin d'apprécier la solvabilité du client. Avec le machine Learning, ces informations sont utilisées pour entraîner un modèle capable d'identifier des régularités dans les données passées et de prévoir un événement

futur : défaut, retard de paiement, non-paiement, sinistre, classe de risque ou montant de couverture à recommander.

a) Amélioration de la capacité prédictive :

(TheWorldBankGroup, 2019) Indique que les méthodes innovantes de credit scoring peuvent améliorer l'exactitude des modèles utilisés pour évaluer le risque. Dans le même sens, (EuropeanBankingAuthority, 2023) précise que, dans le risque de crédit, les modèles de machine Learning peuvent améliorer le pouvoir prédictif, notamment pour l'estimation de la probabilité de défaut dans les modèles internes de notation.

b) L'élargissement des données utilisées :

(Finlay, 2012) Les modèles prédictifs appliqués au crédit et à l'assurance ne reposent pas seulement sur des informations financières classiques, mais nécessitent aussi une phase de collecte, de préparation et de transformation de données variées. Dans cette logique, les données comportementales, transactionnelles et commerciales peuvent compléter les indicateurs traditionnels afin d'améliorer la qualité de l'évaluation du risque. Par contre (TheWorldBankGroup, 2019) mentionne que le passage vers des méthodes plus innovantes, notamment fondées sur le machine Learning, a élargi le champ des informations considérées comme utiles pour les modèles et les décisions de crédit. Ainsi, l'évaluation du risque ne repose plus uniquement sur les données financières classiques ou l'historique de crédit, mais peut aussi intégrer des données transactionnelles, comportementales et commerciales, afin d'améliorer la qualité de l'analyse et de mieux apprécier la solvabilité des clients ou des entreprises.

c) La détection des relations complexes et non linéaires :

Les techniques de machine Learning sont capables de refléter des relations non linéaires entre les variables. Ces techniques peuvent être particulièrement utiles pour améliorer la différenciation du risque, surtout lorsque les modèles classiques présentent une performance limitée. Dans ce sens, le machine Learning permet non seulement d'améliorer le pouvoir discriminant du modèle, mais aussi d'identifier les facteurs de risque les plus pertinents dans l'évaluation d'un emprunteur ou d'un acheteur étranger. (EuropeanBankingAuthority, 2023)

d) L'automatisation et le gain d'efficacité :

L'automatisation partielle du processus d'évaluation du risque. Le machine Learning aide à traiter plus rapidement les dossiers, à réduire certaines tâches répétitives et à accélérer les

décisions de crédit. L'automatisation est des processus comme l'une des opportunités offertes par les méthodes innovantes de credit scoring. (TheWorldBankGroup, 2019)

e) La meilleure segmentation des profils de risque

Le credit scoring ne sert plus uniquement à accepter ou refuser une demande de crédit, mais qu'il intervient dans plusieurs étapes du processus de crédit : fixation des limites, tarification selon le risque et ciblage de certains clients potentiels. Cette évolution permet aux institutions financières de classer les emprunteurs selon des niveaux de risque plus fins, au lieu de les traiter comme un groupe homogène. La combinaison de nouvelles sources de données avec des algorithmes innovants permet une évaluation plus détaillée et plus segmentée de la solvabilité des consommateurs et des entreprises. (TheWorldBankGroup, 2019)

f) L'amélioration de l'accès au crédit :

(TheWorldBankGroup, 2019) Indique que les méthodes innovantes de credit scoring peuvent favoriser l'inclusion financière et l'accès au crédit, notamment lorsque les données classiques sont insuffisantes. Dans ce sens, le machine Learning peut compléter les méthodes traditionnelles en utilisant d'autres informations pour mieux apprécier la solvabilité d'un client.

g) L'appui à la gestion du risque tout au long du cycle de crédit

(TheWorldBankGroup, 2019) Montre que le credit scoring ne se limite plus à une fonction simple de sélection, c'est-à-dire accepter ou refuser une demande de crédit. Il constitue aujourd'hui un outil d'aide à la décision mobilisé à plusieurs étapes du cycle de crédit. Il peut notamment servir à évaluer le profil de risque du client, à fixer une limite de crédit adaptée, à ajuster la tarification selon le niveau de risque et à améliorer la gestion de la relation client

h) L'explicabilité comme condition d'utilisation :

(EuropeanBankingAuthority, 2023) Indique que la complexité des modèles de machine Learning peut améliorer leur performance, mais elle réduit parfois leur explicabilité et la compréhension de leur fonctionnement. L'autorité souligne également que cette complexité entraîne des difficultés liées à l'explication des résultats, à la traçabilité et à la compréhension du modèle par les fonctions de management. D'autre part (TheWorldBankGroup, 2019) mentionne que les décisions fondées sur le credit scoring doivent rester explicables, transparentes et équitables au sein des institutions doivent être

capables d'expliquer aux consommateurs les décisions prises à partir du score, ainsi qu'aux régulateurs la manière dont le scoring est intégré dans leurs processus et la logique générale de son fonctionnement.

3.3. Les limites du machine Learning dans un domaine sensible :

Dans un domaine sensible comme l'assurance-crédit à l'export, le machine Learning ne doit pas être présenté uniquement comme un outil d'amélioration de la prédiction. Il comporte aussi plusieurs limites, car ses résultats peuvent influencer des décisions importantes : acceptation ou refus d'un crédit, fixation d'une limite, tarification du risque ou surveillance d'un client.

Premièrement, son efficacité est dépendante fortement de la **qualité et de la disponibilité des données** : si les données sont incomplètes, anciennes ou biaisées, le modèle peut produire des résultats peu fiables. (FinancialStabilityBoard, 2017).

Ensuite, les modèles peuvent être confrontés au **déséquilibre des classes**, car les cas de défaut ou de sinistre sont souvent moins nombreux que les cas normaux. Cela peut conduire le modèle à mieux reconnaître les bons clients qu'à identifier les clients réellement risqués. Une autre limite est le **surapprentissage** c'est à dire les institutions financières rencontrent des difficultés à traiter ce problème statistique, surtout avec les modèles complexes (EuropeanBankingAuthority, 2023)

La **difficulté d'interprétation** constitue aussi une limite importante. Les modèles avancés peuvent donner de bonnes prédictions, mais sans toujours expliquer clairement les raisons de la décision. (BankOfEngland, 2019)

Enfin, le modèle dépend du **contexte dans lequel il a été entraîné**. Un modèle construit sur un pays, un secteur ou une période économique donnée peut devenir moins fiable dans un autre contexte. La représentativité des données et les exigences opérationnelles liées aux données sont des difficultés importantes lors de l'utilisation du machine Learning dans les modèles de crédit. (EuropeanBankingAuthority, 2023)

4. L'explicabilité comme condition d'acceptation des modèles prédictifs

Avant d'examiner les outils permettant de rendre les modèles prédictifs plus transparents et les exigences éthiques, managériales et réglementaires qui encadrent leur utilisation, il est nécessaire de clarifier ce que recouvrent les notions d'explicabilité et d'interprétabilité, en les distinguant de la performance prédictive pure, et en montrant pourquoi la complexité

croissante des algorithmes soulève des questions qui dépassent le seul cadre technique pour toucher directement à la responsabilité des décisions prises à partir de ces modèles.

4.1. Définition de l'explicabilité et de la "boîte noire" :

(Guidotti, et al., 2018) Définissent la notion de **boîte noire** comme un modèle dont le fonctionnement interne est difficile à comprendre pour les utilisateurs, même lorsque ses résultats semblent performants. Dans ce cas, l'utilisateur observe les données d'entrée et la prédiction finale, mais il ne comprend pas clairement comment le modèle a relié les variables entre elles pour produire cette décision. Des nombreux systèmes d'aide à la décision performants sont construits comme des boîtes noires, c'est-à-dire des systèmes qui cachent leur logique interne, ce qui pose à la fois un problème pratique et un problème éthique.

Selon (Rudin, 2019) L'interprétabilité correspond à la capacité de comprendre directement le fonctionnement du modèle. Les modèles simples, comme la régression logistique, les arbres de décision simples ou les systèmes de règles, sont généralement considérés comme plus interprétables, car il est possible de suivre plus facilement la logique qui relie les variables explicatives au résultat.

Pour (Molnar, 2025) l'explicabilité est une notion plus large. Elle désigne l'ensemble des méthodes permettant de rendre compréhensible le résultat d'un modèle, même lorsque celui-ci est complexe. Un modèle peut donc être peu interprétable dans sa structure interne, comme une forêt aléatoire, un gradient boosting ou un réseau de neurones, mais devenir partiellement explicable grâce à des outils comme SHAP, LIME, l'importance des variables ou les analyses de sensibilité. On dit que le machine Learning interprétable est un ensemble de méthodes permettant de comprendre les modèles et leurs décisions, notamment lorsque les algorithmes ne fournissent pas spontanément d'explication.

Entre ces deux notions « interprétabilité, explicabilité » on distingue entre la performance et la compréhension. La performance mesure la capacité du modèle à prédire correctement un événement, comme le défaut, le sinistre ou le niveau de risque. La compréhension mesure la capacité des utilisateurs à savoir pourquoi cette prédiction a été produite. (EuropeanBankingAuthority, 2023) Indique clairement que l'augmentation de la complexité peut améliorer la performance du modèle, mais au prix d'une baisse de l'explicabilité et de la compréhension de son fonctionnement. Alors dans le domaine de crédit le modèle doit aussi être explicable, documenté et contrôlable.

4.2. Les outils d'explicabilité mobilisables :

L'explicabilité peut être abordée à travers plusieurs outils. Le premier outil est l'importance des variables. Selon (Molnar, 2025) une variable est considérée comme importante lorsque la modification ou la permutation de ses valeurs entraîne une dégradation de la performance prédictive du modèle. Dans l'assurance-crédit, cette méthode peut montrer, que le score pays, le ratio d'endettement, la rentabilité, l'historique de paiement ou le secteur d'activité sont les variables les plus déterminantes dans l'évaluation de l'acheteur.

Un autre outil qui s'appelle **SHAP**, ou *Shapley Additive exPlanations* c'est un outil plus avancé. (Lundberg & Lee, 2017) Décrit SHAP comme un cadre unifié permettant d'interpréter les prédictions des modèles en attribuant à chaque variable une valeur d'importance pour une prédiction donnée. Autrement dit, SHAP permet de voir comment chaque variable contribue positivement ou négativement au score final d'un phénomène.

Cette méthode rend le modèle sous la forme $f(\mathbf{x}) = \phi_0 + \sum_{i=1}^M \phi_i$

D'où : où ϕ_0 représente la valeur moyenne de référence du modèle et ϕ_i représente la contribution de chaque variable à la prédiction finale. Cette logique permet d'expliquer pourquoi un acheteur donné obtient un score de risque élevé ou faible, en montrant l'effet propre de chaque facteur sur la décision du modèle.

(Bussmann, Giudici, Marinelli, & Papenbrock, 2020) Considèrent que les valeurs de Shapley sont un outil important d'explicabilité dans la gestion du risque de crédit. Leur intérêt réside dans leur capacité à décomposer une prédiction globale en contributions individuelles des variables explicatives.

D'autre part un autre outil souvent mobilisé est **LIME** *Local Interpretable Model-Agnostic Explanations* est une méthode proposée par (Ribeiro, Singh, & Guestrin, 2016) pour expliquer les prédictions d'un classificateur de manière locale et interprétable. Elle est dite locale, car elle ne cherche pas à expliquer le fonctionnement global du modèle, mais uniquement la prédiction donnée pour une observation précise. Elle est aussi dite model-agnostic, car elle peut être appliquée à plusieurs types de modèles sans connaître leur structure interne. Formellement, LIME cherche à trouver un modèle simple g , appartenant à une famille de modèles interprétables \mathcal{G} , qui approxime localement le modèle complexe f autour de l'observation \mathbf{x} . Cette idée peut être représentée par la formule suivante

$$\xi(\mathbf{x}) = \underset{g \in \mathcal{G}}{\operatorname{argmin}} \{ \mathcal{L}(f, g, \pi_{\mathbf{x}}) + \Omega(g) \}$$

Dans cette formule, $\xi(\mathbf{x})$ représente l'explication locale de l'observation \mathbf{x} , \mathbf{f} désigne le modèle complexe à expliquer, \mathbf{g} représente le modèle simple utilisé pour l'explication, $L(\mathbf{f}, \mathbf{g}, \pi_x)$ mesure l'écart entre les prédictions du modèle complexe et celles du modèle simple, π_x donne plus de poids aux observations proches de \mathbf{x} , et $\Omega(\mathbf{g})$ mesure la complexité du modèle explicatif. L'objectif est donc de choisir un modèle \mathbf{g} qui soit à la fois fidèle localement au modèle complexe et suffisamment simple pour être compris par l'utilisateur. Le poids de proximité peut être représenté par une fonction de noyau comme :

$$\pi_x(\mathbf{z}) = e^{\frac{-D(\mathbf{x}, \mathbf{z})^2}{\sigma^2}}$$

$D(\mathbf{x}, \mathbf{z})$ mesure la distance entre l'observation initiale \mathbf{x} et une observation perturbée \mathbf{z} , tandis que σ contrôle la largeur du voisinage local. Plus une observation \mathbf{z} est proche de \mathbf{x} , plus son poids est élevé dans la construction de l'explication. Ensuite, LIME construit souvent un modèle linéaire simple autour de l'observation étudiée :

$$\mathbf{g}(\mathbf{z}) = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p$$

Dans cette équation, les coefficients $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ indiquent l'influence locale des variables sur la prédiction expliquée. Ainsi, LIME permet d'identifier les variables qui ont le plus contribué à une prédiction particulière, tout en rendant la décision du modèle plus compréhensible pour l'utilisateur.

Il existe deux types d'explication : **l'explication globale et l'explication locale**. (Molnar, 2025) L'explication globale désigne l'ensemble des méthodes qui permettent de comprendre le comportement général d'un modèle de machine Learning sur l'ensemble des données. Elle cherche à répondre à la question suivante : *comment le modèle fonctionne-t-il de manière générale ?* Elle permet notamment d'identifier les variables les plus importantes, les relations générales entre les variables et la prédiction, ainsi que la logique dominante suivie par le modèle.

On passe à **l'explication locale** désigne les méthodes qui expliquent une prédiction particulière produite par un modèle pour une observation précise. Elle ne cherche pas à expliquer tout le modèle, mais à répondre à la question suivante : pourquoi le modèle a-t-il donné ce résultat pour ce cas précis ? Les méthodes locales visent à expliquer une prédiction $f(\mathbf{x})$ à partir d'une seule entrée \mathbf{x} , c'est-à-dire à comprendre la décision du modèle pour une observation individuelle. (Lundberg & Lee, 2017)

4.3. L'explicabilité comme exigence éthique, managériale et réglementaire :

Sur le plan **éthique**, l'explicabilité permet d'éviter que les décisions fondées sur des modèles prédictifs soient perçues comme arbitraires ou injustes. Lorsqu'une personne est concernée par une décision automatisée ou semi-automatisée, elle doit pouvoir comprendre les raisons principales qui ont conduit au résultat. Cette compréhension protège l'autonomie de l'individu, car elle lui donne la possibilité de demander des clarifications, de contester une décision ou de signaler une erreur dans les données utilisées. C'est à dire que les systèmes d'intelligence artificielle doivent être transparents et fournir des informations compréhensibles sur les données, les facteurs et la logique ayant conduit à une prédiction ou à une décision. (EuropeanBankingAuthority, 2023)

L'explicabilité est aussi liée à la **justice** et à la non-discrimination. Un modèle peut produire un résultat précis sur le plan statistique, mais utiliser des variables qui créent des effets injustes sur certains groupes ou certaines catégories d'individus. Sans explication, il devient difficile de repérer ces biais ou de vérifier si la décision repose sur des critères pertinents. Alors l'interprétabilité est nécessaire lorsque les systèmes prédictifs sont utilisés pour évaluer des critères comme la sécurité ou la non-discrimination. Ainsi, l'explicabilité ne se limite pas à une question technique ; elle devient un moyen de contrôle éthique des décisions algorithmiques. (Doshi-Velez & Kim, 2017)

Sur le plan **managérial**, l'explicabilité est nécessaire pour permettre aux responsables de comprendre les modèles qu'ils utilisent dans leurs décisions. L'EBA souligne que les modèles de machine learning peuvent améliorer la capacité prédictive, mais que leur complexité pose des difficultés d'explication, de traçabilité et de compréhension par les fonctions de management. Cela montre que l'explicabilité devient une condition de bonne gouvernance interne. (EuropeanBankingAuthority, 2023)

(OCDE, 2023) recommande d'intégrer des mécanismes de supervision humaine afin de réduire les risques liés aux usages prévus ou aux mauvais usages possibles des systèmes algorithmiques. Dans cette logique, l'explicabilité permet au décideur humain de garder un rôle actif dans l'analyse et la validation de la décision.

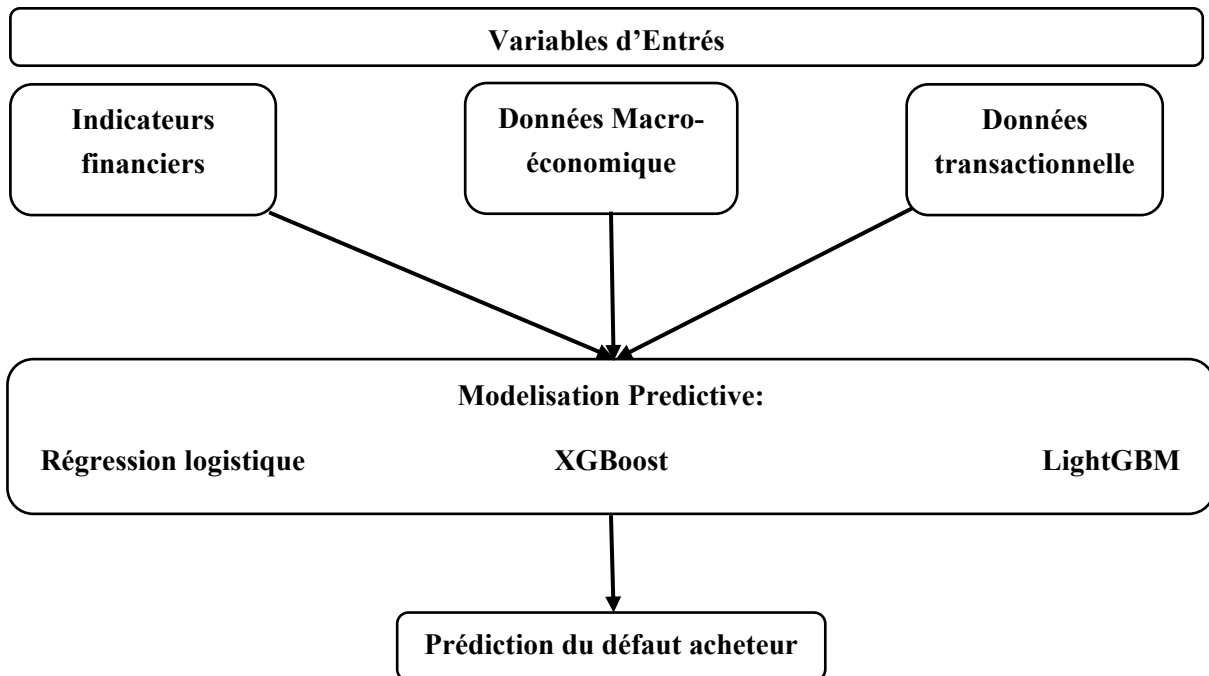
L'explicabilité devient une exigence **réglementaire** lorsque les modèles de machine learning sont utilisés dans des décisions automatisées qui peuvent produire des effets importants sur les personnes, notamment dans le domaine économique ou financier, car le

RGPD encadre les décisions fondées exclusivement sur un traitement automatisé, y compris le profilage. (CNIL, 2016)

La réglementation insiste aussi sur la supervision humaine. Dans les systèmes à haut risque, il ne suffit pas que le modèle soit techniquement performant ; il faut prévoir des mécanismes permettant à des personnes compétentes de surveiller son fonctionnement, d'interpréter ses résultats et d'intervenir lorsque cela est nécessaire. L'article 14 de l'AI Act prévoit que la supervision humaine doit permettre de prévenir ou de réduire les risques pour la santé, la sécurité ou les droits fondamentaux. Cette exigence rejoint directement l'idée selon laquelle l'explicabilité sert à rendre le contrôle humain réellement possible. (EU Artificial Intelligence Act , 2021)

(EuropeanBankingAuthority, 2023) rappelle que l'utilisation des modèles complexes doit rester alignée avec les exigences prudentielles, notamment lorsqu'ils sont utilisés dans des modèles internes de risque. Elle souligne que les difficultés d'explication et de traçabilité peuvent devenir des obstacles à leur adoption.

5. Modèle théorique :



6. Hypothèses :

À partir de la revue de littérature, le risque de défaut de l'acheteur étranger apparaît comme un phénomène multidimensionnel, influencé à la fois par la situation financière de

l'entreprise acheteuse, son environnement macroéconomique, son historique de paiement et les caractéristiques de la relation commerciale. Sur cette base, on peut formuler ces hypothèses suivantes :

- ✓ **H1 : Les indicateurs financiers influencent directement l'évaluation de défaut des acheteurs étrangers.**

(Yazdi, Hanne, & Wang, 2019) Constituent le premier appui de cette hypothèse. Ils montrent que les agences de crédit à l'export évaluent traditionnellement les acheteurs à partir d'un ensemble d'indicateurs financiers précis : le ratio dette/capitaux propres, le ratio dette/EBITDA, le DSCR (taux de couverture des dettes), la couverture des intérêts, la rentabilité, la liquidité et la rotation de l'actif. Ces indicateurs sont présentés comme les variables principales de modèle en permettant d'apprécier la capacité de l'acheteur à honorer ses engagements financiers.

(Petrová, Krügerová, & Koziel, 2020) viennent renforcer cette idée en rappelant que l'évaluation finale du risque combine plusieurs éléments, dont la qualité de crédit du débiteur figure en bonne place, aux côtés du risque pays et de la durée du risque. Cela signifie que la solidité financière de l'acheteur est une composante irréductible de l'analyse, même dans les approches les plus globales.

(Bärtland & Krummaker, 2020) apportent un éclairage complémentaire sur les données de Burnes Union montre que les modèles de machine learning, lorsqu'ils sont appliqués à l'assurance, s'appuient précisément sur des variables financières pour prédire la survenue et la gravité des sinistres.

- ✓ **H2 : Les données macroéconomiques et le risque pays influencent significativement l'évaluation de défaut des acheteurs étrangers.**

(Agarwal, et al., 2023) fournissent le fondement théorique de cette hypothèse. Ils mettent en évidence que les frictions informationnelles constituent un obstacle central au commerce international d'où l'exportateur ne dispose pas d'une information complète, fiable et vérifiable sur son acheteur étranger, notamment en raison de la distance géographique, institutionnelle et économique qui les sépare. Cette incertitude est d'autant plus forte que l'environnement macroéconomique du pays de l'acheteur est instable ou opaque. Les frictions informationnelles sont ainsi présentées comme un obstacle au commerce extérieur comparable aux coûts de transit ou aux barrières tarifaires.

(Petrová, Krügerová, & Kozieł, 2020) précisent que le risque à l'export ne peut jamais être réduit à la seule lecture financière de l'acheteur. Ils décomposent le risque territorial en risque politique, risque lié à la politique commerciale, et risque monétaire et de change, en soulignant que ces dimensions sont fortement liées entre elles et doivent être gérées de façon globale. Ils rappellent également que l'OCDE utilise un modèle fondé sur des données macroéconomiques et financières, pour produire une classification internationale des risques pays, qui sert de référence aux assureurs de crédit à l'exportation.

(Yazdi, Hanne, & Wang, 2019) confirment cette logique en intégrant explicitement le risque pays parmi les indicateurs principaux de leur modèle d'évaluation, au même titre que les ratios financiers.

✓ **H3 : Les données transactionnelles contribuent significativement à la prédiction du défaut acheteur.**

(Yazdi, Hanne, & Wang, 2019) intègrent dans leur cadre d'évaluation plusieurs critères non financiers directement liés au comportement de l'acheteur dans la relation commerciale : l'historique de paiement, la gouvernance, la structure de management et la position de l'entreprise sur son marché. Ces éléments montrent que le comportement passé de l'acheteur constitue une information précieuse pour anticiper son comportement futur, indépendamment de ses états financiers.

(Bärtland & Krummaker, 2020) apportent un appui supplémentaire, les modèles de machine learning les plus performants sont ceux capables d'exploiter des données variées sur les comportements passés des assurés, ce qui correspond à la logique des données transactionnelles appliquées à l'évaluation du risque.

✓ **H4 : Les modèles de machine learning, notamment XGBoost et LightGBM, présentent une meilleure performance prédictive que la régression logistique.**

(Noriega & Rivera, 2023) constituent le principal appui de cette hypothèse. Ils montrent que les modèles de type boosting, notamment XGBoost et LightGBM, Ils précisent que l'évaluation de ces modèles repose sur des indicateurs tels que l'AUC, l'accuracy, le recall, la précision et le score F1. Ils soulignent également que dans le cas de l'assurance-crédit à l'export, le machine learning peut améliorer la qualité de la prédiction du risque, à condition de construire une bonne base de données, de choisir les variables pertinentes et d'interpréter correctement les résultats.

✓ **H5 : L'utilisation des méthodes d'explicabilité permet de mieux comprendre les facteurs déterminants du défaut acheteur :**

La dernière hypothèse mobilise les apports convergents de la littérature sur l'intelligence artificielle explicable (*XAI*). (Liao, Jiao, & Zhang, 2025) proposent les 2 méthodes SHAP et LIME ,permettant d'identifier les variables décisives à l'échelle globale du modèle, mais aussi d'expliquer individuellement chaque décision de classification.

Aussi, (Černevičienė & Kabašinskas, 2024) confirment que dans les domaines sensibles tels que le crédit, la performance prédictive est une condition nécessaire mais non suffisante : la compréhension du processus décisionnel est tout aussi fondamentale.

Conclusion :

Ce chapitre a permis d'établir les fondements théoriques de notre étude, en définissant les concepts clés relatifs à l'évaluation des acheteurs étrangers en assurance-crédit à l'export. Ces apports théoriques fournissent désormais les bases nécessaires pour conduire notre analyse empirique, qui portera sur l'impact des facteurs financiers, du risque pays, des données transactionnelles, des performances des modèles prédictifs et de l'utilisation des outils d'explicabilité. La prochaine étape consistera à présenter notre méthodologie de recherche et le contexte spécifique de l'étude, afin de tester la validité de ces concepts dans un cadre pratique.

CHAPITRE II
CADRE METHODOLOGIQUE
ET ORGANISATIONNEL

Introduction :

Le cadre théorique évoqué dans le premier chapitre, se suit par un second qui définit la démarche scientifique adoptée pour répondre à notre question de recherche principale : comment le Machine Learning peut-il contribuer à l'évaluation de l'acheteur étranger dans le cadre de l'assurance-crédit à l'export ? Il s'agit de préciser les fondements épistémologiques qui orientent notre travail, de décrire l'approche et le design retenus, de présenter les données mobilisées ainsi que les modèles choisis, et d'expliquer les méthodes d'évaluation et d'interprétation des résultats. Ensuite, nous exposerons le cadre organisationnel en présentant l'entreprise d'accueil et le contexte de l'étude de cas.

I. Cadre méthodologique :

1. Posture épistémologique :

Notre recherche s'inscrit dans un **paradigme positiviste**. Nous trouvons que le risque de non-paiement de l'acheteur étranger existe vraiment dans la réalité. On peut mesurer ce risque sur la base des données financières, le comportement de l'acheteur et les conditions économiques internationales. Notre objectif alors est de construire un modèle prédictif capable d'identifier les déterminants de ce risque à partir de données historiques et de produire des résultats vérifiables et fiables. (GUBA & LINCOLN, 1994)

Ce positionnement est aligné avec la nature quantitative de notre démarche et avec l'utilisation d'algorithmes de machine Learning, d'où la finalité est précisément d'extraire des régularités statistiques à partir de données empiriques.

• L'approche méthodologique :

Par rapport l'approche méthodologique nous adoptons une approche **hypothético-déductive**. Partant des hypothèses formulées dans le cadre conceptuel du premier chapitre, notre démarche consiste à tester empiriquement ces hypothèses sur un jeu de données réelles issues du domaine de l'assurance-crédit à l'export. Nous allons essayer de mettre en pratique les idées que nous avons présentées au début par l'application des données réelles provenant du domaine de l'assurance-crédit à l'export. Notre objectif alors est de voir si ces idées sont vraies en pratique. Nous allons utiliser des données réelles pour étudier ce sujet. L'assurance-crédit à l'export est le domaine que nous avons choisi pour cette étude. (CRESWELL, 2009)

Sur le plan méthodologique, il s'agit d'une recherche **quantitative**. Elle repose sur la collecte de données structurées, leur traitement algorithmique et l'interprétation statistique des résultats. Cette approche est justifiée par la nature des données disponibles dans le domaine

de l'assurance-crédit, qui sont principalement des ratios financiers, des scores de risque-pays et des historiques de paiement, tous sont sous forme numérique.

2. Base de données et unité d'analyse

Les données mobilisées dans le cadre de cette recherche proviennent de la **Compagnie Algérienne d'Assurance et de Garantie des Exportations (CAGEX)**, principal organisme public d'assurance-crédit à l'exportation en Algérie. La CAGEX constitue une source de première main pour notre étude, car elle détient les dossiers de risque acheteur, les historiques de sinistres et les informations financières et qualitatives relatives aux acheteurs étrangers couverts.

Les données collectées sont de deux natures :

- Données quantitatives : ratios financiers des acheteurs (bilans, comptes de résultat), scores de risque-pays, délais de paiement, montants de créance.
- Données qualitatives : secteur d'activité, type de contrat, notation interne de la gouvernance, historique de comportement de paiement.

Dans le cadre de cette étude, et à cause de l'absence d'un accès direct à une base de données institutionnelle, la base utilisée dans ce travail est une **base semi-synthétique** de 300 observations, construite à des fins académiques. Elle regroupe des variables financières, des variables qualitatives, des variables pays et des variables relatives aux conditions de crédit. En ce sens, elle vise à reproduire, dans un cadre de recherche, la logique multidimensionnelle de l'évaluation du risque en assurance-crédit., tel que des données simulées construites selon les caractéristiques documentées dans la revue de littérature (Bärtland & Krummaker, 2020) (Yazdi, Hanne, & Wang, 2019). Cette approche mixte permet de pallier les contraintes d'accès aux données confidentielles tout en conservant la cohérence scientifique de la démarche.

La construction de cette base repose sur une articulation entre un niveau **microéconomique** et un niveau **macro-institutionnel**. Le niveau microéconomique concerne les données propres à l'acheteur : chiffre d'affaires, actif, passif, dette, capitaux propres, résultat net, EBIT, ratios de liquidité, d'endettement et de rentabilité. Le niveau macro-institutionnel introduit la dimension pays, notamment à travers la variable de risque pays. La référence à la méthodologie de classification du risque pays de l'OCDE, car l'OCDE indique que cette classification repose sur une approche en deux étapes combinant une évaluation quantitative

du risque et un examen par les participants à l'Arrangement sur les crédits à l'exportation bénéficiant d'un soutien public. (OCDE, n.d.)

3. Logique de collecte et structuration des données :

Les données ont été structurées dans un fichier Excel comportant plusieurs feuilles, dont une feuille principale de travail consacrée aux observations exploitables pour l'analyse. L'importation et la lecture de cette base ont été effectuées sous Python à l'aide de la bibliothèque **pandas**.

La collecte, dans ce sens, ne doit pas être comprise comme une extraction directe d'un portefeuille réel d'assureur-crédit, mais comme une **constitution raisonnée d'une base de recherche**. Cette distinction mérite d'être explicitée dans le mémoire. Elle signifie que la base ne prétend pas refléter parfaitement une base opérationnelle complète ; elle sert plutôt à tester une démarche d'évaluation du risque et à construire un cadre empirique de comparaison entre méthodes traditionnelles et méthodes de machine Learning.

4. Définition de la variable cible :

La variable cible retenue est **default-flag**, convertie en variable binaire **default_num**, où 1 signifie défaut et 0 signifie non-défaut. Ce choix est méthodologiquement pertinent, car il permet de traiter directement la question centrale du mémoire : identifier les acheteurs étrangers présentant une probabilité plus forte de défaut. Sur le plan statistique, cela place la recherche dans un cadre de *classification binaire*, ce qui justifie l'usage initial de la régression logistique, puis la comparaison avec des modèles d'arbres et de boosting.

5. Sélection des variables explicatives :

Le choix des variables explicatives a été guidé par trois critères. Le premier est un critère **théorique**, ont été retenues les variables qui, dans la littérature sur le risque de crédit et dans la pratique de l'évaluation financière, sont susceptibles d'éclairer la solvabilité ou la fragilité de l'acheteur : chiffre d'affaires, taille, niveau d'endettement, liquidité, rentabilité, couverture des intérêts, contexte pays, secteur d'activité, durée du crédit, type de paiement et garanties.

Le deuxième est un critère **méthodologique**. Les variables purement identificatoires, comme l'identifiant ou le nom de l'acheteur, ont été exclues. Les variables susceptibles de créer une fuite d'information ont été supprimées du modèle principal. Cette précaution vise à éviter que le modèle n'apprenne artificiellement une réponse déjà incorporée dans la base, ce qui affaiblirait sa portée analytique.

Le dernier est un critère **empirique**. Les variables ont été examinées au regard de leur qualité, de leur cohérence économique, de leur niveau de redondance et de leur disponibilité au moment où une décision d'assurance-crédit serait censée être prise. Cette logique est importante, car un bon modèle n'est pas seulement un modèle performant ; c'est aussi un modèle fondé sur des informations réellement mobilisables dans la pratique.

6. Préparation des données :

La préparation des données a constitué une étape essentielle. Elle a d'abord consisté à identifier les valeurs manquantes, les types de variables, les éventuelles incohérences de libellés et les valeurs extrêmes. Ensuite, les variables ont été séparées entre variables numériques et variables qualitatives. Cette distinction a permis d'appliquer à chaque groupe un traitement spécifique.

Pour les valeurs manquantes, le travail a mobilisé une imputation simple. Dans ce mémoire, cette option est justifiée par la taille relativement limitée de la base et par la volonté de conserver une procédure lisible et reproductible. Pour les variables qualitatives, l'encodage a été réalisé par **OneHotEncoder**, un mécanisme de codage transformant les catégories en colonnes binaires. Cette étape est nécessaire pour rendre les variables qualitatives exploitables par de nombreux algorithmes de classification.

7. Outils mobilisés :

L'environnement de travail retenu est **Python**. La structuration des données a été réalisée avec **pandas**, la visualisation descriptive avec **matplotlib** et **seaborn**, et la modélisation avec **XGBoost**, **LightGBM** et **Régression logistique**. Le choix de Python s'explique par la complémentarité de ces bibliothèques et par leur large utilisation dans les travaux appliqués de machine Learning.

- **Régression logistique** : est une méthode statistique utilisée lorsque la variable à expliquer est de nature qualitative, le plus souvent binaire. Elle permet d'estimer la probabilité d'appartenance d'une observation à une catégorie donnée, par exemple « acheteur risqué » ou « acheteur non risqué ». Ce modèle est particulièrement adapté pour analyser la relation entre une variable dépendante et un ensemble de variables explicatives, comme des ratios financiers, des scores pays ou des caractéristiques transactionnelles. (Hosmer, Lemeshow, & Sturdivant, 2013)

- **XGBoost** : ou Extreme Gradient Boosting, est un algorithme de machine Learning fondé sur le principe du boosting par arbres de décision. Contrairement au Random Forest, qui construit plusieurs arbres de manière relativement indépendante, XGBoost construit les arbres de façon séquentielle : chaque nouvel arbre cherche à corriger les erreurs commises par les arbres précédents. XGBoost est un système de tree boosting évolutif, largement utilisé pour atteindre de très bonnes performances prédictives dans plusieurs problèmes de machine Learning. (Chen & Guestrin, 2016) . Dans cette recherche, XGBoost est mobilisé pour améliorer la capacité de prédiction du risque, surtout lorsque les relations entre les variables sont complexes ou non linéaires
- **LightGBM : Light Gradient Boosting Machine**, est une méthode de gradient boosting par arbres de décision capable d'accélérer le processus d'entraînement tout en conservant une performance prédictive élevée. (Ke, et al., 2017) . Dans cette étude, LightGBM est utilisé pour comparer sa performance avec celle de la régression logistique, du Random Forest et de XGBoost. Son intérêt réside principalement dans sa rapidité d'exécution et sa capacité à traiter des relations complexes entre les variables explicatives du risque.

8. Procédure de modélisation et validation :

La base a été divisée en trois sous-ensembles : **entraînement**, **validation** et **test**. L'ensemble d'entraînement sert à ajuster les modèles ; l'ensemble de validation sert à régler certains choix, notamment le seuil de décision ; et l'ensemble de test sert à fournir une évaluation finale hors apprentissage. En complément, une **validation croisée stratifiée** a été utilisée lors du tuning. Cette méthode conserve la proportion des classes dans les plis de validation, ce qui est particulièrement utile dans les problèmes de classification où les classes ne sont pas parfaitement équilibrées.

9. Ajustement du seuil de décision et critères d'évaluation :

L'évaluation des modèles ne s'est pas limitée à l'accuracy. Dans un problème de défaut, cette métrique peut être trompeuse, surtout si la classe majoritaire domine. L'analyse est donc appuyée sur plusieurs indicateurs : precision, recall, F1-score, matrice de confusion, courbe ROC.

10. Portée et limites méthodologiques :

Cette méthodologie présente plusieurs atouts. Elle relie clairement la posture de recherche, la structure de la base, les outils de préparation des données et la logique de modélisation. Elle permet aussi de comparer des modèles de complexité différente dans un cadre homogène d'évaluation. En même temps, certaines limites doivent être reconnues. La première concerne la nature semi-synthétique de la base, qui ne reproduit pas toute la richesse d'un portefeuille réel. La deuxième tient à la taille de l'échantillon, qui reste relativement modeste pour certains algorithmes avancés. La troisième concerne le fait que certaines variables de la base résultent d'une logique de construction académique et ne doivent pas être assimilées à des données de marché observées directement. Ces limites n'enlèvent pas l'intérêt du travail.

II. Contexte de la recherche

Dans cette partie on va voir le positionnement de l'assurance-crédit à l'export international et l'intégration des modèles avancés comme le machine Learning.

- **L'assurance-crédit à l'exportation, réglementation et adoption des modèles avancés :**

Au niveau international, L'assurance-crédit à l'exportation est fortement encadrée, surtout lorsque l'assurance ou la garantie est soutenue par l'État. (OCDE, n.d.) joue un rôle central dans la régulation des crédits à l'exportation officiellement soutenus, car elle constitue le principal forum où sont négociées, appliquées et surveillées les disciplines internationales en matière de crédits à l'exportation. Ces disciplines visent à garantir une concurrence équitable entre exportateurs et à éviter que les agences publiques ne deviennent un instrument de subvention déguisée.

L'Arrangement de l'OCDE s'applique notamment aux crédits à l'exportation officiellement soutenus dont la durée de remboursement est égale ou supérieure à deux ans, tout en excluant certains domaines comme les équipements militaires et les produits agricoles. Les crédits conformes à cet Arrangement bénéficient également d'une forme de sécurité juridique dans le cadre de l'Accord de l'OMC sur les subventions et les mesures compensatoires. Ainsi, l'assurance-crédit à l'exportation se situe au croisement de la politique commerciale, de la gestion du risque et de la réglementation internationale. (OCDE, n.d.)

En Europe, le marché est marqué par une articulation entre assureurs privés, agences publiques de crédit à l'exportation et règles prudentielles. Pour les risques à court terme considérés comme « commercialisables », notamment à l'intérieur de l'Union européenne et dans certains pays développés, la Commission européenne considère que la couverture doit en principe être laissée aux assureurs privés, ou à des organismes publics agissant comme des opérateurs de marché. (European Commission, 2021). D'autre part les compagnies d'assurance et de réassurance européennes sont soumises au régime **Solvabilité II**, le régime prudentiel applicable aux entreprises d'assurance et de réassurance dans l'UE. Elle est entrée en vigueur en janvier 2016. Il adopte une approche fondée sur les risques qui permet d'évaluer la « solvabilité globale » des entreprises d'assurance et de réassurance au moyen de mesures quantitatives et qualitatives. Les principales caractéristiques du cadre réglementaire Solvabilité II sont : la cohérence avec le marché d'où les actifs et passifs sont évalués au montant pour lequel ils peuvent être échangés, transférés ou réglés sur le marché ; Basé sur le risque si des risques plus élevés entraîneront une exigence de capital plus élevée pour couvrir les pertes imprévues ; ainsi la proportionné : les exigences réglementaires sont appliquées d'une manière qui est proportionnée à la nature, à l'ampleur et à la complexité des risques inhérents à l'activité des entreprises d'assurance et de réassurance. Enfin la surveillance du groupe : les autorités de surveillance renforcent la coordination et l'échange d'informations au sein des collègues d'autorités de surveillance afin d'améliorer la surveillance transfrontalière des groupes d'assurance et de réassurance. (EIOPA , n.d.)

Dans ce contexte, l'adoption des modèles avancés, notamment le machine learning, devient progressivement un sujet stratégique. (European Banking Authority, 2023) indique que l'augmentation de la disponibilité des données, des capacités de stockage et de la puissance de calcul crée de nouvelles possibilités pour utiliser des modèles de machine learning dans le risque de crédit. Toutefois, elle souligne aussi que ces modèles sont souvent plus complexes que les méthodes traditionnelles, comme la régression ou les arbres simples, et qu'ils peuvent être moins transparents.

On passe à l'Angleterre, dans un rapport public, il est mentionné que le Machine Learning est devenu un moteur de performance incontournable pour deux tiers des institutions financières britanniques, le ML s'est solidement ancré dans les activités bancaires et assurantielles, dépassant les simples prototypes pour automatiser des fonctions critiques tant en front-office (relation client) qu'en back-office (lutte contre le blanchiment et détection de la fraude). Cette transformation n'est pas sans défis : si la réglementation n'est pas perçue

comme un frein, les systèmes informatiques obsolètes et la complexité des modèles imposent une vigilance accrue. Pour prévenir toute dérive algorithmique, les firmes privilégient une approche de « l'humain dans la boucle » et une gouvernance interne rigoureuse, illustrant ainsi une volonté de concilier l'agilité de l'intelligence artificielle avec les impératifs de stabilité et de responsabilité fiduciaire. (Bank of England , 2019)

Aux États-Unis l'Export-Import Bank of the United States représente l'un des principaux instruments publics de soutien aux exportations. Son fondement juridique repose sur l'Export-Import Bank Act 1945, qui crée l'EXIM comme agence des États-Unis chargée de faciliter les exportations de biens et services américains. L'EXIM propose une assurance-crédit export couvrant les créances étrangères contre les risques commerciaux et politiques, ce qui permet aux entreprises américaines d'accorder des conditions de paiement plus souples à leurs acheteurs étrangers. Son autorité pour conduire de nouvelles opérations a été prolongée jusqu'au 31 décembre 2026 par l'Export-Import Bank Extension adopté dans le cadre du Further Consolidated Appropriations Act de 2020. (The Charter of the Export-Import Bank of the United States, 2021)

Au Japon, **NEXI Nippon Export and Investment Insurance** couvre les pertes subies lorsqu'une entreprise japonaise ne peut pas exporter ou ne peut pas recouvrer ses créances après expédition, à cause d'événements tels que la guerre, les restrictions d'importation, les catastrophes naturelles ou la faillite de l'acheteur étranger. (NEXI -Nippon Export and Investment Insurance -, n.d.)

En Afrique, le besoin d'assurance-crédit et de garanties est particulièrement important, car les opérations commerciales sont souvent exposées à des risques politiques, de transfert, de liquidité, de change et de solvabilité des acheteurs. **ATIDI African Trade & Investment Development Insurance**, fondée en 2000 par pays d'Afrique orientale, elle facilite les investissements et les échanges en Afrique en fournissant des couvertures contre les risques de commerce et d'investissement. Elle se présente comme un outil de réduction du risque pour accompagner le commerce africain et la mise en œuvre de la ZLECAf (Zone de Libre-Échange Continentale Africaine) (ATIDI, n.d.). En Afrique du Sud, elle dispose également de l'**ECIC Export Credit Insurance Corporation of South Africa**, est une agence publique créée en 2001 sous l'Export Credit and Foreign Investments Insurance Act de 1957, afin de couvrir les risques politiques et commerciaux des exportateurs sud-africains de biens d'équipement et de services associés. (ECIC, n.d.)

L'Union africaine a adopté en 2024 une **stratégie continentale de l'intelligence artificielle** afin d'encadrer le développement et l'utilisation de l'IA selon une vision africaine, responsable et orientée vers le développement. Cette stratégie considère l'IA comme un levier de transformation économique et sociale, capable de soutenir l'innovation, l'emploi, les services publics, la santé, l'éducation, l'agriculture, la finance. Elle insiste cependant sur la nécessité d'une IA adaptée aux réalités africaines, fondée sur les données, les langues, les besoins et les priorités du continent. Aussi sur l'éthique, la protection des droits humains, la non-discrimination, la transparence, la responsabilité et le contrôle humain, surtout dans les secteurs sensibles comme la finance, l'assurance. La stratégie recommande également de renforcer les cadres juridiques liés à la protection des données, à la cybersécurité et à la gouvernance numérique, tout en développant les compétences, les infrastructures et la recherche en IA. Ainsi, cette stratégie base sur l'adoption de l'intelligence artificielle en Afrique ne doit pas seulement rechercher la performance technique, mais aussi garantir une utilisation sûre, équitable, explicable et conforme aux objectifs de développement du continent. (UnionAfricaine, 2024)

Dans ce mouvement général, l'Algérie occupe une position particulière en assurance-crédit par La Compagnie Algérienne d'Assurance et de Garantie des Exportations, CAGEX, s'inscrit dans ce cadre comme l'acteur spécialisé de l'assurance-crédit à l'exportation de l'Etat .

- **Présentation de la CAGEX (CAGEX , n.d.)**

La CAGEX est l'**Agence Officielle de Crédit à l'Exportation (ECA)** de la République Algérienne Démocratique et Populaire. Il s'agit d'une entreprise publique spécialisée dans l'assurance-crédit à l'exportation, la garantie des risques commerciaux et politiques, ainsi que l'accompagnement des opérateurs économiques algériens dans leur internationalisation. Elle intervient exclusivement sur les exportations des **biens et services hors hydrocarbures**, en couvrant les risques d'impayés liés aux ventes à crédit et en facilitant l'accès des exportateurs au financement bancaire grâce à la cessibilité des polices d'assurance.

La CAGEX positionne son action au cœur de la stratégie nationale de diversification de l'économie algérienne, en particulier dans le cadre de la Zone de Libre-Échange Continentale Africaine (ZLECAf) et des forums économiques panafricains et islamiques.

- **Cadre et forme juridique**

- a) **Cadre juridique :**

La base légale de la création de la CAGEX est l'**Ordonnance n° 96/06** du 10 janvier 1996 relative à l'assurance-crédit en Algérie et à la gestion des risques liés aux crédits à l'exportation (à l'exclusion des hydrocarbures), ainsi que ses textes d'application (décrets exécutifs et arrêtés interministériels). (JOURNALOFFICIEL, 1996)

Selon l'article 01 de ladite ordonnance : « L'assurance-crédit à l'exportation garantit dans les conditions prévues par cette ordonnance et par le contrat d'assurance, le recouvrement des droits liés aux opérations d'exportation, contre les risques commerciaux, politiques, de non-transfert et des catastrophes. »

Cette ordonnance reste le texte fondateur qui :

- Définit le monopole de la CAGEX en matière d'assurance-crédit à l'exportation ;
- Distingue explicitement les risques commerciaux (pris en charge par la CAGEX pour son propre compte) des risques politiques (pris en charge pour le compte de l'État) ;
- Encadre les procédures d'indemnisation, les taux de couverture et les relations avec les banques et la Banque d'Algérie.

- b) **Forme juridique :**

La CAGEX est une Société par Actions (SPA) de droit algérien, avec un statut d'Une Entreprise Publique Economique, de capital social a 100% public, placée sous la tutelle de l'État (Ministère des Finances / Ministère du Commerce).

Son capital social est de 10 Milliards de Dinars Algériens (10000000000DA), et ses actionnaires sont les institutions financières publiques, majoritaire par le trésor public 40% c'est-à-dire 4 milliards de dinars algériens et le reste de 60% répartie entre 5 Banques et 5 compagnies d'assurance Publiques 6 milliards de dinars algériens. On peut dire que son évolution de Capital Social est répartie comme suit

- Dès que Création (1996) : 2 milliards dinars répartis à parts égales (10% chacun)
- Avant 2024 : Le capital social a été augmenté pour atteindre un montant de 3 milliards dinars algériens.

- Depuis début 2024 : 10 milliards dinars algériens avec le Trésor public comme actionnaire principal à 40%

On peut dire que la CAGEX opère donc en double casquette (double nature d'intervention)

- Commerciale (risques commercial) ;
- D'intérêt public (risques politiques pour le compte de l'État).

Enfin, Elle est soumise à la réglementation des entreprises publiques économiques et à la gouvernance des sociétés par actions.

• **Mission et objectif**

La promesse de la compagnie est comme suit « *LA CAGEX À VOS CÔTÉS, LE MONDE À VOTRE PORTÉE* »

La CAGEX offre une couverture pour les risques suivants :

- Risques commerciaux : Pour son propre compte et sous la supervision de l'État, elle couvre les créances découlant d'acheteurs privés, au profit des exportateurs et des opérateurs nationaux actifs sur le marché local.
- Risques politiques : Pour le compte de l'État et sous sa supervision, elle couvre les risques politiques, les risques de transfert ainsi que les risques catastrophiques au profit des exportateurs algériens.
- Risque de rupture de marché : C'est le risque qu'une opération soit interrompue après que l'opérateur a déjà commencé la production et constitué des stocks destinés à l'exportation conformément aux termes d'un contrat, mais avant la livraison et la facturation. Ce risque ne concerne pas les créances, mais la production en cours qui, si elle est spécifique, ne peut être revendue à un autre client.

D'où viendra **ses objectifs** stratégiques comme suit :

- Couvrir l'ensemble des risques (commerciaux, politiques, rupture de marché) ;
- Favoriser la cessibilité des garanties auprès des banques pour débloquer des financements ;
- Fournir des services à valeur ajoutée (information, recouvrement, partenariats) ;
- Contribuer activement à la mise en œuvre de la ZLECAf et aux forums économiques internationaux.

- **Activités et produits proposés :**

On résume dans ce tableau représentatif :

Tableau 4: Représente les produits fournis par la CAGEX

Produit	Description	Bénéficiaires
Police globale d'assurance-crédit à l'exportation	Destinée aux exportateurs de biens et de services, qui réalisent des ventes répétitives à crédit (ventes associées à des délais de paiement), avec plusieurs acheteurs étrangers.	Exportateurs réguliers multi-acheteurs
Police individuelle d'assurance-crédit à l'exportation	Est destinée à couvrir les opérations ponctuelles d'exportation de biens et de services avec un seul acheteur étranger contre les risques d'impayés.	Exportateurs pour contrats isolés
Police d'assurance-crédit domestique	Couvre les risques de non-paiement des factures, à leur échéance, pour les transactions commerciales sur le marché national.	Opérateurs sur le marché national
La vente d'informations	Fournit, sur demande, des informations financières, économiques et commerciales aux clients, qu'ils soient assurés ou non.	Toute entreprise Et partenaires
Le recouvrement des créances	En plus de fournir une assurance contre les risques de non-rapatriement des créances, la CAGEX offre des services de recouvrement pour aider les exportateurs et les opérateurs économiques qui n'ont pas assuré	Exportateurs ayant ou non une police

	leurs exportations ou ventes à crédit, contre les risques d'impayés	
Police Foire et exposition	Est destinée aux exportateurs et entreprises de droit algérien qui participent aux foires, salons spécialisés et expositions - ventes à l'étranger.	Participants Algériens à des événements internationaux
Partenariats & Mise en relation	Une solution unique en Algérie qui combine prospection commerciale et gestion des risques.	Entreprises cherchent des partenariats étrangers

Source : Elaboré par nous-mêmes à partir de site officiel de la CAGEX :

<https://www.cagex.dz/fr/Produits>

- **Gouvernance et organisation interne**

La CAGEX est dirigée par un Président Directeur Général nommé par l'État. Elle dispose d'une structure classique d'entreprise publique avec des pôles fonctionnels (Operations techniques, Commerciale et Marketing, Activités support), comme indique l'organigramme. (Voir Annexe unique)

- **Direction Arbitrage et information commerciale**

Cette direction est le métier cœur de la CAGEX, où elle évalue l'acheteur quel que soit au niveau international ou national, elle est divisée en deux sous-directions : la première est information commerciale, et la deuxième est arbitrage

- Sous-direction information commerciale** : sa fonction principale est la collecte et l'analyse des données administratives, commerciales, économiques et financières ainsi que le développement de la base de données des opérateurs économiques de la CAGEX. Cette dernière est scindée en deux services distincts, un est concentré sur la vente de l'information commerciale au profit des assureurs crédits internationaux des opérateurs nationaux tel que la : COFACE, EXIM BANK TURK, ICIEC, Elissphere ...etc., et le deuxième service a pour objet, la gestion de la base de données des opérateurs et les études économiques.

Figure 3: Représente Les partenaires de la CAGEX



Source : site officiel de la CAGEX <https://www.cagex.dz/en/home>

b) Sous-direction Arbitrage : elle est chargée de souscrire et d'évaluer les risques, collecter et analyser les informations concernant les entreprises, veuille à la constitution de banques de données sur les acheteurs, établir des bilans sur la situation économique par secteur et par pays, suivre et gérer les engagements de la Compagnie. Le département est divisé en deux différents services, le premier s'occupe de l'arbitrage des acheteurs sur le marché local et le second se concentre sur l'arbitrage des acheteurs étrangers. Au niveau de la CAGEX, les arbitres sont affectés selon la nature d'activité des Polices d'assurance en leur portefeuille. Il y a ceux qui sont chargés des arbitrages à export et d'autres qui se spécialisent dans les arbitrages domestiques, ces derniers sont affectés selon des secteurs d'activités bien déterminés. Ceci leur permet d'avoir une connaissance approfondie du marché et de ses spécificités et prendre des décisions plus efficaces et plus rapides.

Conclusion :

Ce chapitre a donc posé les fondations pratiques de notre recherche en présentant la démarche méthodologique adoptée ainsi que le contexte organisationnel de l'étude. Il a permis de clarifier la posture de recherche, la nature des données utilisées, les outils mobilisés et les modèles retenus pour l'analyse. De plus, la présentation de la CAGEX et de sa Direction Arbitrage et Information Commerciale a permis de mieux situer le terrain de recherche. Ces éléments constituent une base essentielle pour aborder, dans le chapitre suivant, l'analyse des données et l'interprétation des résultats en lien direct avec notre problématique.

CHAPITRE III
RESULTATS ET DISCUSSION

Introduction :

Ce chapitre présente les résultats obtenus à partir de la modélisation prédictive du risque de défaut des acheteurs étrangers. Après avoir décrit la base de données et les variables mobilisées, nous analysons les performances des différents modèles testés, à savoir la régression logistique, Random Forest, XGBoost et LightGBM. L'objectif est d'identifier le modèle le plus adapté à l'évaluation du risque en assurance-crédit à l'exportation, en tenant compte non seulement de la performance globale, mais aussi de la capacité du modèle à détecter les acheteurs réellement risqués. Enfin, les résultats sont discutés à la lumière des travaux antérieurs et des hypothèses formulées dans la recherche.

I. Présentation des résultats :

1. Analyse descriptive de la base :

La base de données utilisée dans cette étude regroupe **300 acheteurs étrangers** observés à travers **35 variables**. Elle a été construite afin d'évaluer le risque de défaut dans le cadre de l'assurance-crédit à l'exportation. La variable cible retenue est **default_flag**, également codée sous forme numérique par **default_num**. Cette variable permet de distinguer deux catégories d'acheteurs : Les acheteurs **non défaillants**, codés par la valeur **0**, et les acheteurs **défaillants**, codés par la valeur **1**.

Le problème étudié correspond à une **classification binaire**, car le modèle cherche à prédire si un acheteur étranger présente ou pas un risque de défaut. La base couvre plusieurs dimensions importantes pour l'évaluation du risque, notamment les **données financières**, les **données pays**, les **données sectorielles**, les **caractéristiques transactionnelles**, les **garanties**, ainsi que **plusieurs ratios financiers liés à la liquidité, la solvabilité, l'endettement et la rentabilité**. Cette diversité de variables permet d'avoir une vision plus globale du profil de risque de chaque acheteur.

Tableau 5: Représente le contenu général de la base de données

Elément	Description
Nombre d'observation :	300 Acheteurs étrangers.
Nombre des variables :	35 Variables.
Variable cible :	Default -flag / Default-Num.
Type de prédiction :	Classification binaire.
Classe 0 :	Acheteur Non défaillant (Non risqué).
Classe 1 :	Acheteur défaillant (risqué).
Nature des Variables :	Financière, sectorielle, transactionnelle, Pays, Garantie.

Source : Elaboré par nous-mêmes selon la base de données

Ce tableau résume les principales caractéristiques de la base avant de passer à l'analyse descriptive et à la modélisation. Il montre que les données utilisées ne sont pas purement des états financiers des acheteurs, mais intègrent également des informations qualitatives et macro-économique pouvant influencer le risque de non-paiement.

2. Dictionnaire des variables :

Après a présentation générale de la base, il est nécessaire de faire un dictionnaire significatif des principales variables utilisées dans l'étude. Ce dernier permet de comprendre la structure de la base, la nature des informations collectées et le rôle de chaque variable dans l'évaluation du risque de défaut des acheteurs étrangers.

Étant donné que la base contient **35 variables**, afin d'alléger la lecture, seules les variables les plus importantes sont présentées dans le tableau suivant :

Tableau 6: Représente le dictionnaire des principales variables de la base

Variable	Description	Type de variable	Rôle dans l'analyse
Buyer-id	Identifiant unique de chaque acheteur étranger	Qualitatif	Identification de l'acheteur
Country	Pays de l'acheteur étranger	Qualitatif	Analyse du risque pays
Sector	Secteur d'activité de l'acheteur étranger	Qualitatif	Analyse du risque sectoriel
Size segment	Taille de l'entreprise acheteuse	Qualitatif	Caractérisation du profil de l'acheteur
Sales-t-usd	Chiffre d'affaires de l'année t	Quantitatif	Mesure la taille de l'activité
Total_assets_t_usd	Total des actifs de l'acheteur à l'année t	Quantitatif	Analyse de la structure financière
Total_liabilities_t_usd	Total de dettes de l'acheteur à l'année t	Quantitatif	Evaluation de l'endettement
Equity_t_usd	Capitaux propres de l'acheteur	Quantitatif	Mesure de la solvabilité
Net_income_t_usd	Résultat net de l'acheteur	Quantitatif	Analyse de la rentabilité
Current_ratio	Ratio de liquidité générale	Quantitatif	Évaluation de la capacité à faire face aux dettes à court terme
Debt_to_assets	Ratio d'endettement par rapport aux actifs	Quantitatif	Mesure du poids de la dette
Debt_to_equity	Ratio dettes / capitaux propres	Quantitatif	Analyse du levier financier
ROA	Rentabilité des actifs	Quantitatif	Mesure de la performance économique
ROE	Rentabilité des capitaux propres	Quantitatif	Mesure de la rentabilité financière
Country_score	Score associé au risque pays	Quantitatif	Évaluation de l'environnement externe
Oecd_country_class	Classe de risque pays selon la classification OCDE	Qualitatif	Appréciation du risque géographique
Requested_amount_usd	Montant demandé pour la couverture	Quantitatif	Analyse du niveau d'exposition

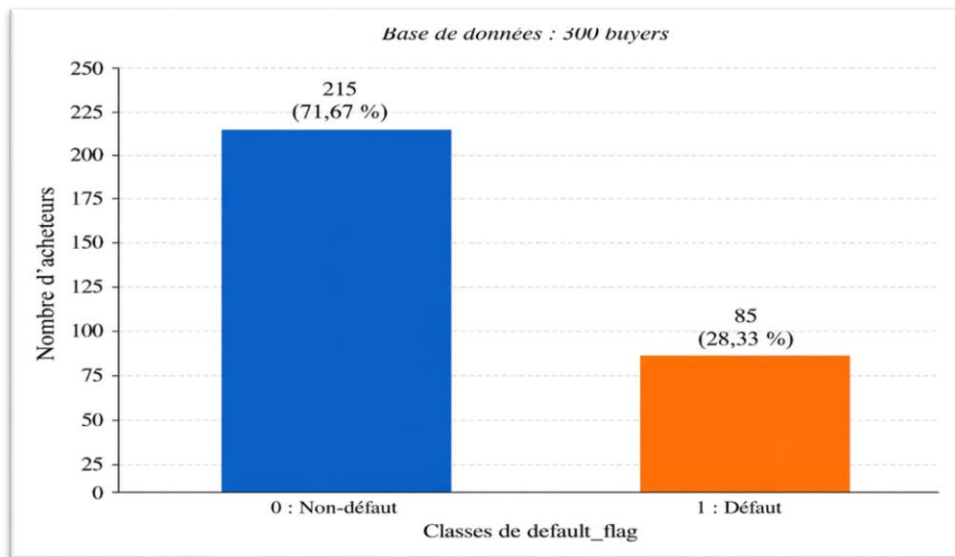
Credit_term_months	Durée du crédit accordé en mois	Quantitatif	Évaluation du risque lié au délai de paiement
Payment_type	Type de paiement utilisé	Qualitatif	Analyse des conditions transactionnelles
Number_guarantees	Nombre de garanties disponibles	Quantitatif	Appréciation du niveau de sécurité
Guarantee_type	Type de garantie fournie	Qualitatif	Analyse de la qualité des garanties
Overall_risk_score	Score global de risque de l'acheteur	Quantitatif	Variable synthétique d'aide à la décision
Insured_amount_usd	Montant assuré estimé	Quantitatif	Résultat lié à la décision de couverture
Premium_rate_pct	Taux de prime estimé	Quantitatif	Évaluation du coût de la couverture
Default_flag / default_num	Indique si l'acheteur est défaillant ou non	Binaire	Variable cible du modèle

Source : Elaborée par nous-mêmes

Ces Variables sont dupliquées en 2 autres années **t-1** et **t-2** : Sales-t-usd, Total_assets_t_usd, Total_liabilities_t_usd, Equity_t_usd et Net_income_t_usd .

Ce dictionnaire montre que la base de données combine plusieurs catégories d'informations. Les variables financières permettent d'évaluer la santé financière de l'acheteur, à travers la liquidité, la rentabilité, la solvabilité et l'endettement. Les variables pays et sectorielles permettent d'intégrer l'environnement externe dans l'analyse du risque. Les variables transactionnelles, comme le montant demandé, la durée du crédit et le type de paiement, renseignent sur les caractéristiques de l'opération commerciale. Enfin, les garanties apportent une information complémentaire sur le niveau de sécurité de la transaction.

Figure 4: Représente la Distribution de la variable cible



Source : Élaboré par nous-mêmes selon le modèle

Commentaire : La distribution de la variable cible montre une domination de la classe des acheteurs non défaillants. En effet, sur les 300 observations de la base, 215 appartiennent à la classe 0, contre 85 à la classe 1. Cette configuration indique un déséquilibre des classes, ce qui peut influencer l'apprentissage des modèles prédictifs et justifie l'utilisation d'indicateurs d'évaluation adaptés à ce type de problème.

3. Préparation des données et choix méthodologiques :

Avant la construction du modèle de prédiction, l'étape de préparation des données a été réalisée afin de rendre la base exploitable par les algorithmes de machine Learning. Cette étape est primordiale, par ce que les modèles ne peuvent pas traiter directement certaines formes de données, notamment les valeurs manquantes ou les variables qualitatives exprimées sous forme textuelle. Dans cette étude, 29 variables explicatives ont été retenues pour prédire la variable cible default_flag / default_num. Ces variables couvrent plusieurs dimensions du risque : la situation financière de l'acheteur, son pays, son secteur d'activité, les caractéristiques de la transaction, ainsi que les garanties associées.

Tableau 7: Représente les variables tenues pour la création de modèle

Catégorie	Variables utilisées
Variables numériques	oecd_country_class, country_score, sales_t_usd, total_assets_t_usd, total_liabilities_t_usd, total_debt_t_usd, current_assets_t_usd, current_liabilities_t_usd, equity_t_usd, net_income_t_usd, ebit_t_usd, interest_expense_t_usd, requested_amount_usd, coverage_ratio, credit_term_months, number_guarantees, current_ratio, debt_to_assets, debt_to_equity, equity_ratio, roa, roe, net_margin, interest_coverage
Variables qualitatives	country, sector, size_segment, payment_type, guarantee_type
Variable cible	default_flag / default_num

Source : élaboré par nous-mêmes

3.1. La démarche :

Le traitement appliqué aux variables dépend de leur nature :

Les **variables numériques** ont été traitées par une **imputation** des valeurs manquantes (permet de remplacer les valeurs absentes afin d'éviter la perte d'observations), puis par une **standardisation** (permet, quant à elle, de mettre les variables sur une échelle comparable). Cette étape est particulièrement importante pour les modèles sensibles à l'échelle des variables, comme la régression logistique.

Les **variables qualitatives** ont été traitées d'une autre manière. Les valeurs manquantes ont été remplacées par **la modalité la plus fréquente**, puis les variables ont été transformées à l'aide d'un encodage **One-Hot**. Ce traitement permet de convertir les catégories textuelles en variables numériques binaires. Par exemple, des modalités comme : Tunisia, Chemicals ou Open account ne peuvent pas être utilisées directement par les modèles. Grâce à l'encodage One-Hot, chaque modalité devient une information exploitable par l'algorithme.

Concernant la séparation des données, deux choix méthodologiques ont été retenus selon le type de modèle utilisé. La base a été divisée en deux parties : **240 observations pour l'apprentissage** et **60 observations pour le test** pour entraîner le modèle sur une partie majoritaire de la base, puis d'évaluer sa capacité de généralisation sur des données non utilisées pendant l'apprentissage.

Pour les modèles optimisés avec **changement de seuil**, une séparation en trois parties a été utilisée : **180 observations pour l'apprentissage**, **60 observations pour la validation** et **60 observations pour le test**. Ce choix est plus rigoureux lorsque l'objectif est d'ajuster le seuil de décision du modèle.

Tableau 8: Représente le découpage de la base de données

Type de découpage	Apprentissage	Validation	Test	Utilisation
Train/ Test	240 Observations	----	60 Observations	Modèles Classiques
Train/Validation/ Test	180 Observations	60 Observations	60 Observations	Modèles avec changement de seuil

Source : Elaboré par nous-mêmes

3.2. Interprétation détaillée des métriques d'évaluation :

Avant d'analyser les résultats des modèles, il est nécessaire de présenter les principales métriques utilisées pour mesurer leur performance. Dans une classification binaire, ces indicateurs ne donnent pas tous la même information. Leur interprétation est donc importante, surtout dans notre cas, car la variable cible est déséquilibrée : les acheteurs non défaillants sont plus nombreux que les acheteurs défaillants.

L'**accuracy** mesure la proportion totale de prédictions correctes réalisées par le modèle. Elle indique donc le pourcentage global d'acheteurs bien classés. Cependant, cette métrique peut être insuffisante lorsque les classes sont déséquilibrées. Dans notre base, les non-défauts représentent environ **71,67 %** des observations. Ainsi, un modèle peut obtenir une accuracy relativement élevée simplement parce qu'il prédit bien la classe majoritaire, tout en détectant mal les acheteurs défaillants.

La **precision de la classe 1** indique, parmi les acheteurs prédits comme défaillants par le modèle, combien sont réellement défaillants. Une precision élevée signifie que le modèle limite les fausses alertes. Autrement dit, lorsqu'il classe un acheteur comme risqué, il se trompe rarement. Cette métrique est utile lorsque l'on veut éviter de refuser inutilement des acheteurs solvables.

Le **recall de la classe 1**, (sensibilité), mesure la capacité du modèle à retrouver les acheteurs réellement défaillants. Il indique, parmi tous les acheteurs qui sont réellement en défaut, combien ont été correctement détectés.

Le **F1-score de la classe 1** combine la precision et le recall. Il permet donc d'évaluer l'équilibre entre la capacité du modèle à détecter les défauts et sa capacité à limiter les

fausses alertes. Cette métrique est pertinente lorsque l'on ne veut pas privilégier uniquement la précision ou uniquement le recall.

Le **ROC-AUC** mesure la capacité générale du modèle à distinguer les acheteurs défaillants des acheteurs non défaillants. Plus sa valeur est proche de **1**, plus le modèle sépare correctement les deux classes. À l'inverse, une valeur proche de **0,5** indique que le modèle ne fait pas mieux qu'un classement aléatoire.

Le **seuil standard de 0,50** est la limite utilisée par défaut pour transformer une probabilité en décision.

Enfin, la **matrice de confusion** permet d'analyser concrètement les erreurs du modèle. Elle distingue les acheteurs correctement classés comme non défaillants, les acheteurs correctement classés comme défaillants, ainsi que les erreurs de classification. Dans le contexte de l'assurance-crédit, les **faux négatifs** sont les erreurs les plus sensibles, car ils correspondent à des acheteurs réellement défaillants que le modèle classe comme non défaillants. Ce type d'erreur peut conduire à une mauvaise décision de couverture et augmenter l'exposition au risque.

4. Interprétation détaillée des résultats par modèle :

Afin d'évaluer la capacité des différents modèles à détecter les acheteurs défaillants, plusieurs méthodes de classification ont été testées.

4.1. Régression logistique :

a. Régression logistique simple :

Pour évaluer la performance de la régression logistique simple, plusieurs indicateurs ont été mobilisés, notamment l'accuracy, la précision, le recall, le F1-score et le ROC-AUC.

La matrice de confusion permet également d'observer plus concrètement les erreurs du modèle, en particulier les défauts non détectés.

Le résultat du modèle :

Figure 5: le résultat du modèle de régression logistique simple

Rapport de classification :				
	precision	recall	f1-score	support
0	0.76	0.86	0.80	43
1	0.45	0.29	0.36	17
accuracy			0.70	60
macro avg	0.60	0.58	0.58	60
weighted avg	0.67	0.70	0.68	60

Source : Élaboré par nous-mêmes selon le modèle

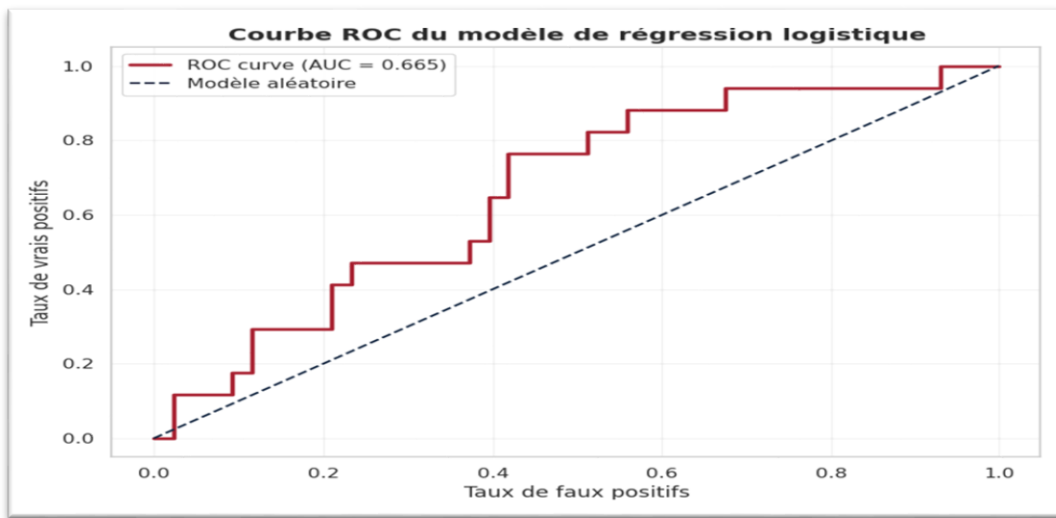
Commentaire :

Les résultats de ce modèle montrent une performance globale moyenne. Il obtient une **accuracy de 0,70**, ce qui signifie qu'il classe correctement **70 % des acheteurs** de la base test. Au premier lieu, ce résultat peut sembler acceptable. Cependant, cette valeur doit être interprétée avec prudence, car la base est déséquilibrée : les acheteurs non défaillants sont plus nombreux que les acheteurs défaillants. Ainsi, un modèle peut obtenir une accuracy correcte en prédisant surtout la classe majoritaire, sans bien détecter les défauts.

L'analyse de la **classe 1**, les acheteurs défaillants, montre une limite importante. Le modèle présente une **precision de 0.45**, un **recall de 0.29** et un **F1-score de 0.36**. Cela signifie qu'il détecte mal les défauts réels. Le modèle identifie correctement **37 non-défauts** et seulement **5 défauts**, mais il laisse passer **12 acheteurs réellement défaillants**.

La courbe ROC-AUC suivante permet d'évaluer le pouvoir discriminant global du modèle .

Figure 6: représente la courbe ROC-AUC de régression logistique



Source : Élaboré par nous-mêmes selon le modèle

Le **ROC-AUC de 0.6648** montre une capacité de discrimination moyenne. La courbe s'éloigne peu de la diagonale aléatoire, confirmant les limites discriminantes du modèle. La régression logistique reste donc utile comme modèle de référence, mais elle est insuffisante pour détecter efficacement les acheteurs risqués. Dans une logique d'assurance-crédit à l'exportation, les 12 défauts non détectés représentent un risque financier direct pour la compagnie, ce qui justifie la recherche d'un modèle plus sensible à la classe minoritaire. Cette limite peut s'expliquer par son caractère linéaire, qui ne permet pas toujours de capter les relations complexes entre les variables financières, pays, sectorielles et transactionnelles.

b. Régression logistique optimisée avec pondération des classes :

Afin de corriger le déséquilibre de classes nous utilisons l'option **class_weight = balanced**. Ce choix méthodologique permet de corriger partiellement le déséquilibre de la variable cible. Avec la pondération des classes, le modèle accorde davantage d'importance à la classe minoritaire (acheteurs défaillants).

Les résultats du modèle :

Figure 7: représente le résultat du modèle de régression logistique optimisé

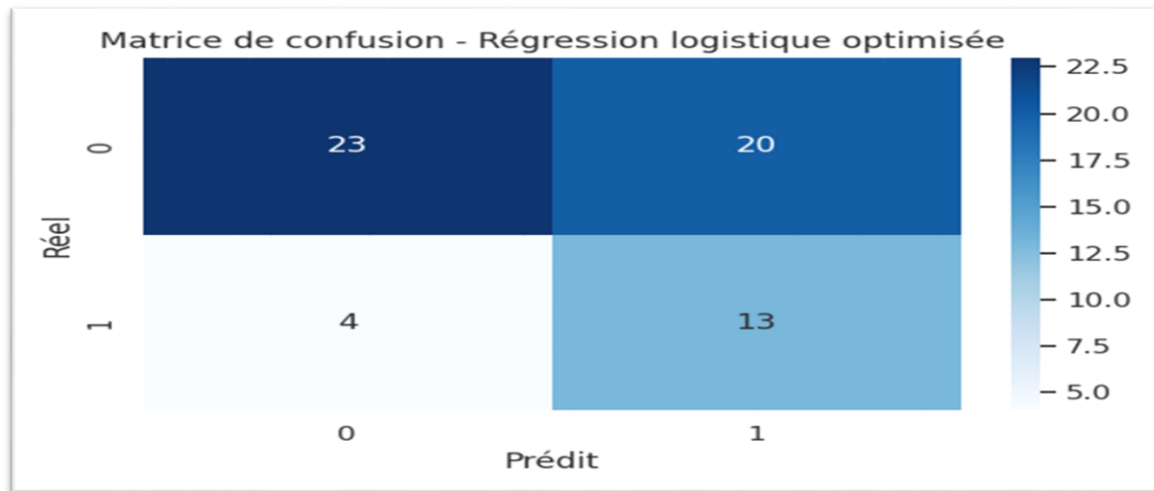
Rapport de classification :					
	precision	recall	f1-score	support	
0	0.85	0.53	0.66	43	
1	0.39	0.76	0.52	17	
accuracy			0.60	60	
macro avg	0.62	0.65	0.59	60	
weighted avg	0.72	0.60	0.62	60	

Source : Élaboré par nous-mêmes selon le modèle

Commentaire :

Les résultats de ce modèle montrent une performance différente de celle du modèle précédent. Il obtient une **accuracy de 0.60**, ce qui signifie qu'il classe correctement **60 % des acheteurs** de la base test. À première vue, cette performance globale semble plus faible que celle de la régression logistique simple. Cependant, cette baisse de l'accuracy doit être interprétée avec prudence, car l'objectif de ce modèle optimisé n'est pas seulement d'augmenter le nombre total de bonnes prédictions, mais surtout d'améliorer la détection des acheteurs défaillants. L'analyse de la **classe 1**, (les acheteurs défaillants), montre une amélioration importante. Le modèle présente une **precision d'environ 0.39**, un **recall de 0.76** et un **F1-score d'environ 0.52**. Cela signifie que le modèle arrive à détecter une grande partie des défauts réels. La matrice de confusion confirme :

Figure 8: représente la matrice de confusion du modèle de régression logistique optimisé



Source : Élaboré par nous-mêmes selon le modèle

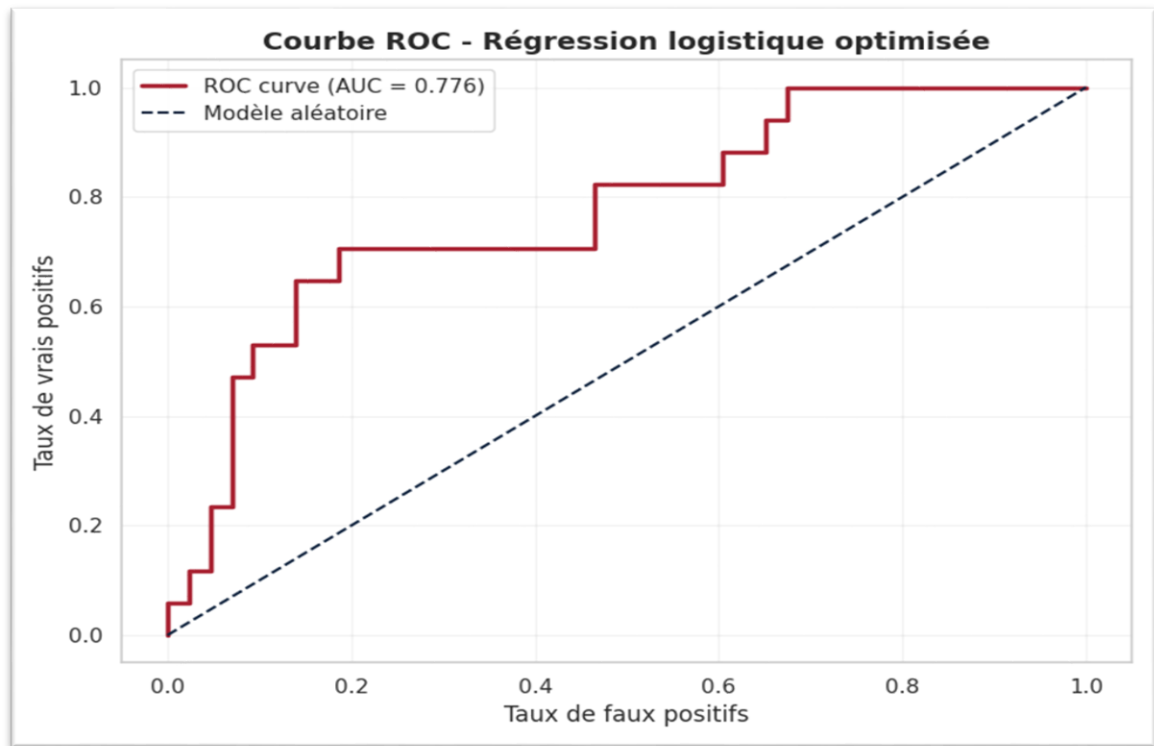
Le modèle identifie correctement **23 acheteurs non défaillants** et **13 acheteurs défaillants**, mais il classe aussi **20 acheteurs non défaillants comme défaillants**. En revanche, il ne laisse passer que

4 acheteurs réellement défaillants, contre **12** dans la régression logistique simple.

Ce résultat montre que la pondération des classes rend le modèle plus sensible à la classe minoritaire. C'est à dire le modèle devient plus prudent et détecte davantage les acheteurs à risque. Cette amélioration est importante dans une logique d'assurance-crédit, car les défauts non détectés peuvent représenter un risque financier élevé pour la compagnie. Cependant, cette prudence entraîne aussi une augmentation des fausses alertes, puisque certains acheteurs solvables sont classés comme risqués.

La courbe ROC-AUC montre ça :

Figure 9: Représente la courbe ROC-AUC de la régression logistique optimisée



Source : Élaboré par nous-mêmes selon le modèle

La courbe **ROC-AUC** confirme cette amélioration. Le modèle optimisé obtient un **ROC-AUC de 0.7756**, contre **0.6648** pour la régression logistique simple. Cela signifie que la régression logistique optimisée sépare mieux les acheteurs défaillants des acheteurs non défaillants. Ainsi, même si son accuracy est plus faible, ce modèle est plus intéressant pour la détection du défaut, car il réduit fortement le nombre d'acheteurs risqués non détectés.

c. Régression logistique avec SMOTE :

Une deuxième approche a été explorée pour traiter le déséquilibre des classes. Elle s'appelle **SMOTE (Synthetic Minority Over-sampling Technique)** est une méthode utilisée lorsque la base de données est déséquilibrée, le cas où l'une des classes est beaucoup moins représentée que l'autre.

Elle ne copie pas simplement les acheteurs défaillants existants, elle crée de nouveaux exemples proches des anciens, à partir de leurs caractéristiques.

Les résultats du modèle :

Figure 10: Représente la courbe ROC-AUC de la régression logistique optimisée

Rapport de classification :				
	precision	recall	f1-score	support
0	0.83	0.56	0.67	43
1	0.39	0.71	0.50	17
accuracy			0.60	60
macro avg	0.61	0.63	0.58	60
weighted avg	0.70	0.60	0.62	60

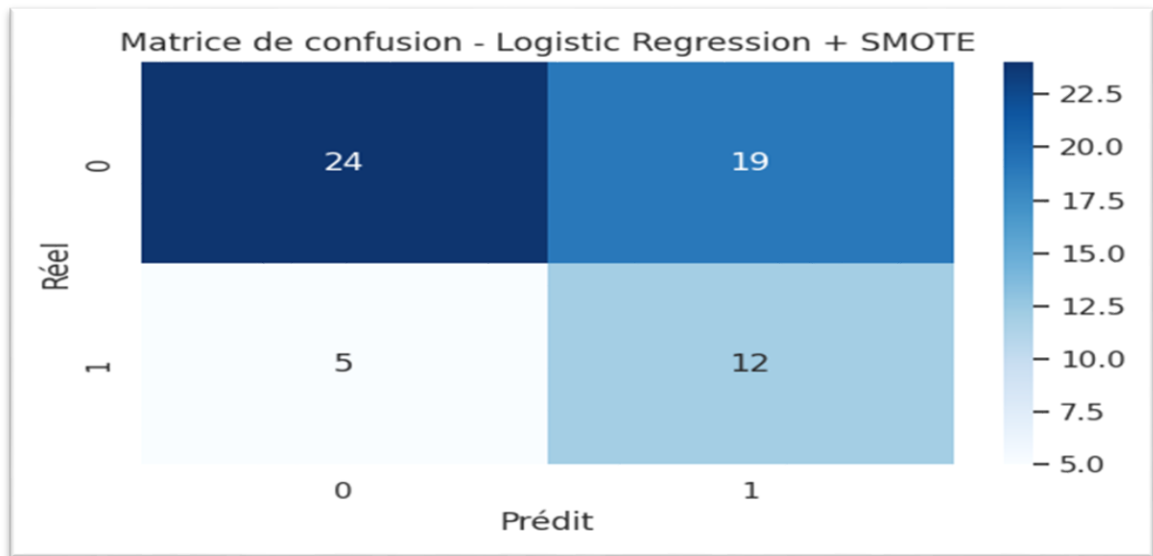
Source : Élaboré par nous-mêmes selon le modèle

Commentaire :

Les résultats de ce modèle montrent une performance globale moyenne. Il obtient une **accuracy de 0.60**, ce qui signifie qu'il classe correctement **60 % des acheteurs** de la base test. Comme pour la régression logistique pondérée, cette accuracy est inférieure à celle de la régression logistique simple. Cependant, cette baisse s'explique par le fait que le modèle cherche davantage à détecter les acheteurs défaillants, au lieu de privilégier uniquement la classe majoritaire.

L'analyse de la **classe 1**(acheteurs défaillants) montre une amélioration par rapport au modèle simple. Le modèle présente une **precision de 0.39**, un **recall de 0.71** et un **F1-score d'environ 0.50**. Cela signifie qu'il détecte **12 défauts sur 17**, ce qui est nettement meilleur que la régression logistique simple, qui ne détectait que 5 défauts. Par matrice de confusion :

Figure 11: Représente la matrice de confusion de la régression logistique +SMOTE

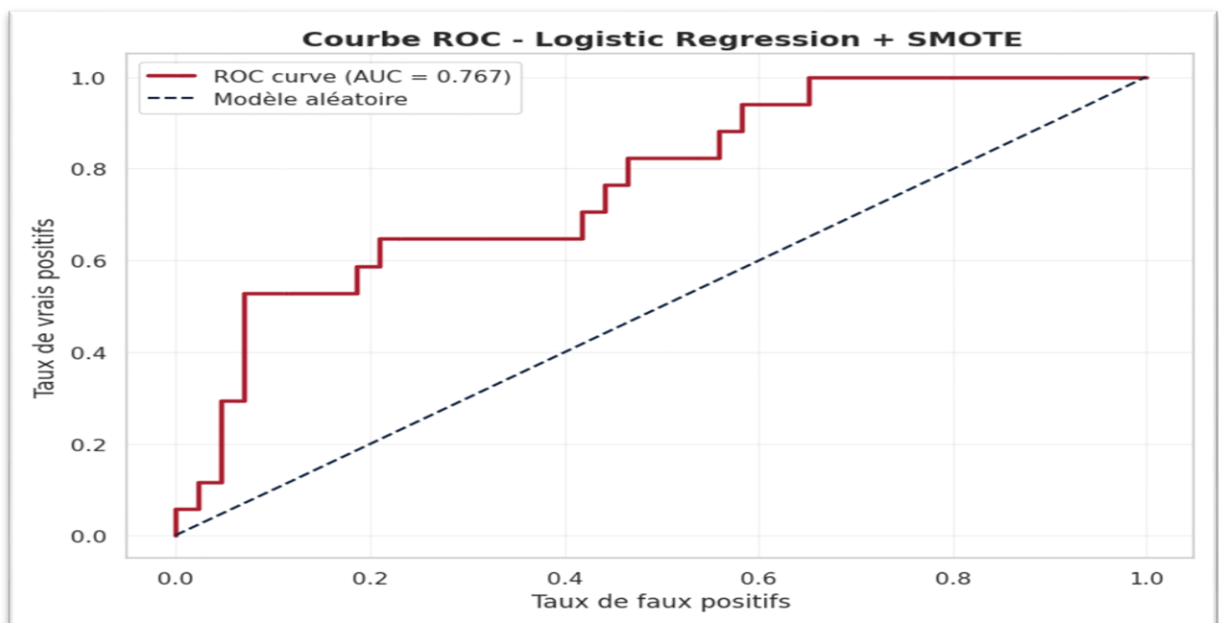


Source : Élaboré par nous-mêmes selon le modèle

Il est confirmé que : le modèle identifie correctement **24 acheteurs non défailants** et **12 acheteurs défailants**, mais il classe aussi **19 acheteurs non défailants comme défailants**. Il laisse également passer **5 acheteurs réellement défailants**.

La courbe ROC-AUC présentée :

Figure 12: Représente la courbe ROC-AUC de la régression logistique +SMOTE



Source : Élaboré par nous-mêmes selon le modèle

Cette représentation graphique **ROC-AUC de 0.7674** montre que le modèle possède une bonne capacité de discrimination, meilleure que celle de la régression logistique simple. Le recours au **SMOTE** a permis de mieux apprendre la classe minoritaire en créant artificiellement de nouveaux exemples d'acheteurs défaillants pendant l'apprentissage. Toutefois, ce modèle reste légèrement moins performant que la régression logistique optimisée avec pondération des classes, qui obtient un recall de 0.76 et un ROC-AUC de 0.7756.

Enfin la régression logistique avec SMOTE améliore la détection des défauts, mais elle génère aussi plusieurs fausses alertes. Elle confirme que le traitement du déséquilibre des classes est nécessaire, mais montre également que le SMOTE n'est pas automatiquement la meilleure solution.

Les résultats des trois variantes confirment que la régression logistique atteint un plafond de performance autour d'une AUC de 0.77, lié à sa frontière de décision linéaire. Afin de capturer les relations non-linéaires entre les variables financières, sectorielles et pays, nous avons recours dans la section suivante à des algorithmes d'apprentissage par ensemble, capables de modéliser des interactions plus complexes entre les facteurs de risque.

4.2. XGBoost :

a. XGBoost avec SMOTE :

Le modèle XGBoost a été testé en combinaison avec le SMOTE

Les résultats du modèle :

Figure 13: Représente les résultats de XGBoost +SMOTE

Rapport de classification :				
	precision	recall	f1-score	support
0	0.77	0.79	0.78	43
1	0.44	0.41	0.42	17
accuracy			0.68	60
macro avg	0.61	0.60	0.60	60
weighted avg	0.68	0.68	0.68	60

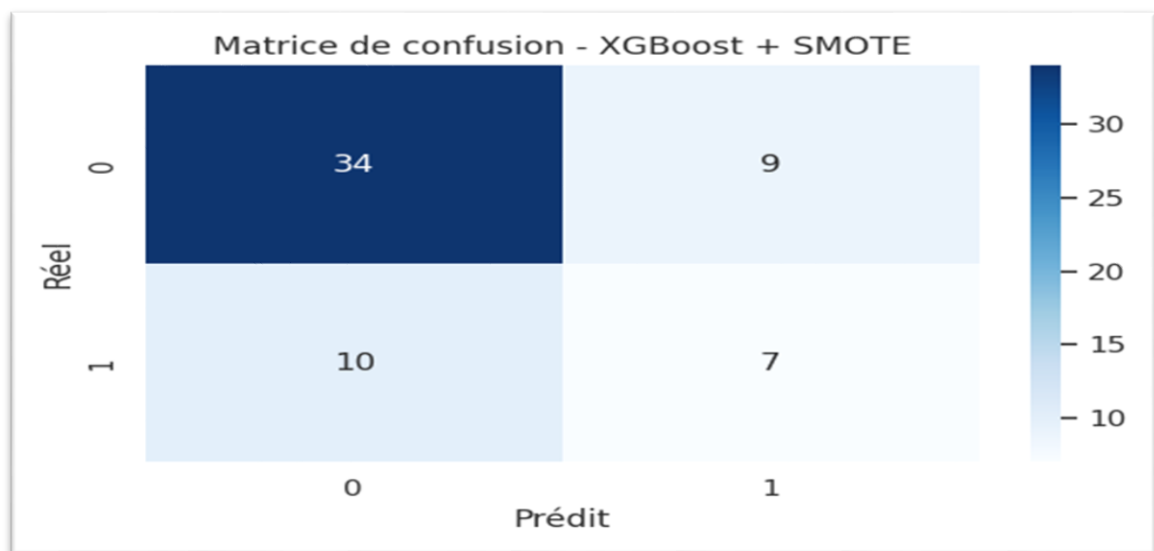
Source : Élaboré par nous-mêmes selon le modèle

Commentaire :

Les résultats montrent une performance globale relativement correcte. Le modèle obtient une **accuracy de 0.6833**, ce qui signifie qu'il classe correctement environ **68 % des acheteurs** de la base test.

Cependant, l'analyse de la **classe 1**, (les acheteurs défaillants) montre une limite importante. Le modèle présente un **recall de 0.41** et un **F1-score de 0.42**, ce qui signifie qu'il ne détecte pas suffisamment les défauts réels qui se trouve dans la matrice de confusion :

Figure 14: Représente la matrice de confusion de XGBoost + SMOTE



Source : Élaboré par nous-mêmes selon le modèle

Le résultat de cette matrice montre que le modèle identifie correctement **34 acheteurs non défaillants** et **7 acheteurs défaillants**, mais il laisse passer **10 acheteurs réellement défaillants** en les classant comme non défaillants.

Ainsi, même si l'accuracy semble acceptable, ce modèle n'est pas le plus adapté à l'objectif principal de l'étude, qui est de mieux repérer les acheteurs risqués. Dans la logique d'assurance-crédit, les **faux négatifs** sont particulièrement sensibles, car ils peuvent conduire à accepter des acheteurs présentant un risque réel de non-paiement.

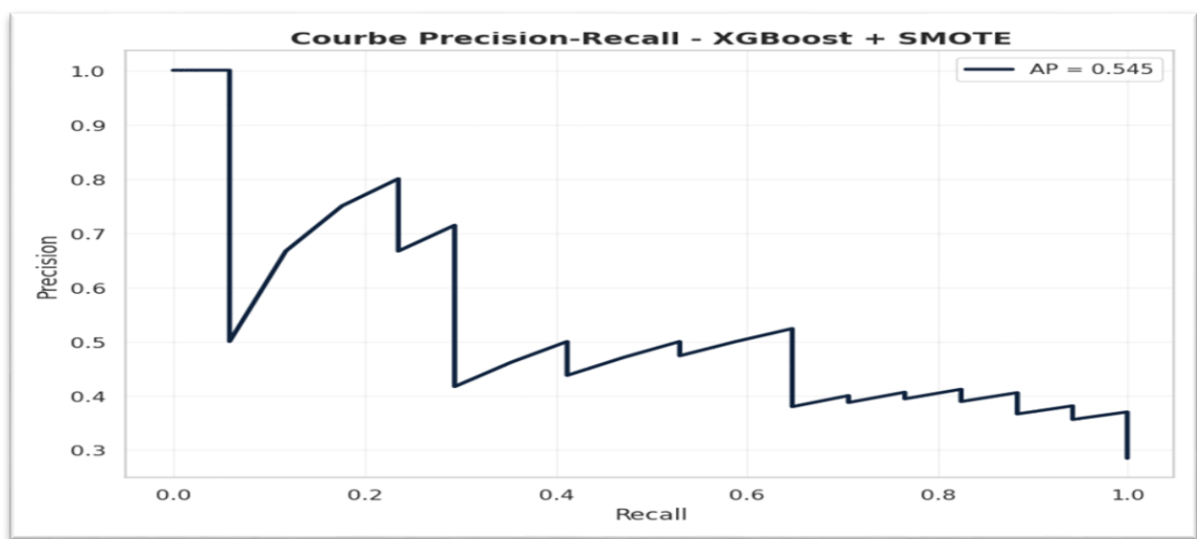
On analyse aussi par une courbe qui mixe le Recall avec precision (**La courbe Precision-Recall**) car la base est **déséquilibrée** : les acheteurs non défaillants sont plus nombreux que

les acheteurs défaillants. Dans ce cas, l'accuracy et même la courbe ROC ne suffisent pas toujours pour bien juger la capacité du modèle à détecter la classe **1 : défaut**.

Cette courbe sert donc à analyser plus précisément la performance du modèle sur la classe défaut. Elle montre le compromis entre deux objectifs :

- Le **recall** : détecter le maximum d'acheteurs réellement défaillants ;
- La **precision** : éviter de classer trop d'acheteurs sains comme défaillants.

Figure 15: Représente la courbe Precision-recall de XGBoost + SMOTE



Source : Élaboré par nous-mêmes selon le modèle

Le résultat obtenu donne un **Average Precision (AP) égale 0.545**, qui indique une performance moyenne dans la détection des défauts. Le modèle arrive à identifier une partie des acheteurs défaillants, mais il ne maintient pas une précision élevée lorsque le recall augmente.

On remarque également que la courbe est irrégulière, qui peut s'expliquer par la taille limitée de la base test, qui contient seulement **60 observations**, dont **17 défauts**. Dans ce cas, chaque bonne ou mauvaise prédiction peut fortement modifier la forme de la courbe.

Alors cette courbe confirme les résultats de la matrice de confusion, ce modèle a détecté 7 défauts sur 17. Même si le modèle présente une accuracy correcte, sa capacité à repérer les acheteurs réellement défaillants reste limitée. La courbe Precision-Recall montre ainsi que

ce modèle nécessite un ajustement du seuil de décision afin d'améliorer la détection de la classe défaut. C'est précisément l'objet de la variante suivante, dans laquelle le seuil de décision est abaissé à **0.2098** afin de maximiser la détection des acheteurs réellement défaillants.

b. XGBoost optimisé avec changement de seuil :

Le modèle de test est XGBoost mais avec le changement de seuil il était **0,5** il devient **0,2098**

Les résultats du modèle :

Figure 16: Représente le résultat de XGBoost +SMOTE avec changement de seuil

Rapport de classification :				
	precision	recall	f1-score	support
0	0.95	0.47	0.62	43
1	0.41	0.94	0.57	17
accuracy			0.60	60
macro avg	0.68	0.70	0.60	60
weighted avg	0.80	0.60	0.61	60

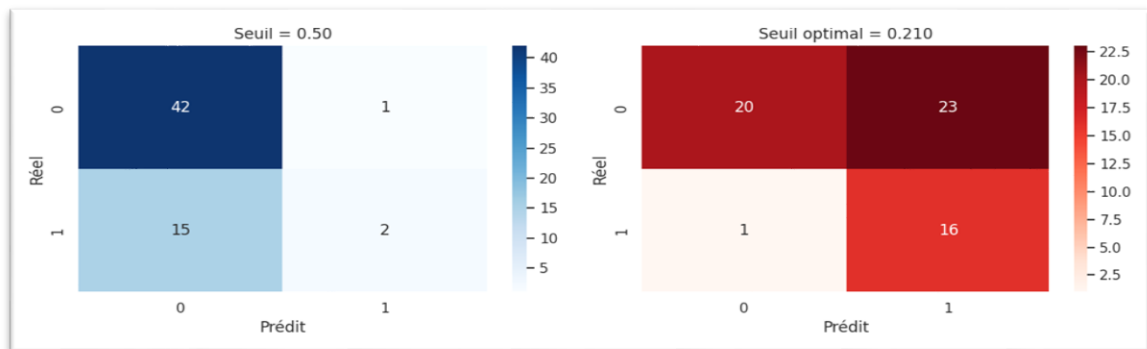
Source : Élaboré par nous-mêmes selon le modèle

Commentaire :

Ce modèle présente une bonne capacité globale de discrimination. Ce résultat montre que le modèle parvient globalement à distinguer les acheteurs défaillants des acheteurs non défaillants.

Selon la matrice de confusion :

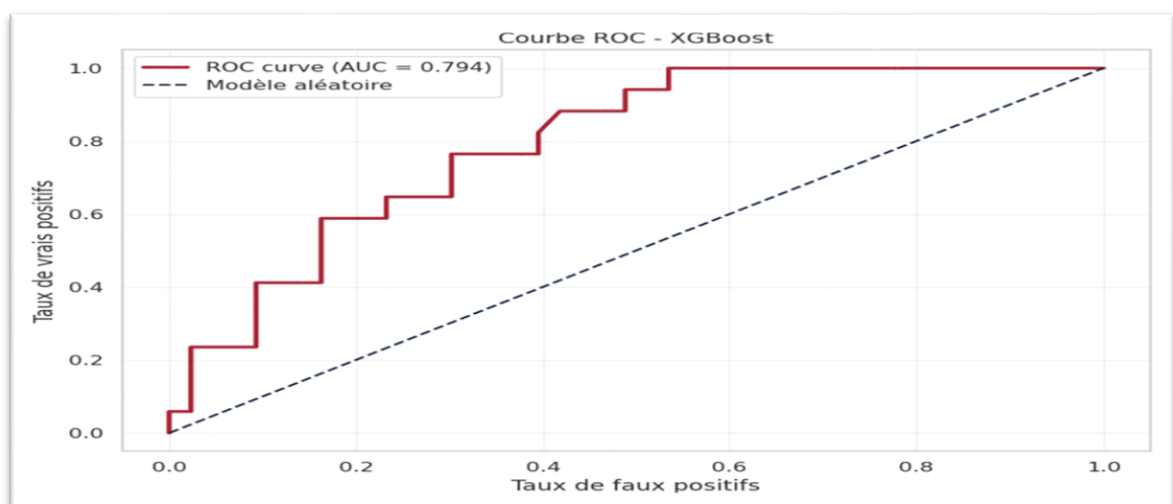
Figure 17: Représente la matrice de confusion avant et après le changement de seuil



Source : Élaboré par nous-mêmes selon le modèle

Après l’ajustement **du seuil à 0.2098**, les résultats changent nettement. Le modèle détecte désormais **16 défauts sur 17**, ce qui porte le **recall de la classe 1 à 0.9412**. Cette amélioration est très intéressante dans une logique de gestion du risque, car elle réduit fortement le nombre de défauts non détectés. Par contre, cette prudence entraîne une augmentation des fausses alertes : **23 acheteurs non défaillants sont classés comme défaillants**. Par conséquent, la **precision de la classe 1 baisse à 0.4103** et l’**accuracy diminue à 0.60**.

Figure 18 : Représente la matrice de confusion avant et après le changement de seuil



Source : Élaboré par nous-mêmes selon le modèle

La courbe ROC met en évidence la capacité discriminante globale du modèle. Une **AUC proche de 0.80** suggère que le modèle possède une capacité raisonnable à distinguer les observations positives des observations négatives.

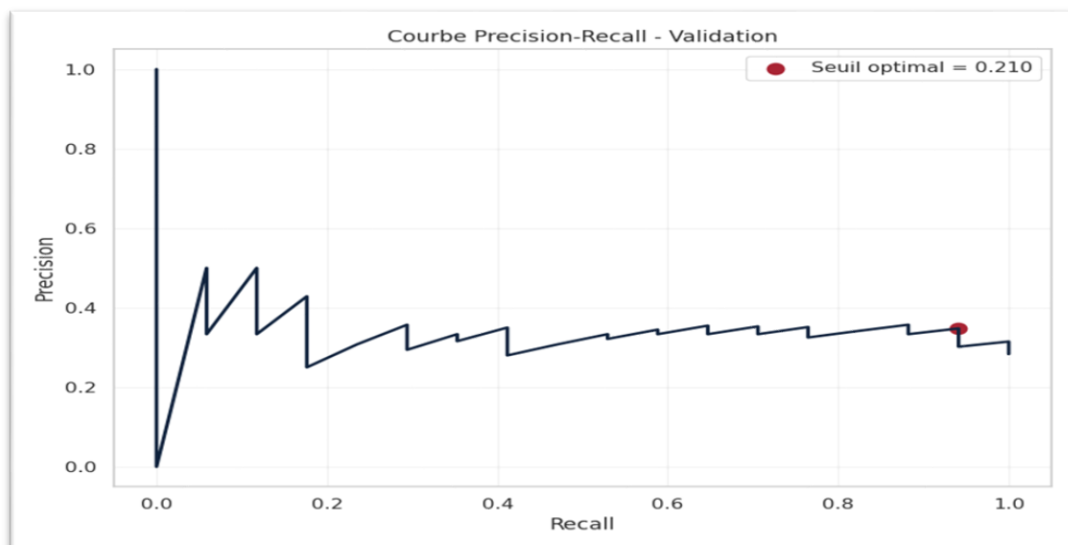
Néanmoins, l'écart avec une performance parfaite ($AUC = 1$) montre que le modèle conserve certaines limites de séparation entre les classes. Cela peut provenir :

- De la qualité des variables explicatives ;
- D'un déséquilibre de classes ;
- Ou bien de la volatilité des données financières.

La courbe ROC confirme que le modèle présente une capacité prédictive exploitable mais encore perfectible.

Aussi :

Figure 19: Représente la courbe precision-recall



Source : Élaboré par nous-mêmes selon le modèle

La courbe Precision-Recall donne une information importante dans les problèmes de classification déséquilibrée. Le seuil optimal identifié à **0.210** correspond à une zone où le recall est élevé, et la précision demeure relativement faible. Cela signifie qu'une proportion importante des alertes générées par le modèle est incorrecte.

D'un point de vue, cette situation est cohérente avec la matrice de confusion : le modèle détecte presque tous les cas positifs ; mais au prix d'un nombre important de faux positifs.

La courbe souligne donc la nécessité d'adapter le seuil de décision aux objectifs économiques réels de l'institution financière.

c. Optimisation du modèle XGBoost : réglage des hyperparamètres et ajustement du seuil de décision

Dans cette partie, nous avons appliqué un tuning des hyperparamètres afin de vérifier si un meilleur réglage du modèle XGBoost pouvait améliorer ses résultats. Cette étape consiste à tester plusieurs combinaisons de paramètres, pour l'objectif de trouver une configuration plus adaptée aux données et de limiter le risque de surapprentissage.

Le surapprentissage apparaît lorsque le modèle apprend trop les détails propres à la base d'entraînement. Dans ce cas, il peut donner de bons résultats sur les données utilisées pour l'apprentissage, mais devenir moins performant sur de nouvelles données. Le tuning permet donc de rendre le modèle plus stable et plus fiable.

Les résultats du modèle :

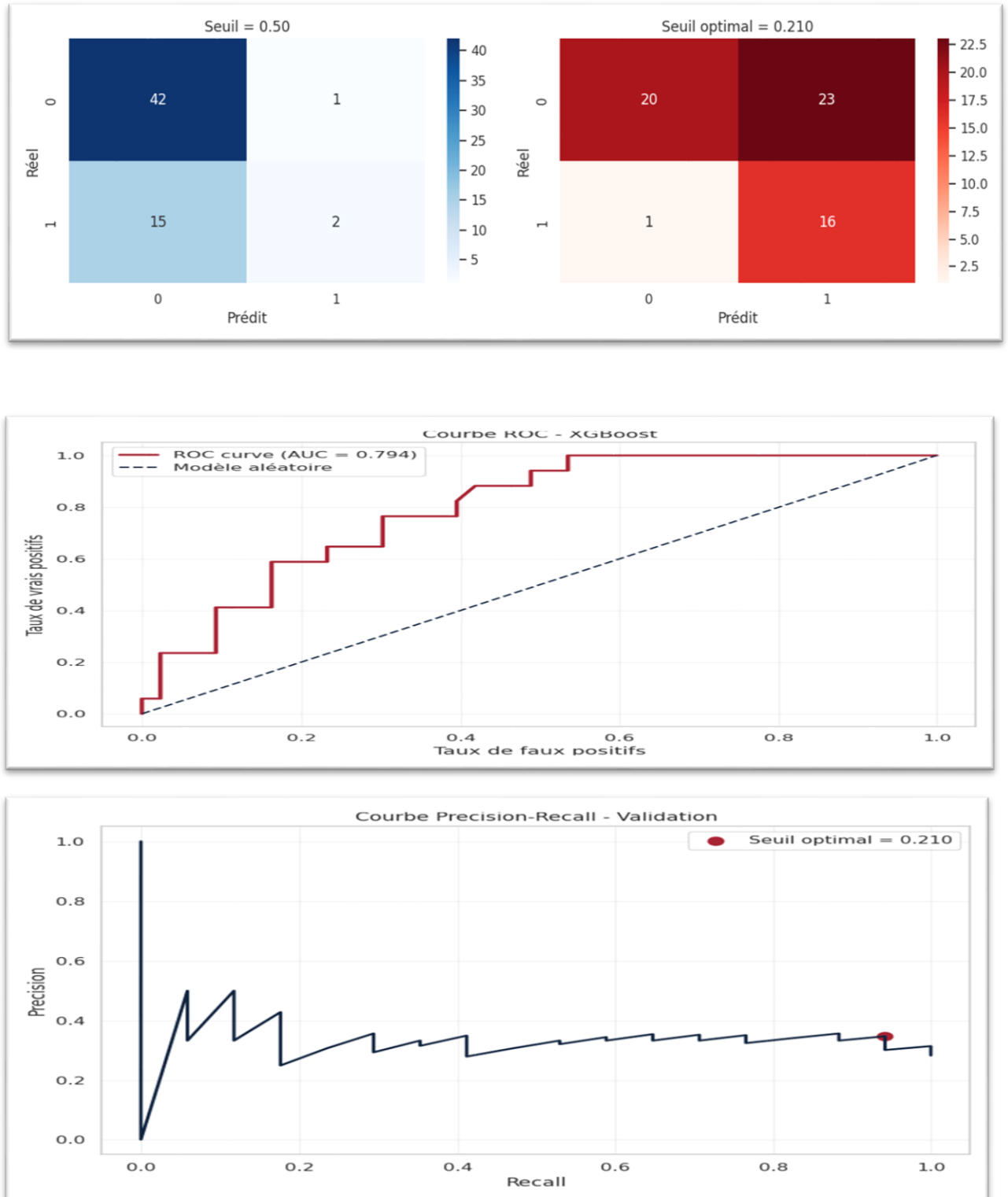
Dans ce cas, les résultats obtenus après le tuning restent très proches de ceux obtenus avant l'optimisation. Cela signifie que le réglage des hyperparamètres n'a pas apporté une bonne amélioration des performances. Toutefois, cette étape reste utile, car elle confirme que le modèle XGBoost est relativement robuste et que ses résultats ne dépendent pas fortement d'un choix arbitraire de paramètres.

L'analyse montre que le point le plus important n'est pas le tuning lui-même, mais plutôt le choix du seuil de décision. Avec le seuil classique de 0.50, le modèle classe correctement la majorité des entreprises non défaillantes. Toutefois, il détecte très peu d'entreprises réellement en défaut. Cela pose un problème si le modèle ne repère pas une entreprise risquée, cela peut entraîner des pertes financières importantes.

En abaissant le seuil à **0.21 (le seuil optimal déterminé par optimisation de la courbe Précision-Rappel)**, le modèle devient plus sensible au risque de défaut. Il détecte beaucoup

plus d'entreprises défaillantes, mais il produit aussi davantage de fausses alertes. Ce résultat reste logique : plus le modèle est prudent, plus il signale d'entreprises comme risquées.

Figure 20: Représente les graphes après le tuning



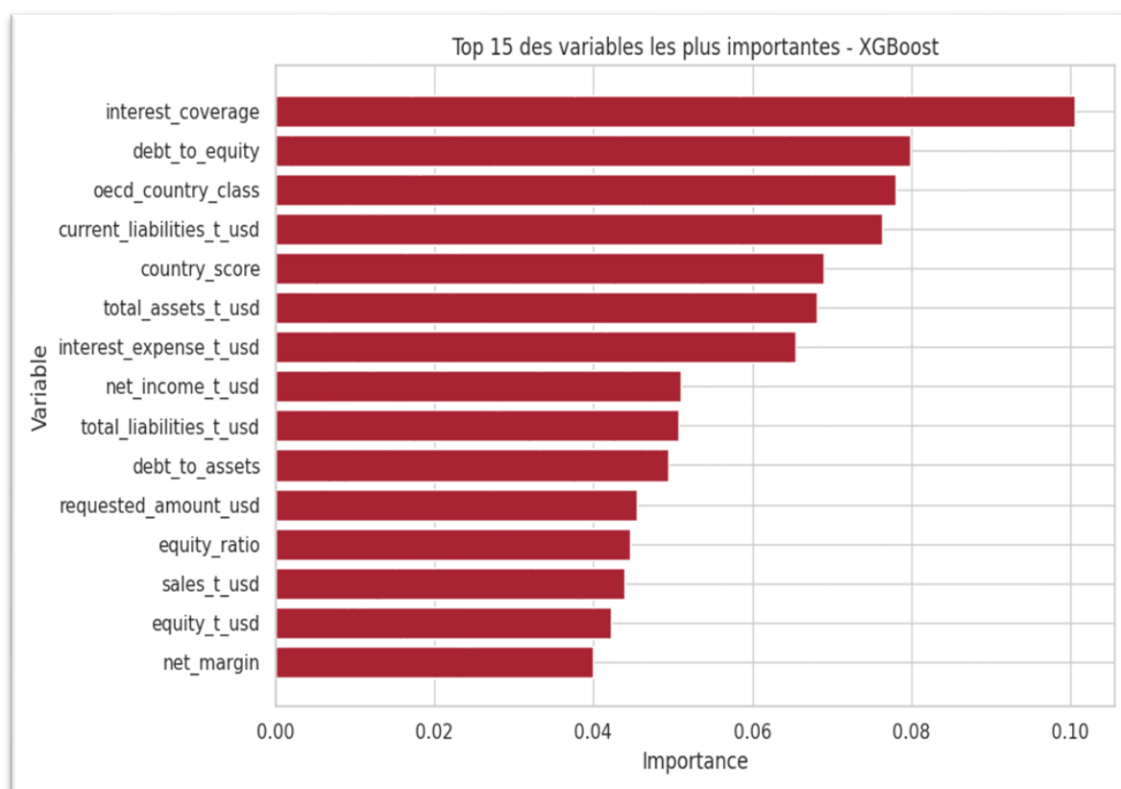
Source : Élaboré par nous-mêmes selon le modèle

La courbe ROC confirme que le modèle a une capacité correcte à distinguer les acheteurs défaillants des acheteurs non défaillants. Avec un AUC de 0.794, le modèle fait mieux qu'un classement aléatoire et présente une performance satisfaisante, mais il n'est pas parfait.

Lorsque le seuil diminue, le recall augmente, ce qui veut dire que le modèle détecte plus de défauts. En revanche, la précision diminue, car certaines entreprises non défaillantes sont classées comme risquées.

Ce modèle nous aide à comprendre qui sont les variables les plus influencé par le modèle lorsqu'il a prédit le défaut des acheteurs :

Figure 21: Représente les variables les plus importantes dans le modèle XGBoost



Source : Élaboré par nous-mêmes selon le modèle

Le graphe montre les variables qui ont une influence directe sur le modèle XGBoost dans la prédiction du défaut. Cette importance est mesurée par le nombre de fois qu'une variable est utilisée dans les arbres de décision du modèle XGBoost, révèle une cohérence forte avec les fondements théoriques de l'analyse du risque de crédit.

On remarque que **interest_coverage** est la variable la plus importante. C'est-à-dire que le modèle accorde beaucoup d'attention à la capacité de l'entreprise à payer ses charges d'intérêts par le résultat d'exploitation de l'exercice. Par logique, une entreprise qui n'a pas la capacité à couvrir ses intérêts présente un risque de défaut plus élevé.

Aussi, les variables liées à l'endettement, comme **debt_to_equity**, **debt_to_assets**, **total_liabilities_t_usd** et **interest_expense_t_usd**, jouent un rôle important. Elles montrent que le niveau de la dette et le poids des charges financières sont des éléments essentiels dans l'évaluation du risque.

Également le **oecd_country_class** et **country_score** sont parmi les variables importantes, qui indique que le risque ne dépend pas seulement de la situation financière de l'entreprise, mais aussi de son environnement pays.

Enfin, le modèle s'appuie sur la capacité de remboursement, l'endettement, la solidité financière et le risque pays pour distinguer les acheteurs défaillants, des acheteurs non défaillants. Ces résultats encourageants nous amènent à tester dans la section suivante le modèle LightGBM, qui repose sur une architecture de gradient boosting optimisée, reconnue dans la littérature pour son efficacité sur des données tabulaires de petite à moyenne taille, précisément le contexte de cette étude.

4.3. LightGBM :

La même démarche d'optimisation appliquée au modèle XGBoost a été reproduite pour le modèle LightGBM, notamment l'ajustement du seuil de décision par maximisation du score F2 sur l'ensemble de validation.

Les résultats obtenus :

Figure 22: Représente les résultats obtenus du modèle LightGBM

Rapport de classification :				
	precision	recall	f1-score	support
0	0.88	0.53	0.67	43
1	0.41	0.82	0.55	17
accuracy			0.62	60
macro avg	0.65	0.68	0.61	60
weighted avg	0.75	0.62	0.63	60

Source : Élaboré par nous-mêmes selon le modèle

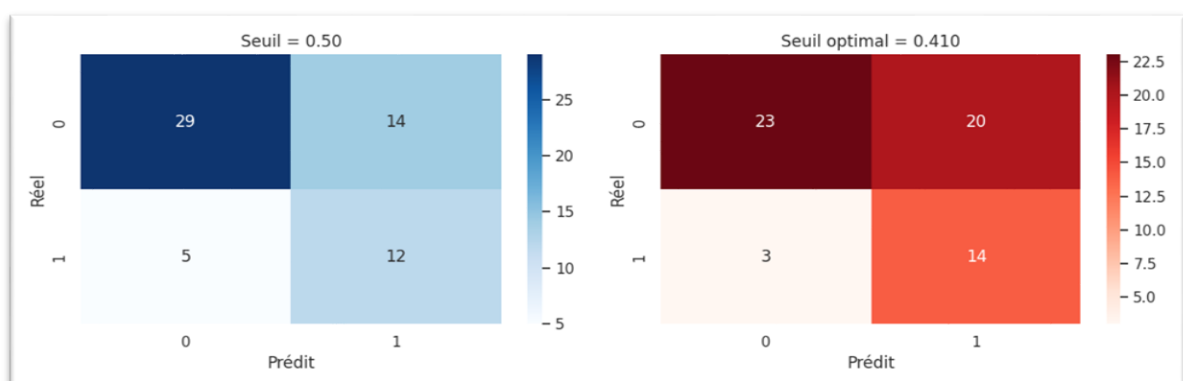
Commentaire :

Après l’ajustement de seuil optimal en 0.41 on a obtenu ce résultat : Pour la classe majoritaire (Les acheteurs non-défaillants), le modèle affiche une precision élevée de 0.88, en indiquant que la grande majorité des prédictions positives pour cette classe sont correctes. Avec un recall de 0.53, qui traduit une tendance du modèle à sous-estimer les vrais négatifs, classant incorrectement près de la moitié.

Pour la classe minoritaire (Les acheteurs défaillants), le recall atteint 0.82, ce qui signifie que le modèle détecte 82 % des cas de défaut réels. Ce résultat est particulièrement pertinent dans le contexte du risque de crédit. La precision relativement faible de 0.41 pour cette classe reflète le déséquilibre de l'échantillon.

D’où vient la matrice de confusion comme suit :

Figure 23: Représente la matrice de confusion avant et après l’optimisation en LightGBM

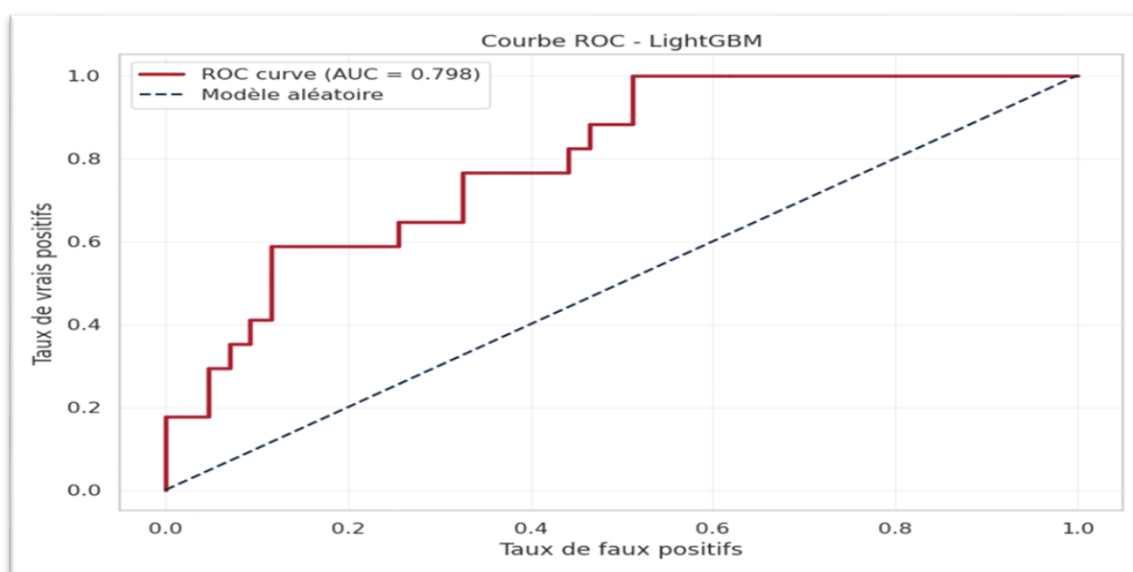


Source : Élaboré par nous-mêmes selon le modèle

Au seuil de **0.50**, le modèle génère 14 faux positifs et 5 faux négatifs, pour un recall de 0.706 sur la classe de défaut. L'abaissement du seuil à **0.410** permet de réduire les faux négatifs à 3, portant le recall à 0.82, au prix d'une augmentation des faux positifs à 20. La non-détection d'un défaut engendre des pertes financières nettement supérieures au coût d'une vérification approfondie d'un dossier sain.

La courbe ROC-AUC :

Figure 24: Représente la courbe ROC-AUC en LightGBM

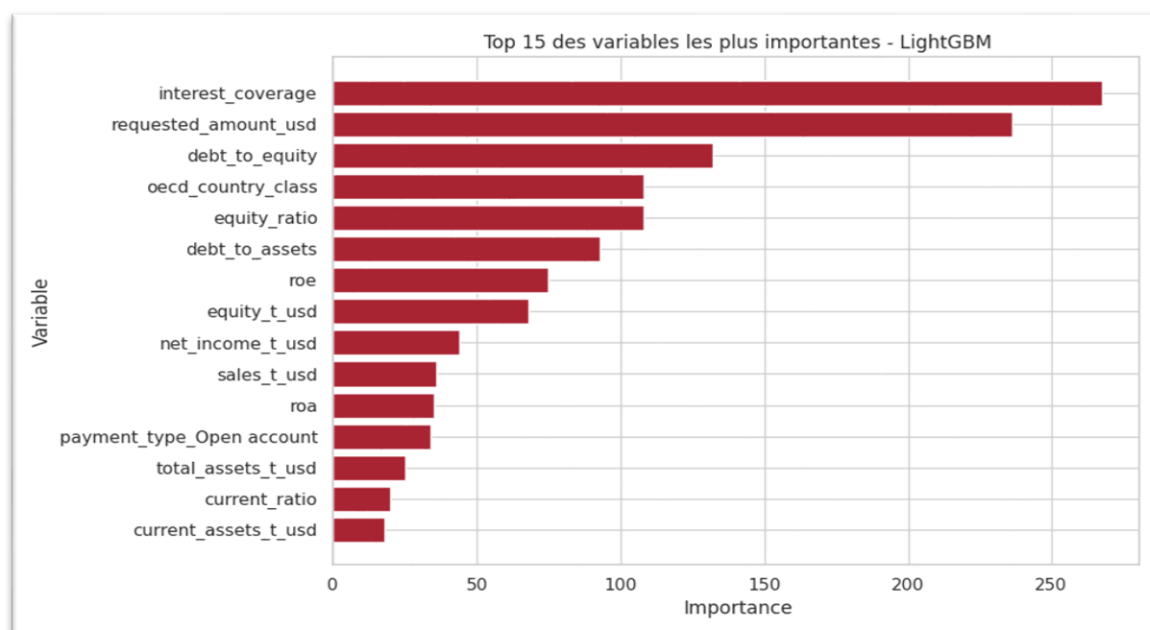


Source : Élaboré par nous-mêmes selon le modèle

Elle montre une progression nettement au-dessus de la diagonale de modèle aléatoire, elle confirme la capacité du modèle à discriminer les deux classes avec une **AUC = 0.798** se situe dans une plage généralement qualifiée **bonne** ce qui signifie la capacité significative du modèle à distinguer les entreprises en défaut des acheteurs solvables. La forme en escalier est une caractéristique des petits échantillons.

Comme le modèle de XGBoost, ce modèle a donné une importance aux variables

Figure 25: Représente les variables importantes en LightGBM



Source : Élaboré par nous-mêmes selon le modèle

En première position, la variable **interest_coverage** s'impose comme le prédicteur le plus discriminant, mesurant directement la capacité de l'acheteur à honorer ses dettes à partir de son résultat d'exploitation.

En deuxième position, `requested_amount_usd` reflète l'exposition directe de l'établissement de crédit au risque de contrepartie : des montants plus élevés sont naturellement associés à des niveaux de risque plus importants. Les ratios de structure financière `debt_to_equity`, `equity_ratio` et `debt_to_assets` — confirment, en troisième et quatrième positions, leur rôle dominant dans l'évaluation de la solvabilité, en mesurant le degré de levier financier de l'entreprise.

Ainsi, la présence de la variable `oecd_country_class` parmi les cinq premières variables souligne l'importance du risque pays dans le contexte du financement du commerce international.

Enfin, la variable `payment_type_Open_account`, issue de l'encodage des modalités de paiement, indique que le type d'instrument de règlement retenu constitue également un signal pertinent du profil de risque de la transaction.

En conclusion on peut dire que le modèle est solide et stable. Il repère bien les entreprises acheteurs sains et en ajustant le seuil, peut devenir plus vigilant pour détecter les acheteurs risqués, ce qui est crucial dans la gestion financière. Les indicateurs comme la couverture d'intérêt et le montant demandé sont essentiels pour comprendre ses décisions.

5. Test des hypothèses :

5.1. HYPOTHÈSE 1 : VALIDATION DES INDICATEURS FINANCIERS

H1 : Les indicateurs financiers influencent directement l'évaluation de défaut des acheteurs étrangers.

a. Résultats empiriques

• Importance des variables dans XGBoost (Figure 21)

L'analyse de l'importance des variables pour le modèle XGBoost révèle que les indicateurs financiers dominant largement les variables explicatives :

- **Interest_coverage** : la variable plus importante qui se positionne dans le premier Rang en termes d'importance. Un acheteur étranger incapable de couvrir ses intérêts par son EBIT présente un risque de défaut plus élevé
- **Debt_to_equity** : Parmi les tops 5 variables les plus importantes. Il mesure le levier financier.
- **Total_liabilities_t_usd** : Parmi les variables influentes, qui mesure l'endettement absolu.
- **Interest_expense_t_usd** : il mesure le coût des dettes, et Reflète les charges financières

Importance des variables dans LightGBM (Figure 25)

Le modèle LightGBM confirme et renforce ces résultats avec un classement similaire :

- **Interest coverage**: Position 1
- **Requested_amount_usd** : Position 2
- **Debt_to_equity, equity_ratio, debt_to_assets**: Positions 3-5

Alors les deux modèles d'ensemble (XGBoost et LightGBM) s'accordent sur l'importance dominante des indicateurs financiers, ce qui renforce la fiabilité du classement.

- **Performance discriminante des modèles basés sur indicateurs financiers**

La capacité des 2 modèles à distinguer les défauts fournit une preuve indirecte :

- **ROC-AUC avant tuning : 0.794** (Figure 20)
- **ROC-AUC LightGBM: 0.798** (Figure 24)
- **Seuil de 0.50: Recall = 0.706, Accuracy = 0.6833** (Figure 13)

Une AUC de 0.79-0.80 signifie que le modèle, construit principalement avec des variables financières, possède une **capacité de discrimination bonne**, meilleure qu'un classement aléatoire (AUC = 0.50).

- **Résultats comparatifs avec la Régression logistique :**

La régression logistique, modèle linéaire basé uniquement sur des combinaisons linéaires d'indicateurs financiers et autres, atteint : **AUC = 0.7756** (Figure 9) et un **Recall classe 1 = 0.76** (avec pondération des classes). Bien que ce résultat soit inférieur aux modèles d'ensemble, il démontre que les indicateurs financiers seuls possèdent un pouvoir discriminant significatif.

b. Conclusion sur H1 : VALIDÉE

Les résultats obtenus valident fortement l'hypothèse H1. Les cinq variables les plus importantes des modèles d'ensemble sont exclusivement des indicateurs financiers, et la régression logistique atteint une AUC de 0,77 avec ces seules variables. L'apport du XGBoost (AUC 0,79–0,80) reste marginal, confirmant que les indicateurs financiers constituent le socle prédictif dominant. Dans la lignée des travaux de Beaver, Altman et Ohlson, ces résultats rappellent que le machine learning prolonge l'analyse financière classique sans la remplacer, et que les ratios financiers doivent demeurer au cœur de toute évaluation du risque en assurance-crédit à l'exportation.

Statut : HYPOTHÈSE FORTEMENT VALIDÉE

5.2. HYPOTHÈSE 2 : VALIDATION DES DONNÉES MACROÉCONOMIQUES ET RISQUE PAYS

H2 : Les données macroéconomiques et le risque pays influencent significativement l'évaluation de défaut des acheteurs étrangers

a. Résultats empiriques

- **Présence des variables pays dans le top 10 d'importance**

Selon les 2 figures : Figure 21 (XGBoost) et Figure 25 (LightGBM) :

- **Oecd_country_class** : figure parmi les **5 variables les plus importantes**.
- **Country_score** : variable explicite du modèle, intégrée dans l'évaluation du risque pays.

Elles sont classées parmi les meilleures variables pour l'évaluation. Dans XGBoost l'oecd_country_class apparaît dans le top 10. Et dans LightGBM : est mentionnée parmi les variables pertinentes.

- **Contribution et Influence des Variables Pays :**

Dans le Dictionnaire des variables (Tableau 6) : nous avons utilisé la **classification OCDE (0-7)** d'une manière explicite dans la base de données. Par l'inspiration directe de « L'OCDE produit annuellement une classification des risques pays en huit catégories (0 à 7), qui sert de référence internationale pour les assureurs de crédit à l'exportation. »

Bien que les variables financières dominant individuellement, l'absence des variables pays dégrade significativement la performance, soulignant leur rôle contextuel.

Comme Matrice de confusion LightGBM (figure 23) Avec un seuil ajusté à 0.41 le **Recall = 0.82** pour la classe défaut. Nous comparons à la régression logistique simple (Recall = 0.29), ceci suggère que les variables non linéairement liées aux indicateurs financiers notamment le risque pays apportent une information supplémentaire dans l'évaluation.

- **Limitation observable : impact modéré dans les modèles**

Les variables pays n'apparaissent pas dans le **top 3** des variables importantes comme les variables financières qui restent dominantes.

D'où les variables pays jouent un **rôle contextuel** ou de multiplicateur de risque, mais ne sont pas des déterminants uniques.

Par contre, sur les 29 variables explicatives, leur présence dans le top 10 montre une **surpondération significative**.

b. Conclusion sur H2 : Partiellement validée

Les résultats empiriques confirment **partiellement** l'importance des variables pays dans la prédiction du défaut. Les variables **oecd_country_class** et **country_score** figurent en effet parmi les dix variables les plus influentes (Figures 21 et 25), la base de données intégrant explicitement le risque pays et les modèles XGBoost et LightGBM exploitant ces informations dans leurs arbres de décision. De plus, la classification OCDE (0-7) est directement implémentée dans la base, renforçant la cohérence avec les pratiques standards. Cependant, ces variables sont surpassées par les indicateurs financiers et, bien qu'elles améliorent la performance globale, elles ne constituent pas des déterminants principaux. Cette influence modérée est confirmée par le Recall limité pour la classe défaut avec le seuil standard (0.71), soulignant que le risque pays agit plutôt comme un facteur contextuel amplificateur que comme un prédicteur unique.

Statut : HYPOTHÈSE PARTIELLEMENT VALIDÉE

5.3. HYPOTHÈSE 3 : VALIDATION DES DONNÉES TRANSACTIONNELLES

H3 : Les données transactionnelles contribuent significativement à la prédiction du défaut acheteur

a. Résultats empiriques

Les Variables transactionnelles dans la base (Tableau 6 et 7) sont :
credit_term_months - payment_type : number_guarantees guarantee_type et
requested_amount_usd

- **Performance des variables transactionnelles dans XGBoost (Figure 21)**

L'analyse de l'importance montre que **requested_amount_usd** est une **variable très importante** car cette dernière rivalise avec les ratios financiers classiques et surpasse certaines variables financières moins critiques qui complète logiquement la couverture d'intérêt (interest_coverage)

- **Importance spécifique dans LightGBM (Figure 25)**

Dans le commentaire : « Après l'ajustement de seuil optimal en 0.41... la variable **payment_type_Open_account** issue de l'encodage des modalités de paiement, indique que le type d'instrument de règlement retenu constitue également un signal pertinent du profil de risque »

Ce passage confirme explicitement que : **payment_type** est parmi les variables pertinentes identifiées par LightGBM qui fournit un **signal prédictif du risque**.

- **Performance comparative : Modèle avec vs. Sans données transactionnelles**

La base originale comprend les variables transactionnelles. Les modèles XGBoost et LightGBM utilisent ces variables. L'AUC observée est : - XGBoost : 0.794 - LightGBM : 0.798 - Régression logistique simple : 0.6648.

Nous remarquons une amélioration significative du modèle XGBoost par rapport à la régression logistique simple (+ 0.13 en AUC) peut s'attribuer partiellement à la capacité des modèles d'ensemble à exploiter les relations non linéaires dans les variables transactionnelles.

b. Conclusion sur H3 : validée

Les variables transactionnelles (requested_amount_usd, payment_type) apparaissent dans le top 5-20 variables d'importance pour XGBoost et LightGBM. LightGBM reconnaît explicitement que payment_type_Open_account constitue un « signal pertinent du profil de risque ». Par contre l'amélioration de performance des modèles d'ensemble par rapport à la régression logistique simple dépend partiellement de l'exploitation des relations non linéaires dans ces variables.

Ainsi, les variables transactionnelles ne dominent pas le classement d'importance, mais elles jouent un rôle complémentaire aux variables financières.

Les **conditions de la transaction** (montant, durée, type de paiement) influencent effectivement le profil de risque. Si un acheteur avec bons indicateurs financiers mais conditions de transaction défavorables verra son risque augmenter. Alors, l'approche multidimensionnelle de l'évaluation du risque est **scientifiquement justifiée**.

Statut : HYPOTHÈSE VALIDÉE

5.4. HYPOTHÈSE 4 : PERFORMANCE COMPARATIVE DES MODÈLES DE MACHINE LEARNING

H 4 : Les modèles de machine Learning, notamment XGBoost et LightGBM, présentent une meilleure performance prédictive que la régression logistique.

- **Organisation des résultats**

Les résultats sont indiqués dans ce petit tableau selon les figures

Tableau 9: Représente Comparaison des modèles prédictifs pour la détection des défauts des acheteurs étrangers

Modèle	ROC-AUC	Recall Classe 1	Faux négatifs	Jugement
Régression logistique simple	0.665	0.29	12	Trop faible - manque la majorité des défauts
Régression logistique pondérée	0.776	0.76	4	Correct - amélioration notable mais perfectible
Régression logistique + SMOTE	0.767	0.71	5	Correct - légèrement moins performante que pondérée
XGBoost seuil 0.50	0.79	0.41	10	Moyen - seuil standard limite la détection
XGBoost optimisé seuil 0.21	0.80	0.94	1	Excellent - détecte presque tous les défauts
LightGBM seuil optimisé 0.41	0.798	0.82	3	Très bon : robuste et fiable pour le défaut

Source : Elabore par nous-mêmes selon les résultats

- **Analyse critique : condition de succès du machine learning**

Observation clé :

Les modèles **XGBoost** et **LightGBM** surpassent la régression logistique **uniquement après ajustement du seuil de décision.**

➤ **Seuil standard (0.50) :**

XGBoost n'améliore pas significativement la détection des défauts (recall = 0.41 vs 0.76 pour la régression logistique pondérée). L'accuracy plus élevée peut donner une

impression trompeuse de performance, masquant la faible capacité à identifier les acheteurs à risque.

➤ **Seuil optimisé :**

Le recall augmente de manière spectaculaire, détectant presque tous les défauts.

Le compromis est acceptable : une légère augmentation des fausses alertes est compensée par une réduction drastique des défauts non détectés.

Alors L'avantage réel des modèles de machine learning ne réside pas seulement dans l'augmentation de l'AUC, mais dans la **flexibilité du seuil de décision**, qui permet d'adapter le modèle aux objectifs métier et aux contraintes de risque.

b. Conclusion sur H4 : Validée

Les résultats empiriques invitent à nuancer le discours dominant sur la supériorité systématique des algorithmes de Machine Learning. En classification de défaut, cette supériorité est **conditionnelle à l'ajustement du seuil de décision**.

Avec un seuil standard de 0.50, l'avantage des modèles d'ensemble peut s'avérer marginal, voire absent : XGBoost affiche dans ce cas un recall inférieur à celui d'une régression logistique pondérée, malgré un meilleur ROC-AUC. C'est donc moins la complexité algorithmique que la capacité à paramétrer finement le seuil de classification qui constitue l'atout décisif du ML. Un modèle sophistiqué mal calibré reste un modèle sous-exploité.

• **Implications managériales**

L'adoption de XGBoost ou LightGBM en assurance-crédit à l'exportation est scientifiquement justifiée, à condition de respecter trois principes opérationnels :

Premièrement, le seuil de décision ne doit pas être traité comme un paramètre technique par défaut, mais comme une variable de pilotage stratégique. Son calibrage doit refléter explicitement les objectifs métier : une tolérance plus grande au risque justifie un seuil conservateur orienté rappel ; une politique de refus préventif appelle un seuil plus sélectif.

Deuxièmement, la régression logistique ne doit pas être écartée. Elle demeure un outil de référence indispensable.

Troisièmement, l'évaluation d'un modèle de crédit ne peut se réduire au ROC-AUC. La performance opérationnelle doit être mesurée au seuil effectivement déployé, en intégrant la structure asymétrique des coûts d'erreur propre à chaque portefeuille.

En définitive, le Machine Learning offre une solution clé en main qu'un espace de décision élargie dont la valeur dépend entièrement de la qualité du pilotage humain qui l'accompagne.

Statut : HYPOTHÈSE VALIDÉE

5.5. HYPOTHÈSE 5 : EXPLICABILITÉ DES MODÈLES PRÉDICTIFS

H5 : L'utilisation des méthodes d'explicabilité permet de mieux comprendre les facteurs déterminants du défaut acheteur.

a. Résultats empiriques

- **Importance des variables — XGBoost (Figure 21)**

L'analyse de l'importance des variables révèle le classement suivant :

1. interest_coverage (couverture d'intérêts)
2. debt_to_equity (ratio dettes / capitaux propres)
3. debt_to_assets (ratio dettes / actifs)
4. total_liabilities_t_usd (dettes totales)
5. interest_expense_t_usd (charges d'intérêts)
6. oecd_country_class (classification de risque OCDE)

Cette importance, mesurée par le nombre de fois qu'une variable est mobilisée dans les arbres de décision du modèle, révèle une cohérence forte avec les fondements théoriques de l'analyse du risque de crédit. Le positionnement d'interest_coverage en tête signifie que la capacité à honorer les charges financières constitue le facteur clé de discrimination ; les ratios d'endettement en positions 2 et 3 confirment le rôle critique du levier financier ; la présence de oecd_country_class en position 6 indique que le risque pays demeure un signal pertinent, bien que secondaire.

- **Importance des variables — LightGBM (Figure 25)**

Le modèle LightGBM confirme et enrichit ces résultats :

interest_coverage s'impose à nouveau comme le prédicteur le plus discriminant, traduisant la centralité de la capacité de remboursement dans l'évaluation de la solvabilité. requested_amount_usd reflète quant à lui l'exposition directe au risque. Les ratios de

structure financière confirment leur rôle dominant, tandis que la présence de `payment_type_Open_account` signale que l'instrument de règlement retenu constitue également un indicateur de risque pertinent.

- **Explicabilité comparative : régression logistique vs. modèles d'ensemble**

La régression logistique offre une interprétabilité directe via ses coefficients, mais se limite aux relations linéaires et atteint un plafond de performance autour d'une AUC de 0.77, lié à sa frontière de décision linéaire. XGBoost et LightGBM, sans méthode d'explicabilité, constitueraient des boîtes noires ; avec la méthode d'importance des variables, ils deviennent interprétables et révèlent des relations non linéaires que la régression logistique ne capte pas.

- b. Conclusion sur H5 : PARTIELLEMENT VALIDÉE**

Explicabilité globale obtenue via importance des variables (Figures 21 et 25) : par les facteurs clés (`interest_coverage`, `debt_ratios`, `country_score`) sont clairement identifiés. L'amélioration tenue nous donne : un modèle XGBoost sans importance = boîte noire, avec importance il devient un modèle explicable

Les méthodes avancées : comme SHAP (Shapley Additive exPlanations) : mentionné en théorie, non utilisé - LIME (Local Interpretable Model-Agnostic) : mentionné en théorie, non utilisé - Implique que seule l'explicabilité GLOBALE est obtenue, pas l'explicabilité LOCALE

Alors H5 confirme que la combinaison de ML + méthodes d'explicabilité améliore la compréhension. Les résultats confirment cela **partiellement**.

Statut : PARTIELLEMENT VALIDÉE

II. Discussion des résultats :

La présente discussion vise à articuler de manière synthétique et critique les principaux résultats empiriques obtenus dans le cadre de ce mémoire, en les confrontant aux hypothèses formulées dans le cadre conceptuel ainsi qu'aux apports de la littérature scientifique mobilisée.

L'analyse empirique de cette recherche confirme que l'évaluation du risque de défaut des acheteurs étrangers repose sur une approche multidimensionnelle, combinant indicateurs financiers, risque pays, données transactionnelles et modèles prédictifs avancés. Concernant l'hypothèse H1, les résultats montrent que les indicateurs financiers constituent le socle de l'évaluation. Les variables **interest_coverage**, **debt_to_equity**, **debt_to_assets**, **equity_ratio** et **total_liabilities_t_usd** dominent les modèles XGBoost et LightGBM, et la régression logistique pondérée atteint une AUC significative de 0,776, confirmant la robustesse de ces indicateurs comme prédicteurs universels de défaut. Ces résultats s'alignent parfaitement avec les travaux de (Beaver, 1966), (Altman, 1968) et (Ohlson, 1980) et sont corroborés par (Yazdi, Hanne, & Wang, 2019) ; (Petrová, Krügerová, & Koziel, 2020) et (Bärtland & Krummacker, 2020), qui montrent que la solidité financière est un déterminant central de l'évaluation du risque de crédit. Dans le contexte algérien, caractérisé par une forte asymétrie informationnelle et des marchés cibles variés, l'automatisation du score financier par machine Learning permet à la CAGEX d'améliorer la précision, d'optimiser le temps d'arbitrage et de réduire la subjectivité humaine.

Pour H2, relative à l'influence des variables macroéconomiques et du risque pays, les résultats sont partiellement validés. Les variables **oecd_country_class** et **country_score** apparaissent dans le top 10 des facteurs importants, mais leur contribution est secondaire par rapport aux indicateurs financiers. La littérature soutient l'importance de ces facteurs pour compléter l'évaluation financière (Agarwal, et al., 2023) et (Petrová, Krügerová, & Koziel, 2020) et (Yazdi, Hanne, & Wang, 2019), soulignant que le risque pays capture l'incertitude macroéconomique et institutionnelle inhérente aux marchés étrangers. Dans le contexte algérien, l'intégration de ces variables est cruciale pour ajuster les décisions de couverture vers des acheteurs situés dans des pays à risque élevé, notamment en Afrique subsaharienne.

H3, portant sur l'effet des données transactionnelles, est validée. Les variables **requested_amount_usd**, **payment_type_Open_account** et **credit_term_months** améliorent la performance des modèles XGBoost et LightGBM en capturant des

comportements à risque non linéaires que la régression logistique ne peut détecter. Ces résultats de (Yazdi, Hanne, & Wang, 2019) et (Bärtland & Krummaker, 2020) sur l'importance des informations comportementales pour prédire le défaut. Pour la CAGEX, ces variables transactionnelles permettent de formaliser et d'automatiser partiellement la détection des dossiers risqués, en priorisant les décisions d'arbitrage selon le profil de la transaction. Les perspectives incluent l'intégration de nouvelles informations transactionnelles et l'évaluation de leur impact sur des portefeuilles réels.

Concernant H4, sur la supériorité des modèles de machine learning, les résultats montrent que XGBoost et LightGBM surpassent la régression logistique uniquement après ajustement du seuil de décision. Par exemple, le recall de XGBoost optimisé atteint 0,94, tandis qu'avec un seuil standard il reste à 0.41. Ces résultats soulignent que la performance du machine Learning dépend du calibrage stratégique du seuil, en accord avec les recommandations de (Noriega & Rivera, 2023) et (Liao, Jiao, & Zhang, 2025). Pour la CAGEX, cette flexibilité permet de moduler la prudence ou la sélectivité des décisions selon les objectifs stratégiques. Les perspectives consistent à adopter un déploiement progressif, combinant régression logistique explicable et machine learning pour les dossiers complexes, et à renforcer la supervision humaine pour le calibrage du seuil.

Enfin, H5, portant sur l'explicabilité, est partiellement validée. L'explicabilité globale via l'importance des variables identifie clairement les facteurs clés (**interest_coverage**, ratios d'endettement, **country_score**), mais l'explicabilité locale (SHAP, LIME) n'a pas été implémentée. (Liao, Jiao, & Zhang, 2025) et (Černevičienė & Kabašinskas, 2024) souligne que la performance prédictive doit être accompagnée de la compréhension du processus décisionnel pour instaurer la confiance et la conformité réglementaire. Dans le contexte algérien, cette explicabilité permet de motiver les décisions auprès des exportateurs et de l'État, tout en facilitant l'appropriation des modèles par les équipes internes.

Notre recherche apporte plusieurs apports, coté théorique elle confirme empiriquement la pertinence durable des travaux fondateurs de Beaver, Altman et Ohlson, démontrant que le machine Learning prolonge plutôt qu'il ne supprime l'analyse financière traditionnelle. Elle démontre également la complémentarité empirique de trois dimensions rarement intégrées : les indicateurs financiers, la classification OCDE du risque pays, et les variables transactionnelles, enrichissant ainsi la théorie du crédit commercial international. Un apport méthodologique majeur réside dans l'établissement que la supériorité des modèles

d'ensemble n'est pas inconditionnelle mais dépend du calibrage du seuil opérationnel, remettant en question l'usage exclusif du ROC-AUC comme indicateur de performance. Enfin, cette étude distingue l'explicabilité globale de l'explicabilité locale, montrant que l'importance des variables demeure insuffisante pour justifier des décisions individuelles en contexte réglementaire.

Sur le plan opérationnel, cette recherche fournit un score de risque objectif et reproductible qui améliore la cohérence décisionnelle et accélère le traitement des dossiers, bien que la taille modeste de l'échantillon limite la confiance sur portefeuille futur. Elle transforme le seuil de classification en levier opérationnel flexible, permettant une adaptation selon l'orientation stratégique prudente ou de croissance, mais cette flexibilité nécessite une validation temporelle absente de cette étude. La présence du risque pays parmi les facteurs influents justifie une stratégie tarifaire et de garanties adaptée au contexte géographique, résultat toutefois spécifique au contexte algérien et difficilement transposable. Enfin, l'étude recommande une architecture hybride alternant régression logistique pour les dossiers standards et modèles complexes pour les cas dure, avec supervision humaine finale, bien que l'absence d'explicabilité locale entrave la justification formelle des décisions de refus.

Cette recherche présente plusieurs limites qu'il convient d'identifier explicitement. La base de données utilisée demeure déséquilibrée entre défaillants et non-défaillants malgré les techniques de rééquilibrage appliquées, et sa taille modeste au regard des standards du machine Learning en scoring de crédit limite la généralisation des résultats. Les variables financières proviennent de bilans déclaratifs dont la fiabilité varie significativement selon les pays, et l'absence de données longitudinales restreint l'analyse à une photographie statique de la situation de l'acheteur. L'étude ne dispose pas de validation temporelle stricte compromettant l'évaluation de la robustesse prédictive en environnement de production, et l'absence d'implémentation de méthodes d'explicabilité locale (SHAP, LIME) demeure significative dans un contexte réglementaire où la justification individuelle des refus est une obligation légale. De plus, le modèle n'intègre pas de variables comportementales (historique de paiement antérieur) ni qualitatives (qualité du management, gouvernance) qui peuvent être déterminantes.

Les perspectives pour la CAGEX incluent le déploiement de modèles avancés de machine Learning sur des données réelles et l'intégration de nouvelles métriques financières locales, afin de renforcer la précision de l'évaluation du risque de défaut, notamment pour les

acheteurs situés dans des pays à risque élevé tels que certains pays d’Afrique subsaharienne ou du Moyen-Orient. La CAGEX dispose d’une équipe solide capable d’analyser chaque dossier d’assurance-crédit, et l’adoption de techniques explicables comme SHAP et LIME facilitera la compréhension et la supervision humaine des décisions. Ainsi l’intégration progressive de modèles hybrides, combinant régression logistique explicable et machine Learning pour les dossiers complexes, permettra de moduler la prudence et la sélectivité des décisions selon les objectifs stratégiques. Par ailleurs, il est important de considérer que le machine Learning, malgré ses performances, doit rester un outil d’aide à la décision : la recommandation de crédit reste encadrée par la supervision humaine et des limites prudentes pour garantir la sécurité financière et la conformité réglementaire. Ainsi, l’ensemble des perspectives souligne la nécessité d’une approche intégrée, combinant performance, explicabilité et expertise humaine, en parfaite adéquation avec la modernisation du secteur financier algérien et la mission stratégique de la CAGEX.

CONCLUSION GENERALE

Au terme de ce travail de recherche consacré à la modélisation prédictive en assurance-crédit à l'exportation, et plus particulièrement à l'évaluation du risque de défaut des acheteurs étrangers au sein de la CAGEX, il apparaît clairement que le Machine Learning représente aujourd'hui un outil important d'amélioration des mécanismes d'évaluation du risque dans le domaine de l'assurance-crédit.

L'analyse théorique a permis de mettre en évidence que l'assurance-crédit à l'exportation constitue un instrument essentiel de sécurisation des échanges internationaux face au risque de non-paiement des acheteurs étrangers. Elle a également montré que les méthodes traditionnelles d'évaluation du risque présentent certaines limites, notamment dans des environnements complexes caractérisés par l'incertitude et la diversité des données.

À travers l'étude empirique basée sur une approche quantitative, il ressort que les modèles de Machine Learning utilisés, notamment XGBoost et LightGBM, permettent d'améliorer la prédiction du risque de défaut des acheteurs étrangers par rapport aux méthodes classiques. Les résultats obtenus montrent également l'importance des variables financières, macroéconomiques et comportementales dans le processus d'évaluation du risque.

Toutefois, les résultats révèlent que l'utilisation de ces modèles reste confrontée à plusieurs contraintes, notamment les limites liées à la qualité des données, le déséquilibre des classes et les exigences d'explicabilité des modèles prédictifs dans un environnement réglementé. Ces éléments influencent directement l'efficacité et l'intégration opérationnelle des outils prédictifs au sein de la CAGEX.

Ainsi, en réponse à la problématique posée, il peut être affirmé que la modélisation prédictive par Machine Learning représente une avancée significative dans l'amélioration de l'évaluation du risque acheteur en assurance-crédit à l'exportation. Cependant, son intégration demeure un processus évolutif nécessitant une amélioration continue des données, des outils analytiques et des compétences techniques.

Enfin, ce travail ouvre des perspectives de recherche intéressantes, notamment à travers l'intégration de nouvelles méthodes d'intelligence artificielle explicable, l'élargissement des bases de données utilisées ainsi que l'application de ces modèles à d'autres domaines liés à la gestion du risque financier et assurantiel.

BIBLIOGRAPHIE

- Allianz Trade . (2023, 05 25). Récupéré sur Allianz Trade Qu'est-ce que le risque commercial ? : <https://www.allianz-trade.fr/blog/risque-commercial.html>
- World Bank Group MIGA . (s.d.). Récupéré sur World Bank Group MIGA : <https://www.miga.org/political-risk-insurance?utm>
- Agarwal, N., Chan, J. M., Lodefalk, M., Tang, A., Tano, S., & Wang, Z. (2023). Mitigating information frictions in trade: Evidence from export credit guarantees. *Journal of International Economics*, pp. 1-3.
- AllianzTrade. (2026). Récupéré sur AllianzTrade Notre analyse des risques par pays: <https://www.allianz-trade.fr/etudes-economiques/risque-pays.html>
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 590-607.
- ATIDI. (s.d.). Récupéré sur <https://www.atidi.africa/>
- Atradius . (s.d.). Récupéré sur Atradius Managing risks enabling trade: <https://group.atradius.com/our-solutions/credit-insurance>
- Atradius. (2025, Decembre 17). Récupéré sur Atradius Misconceptions about credit insurance explained: <https://group.atradius.com/knowledge-and-research/resources/misconceptions-about-credit-insurance-explained?utm>
- Baesens, B., Rösch, D., & Scheule, H. (2016). *Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS*. Hoboken: Wiley.
- Bank of England . (2019, Octobre). Récupéré sur Machine learning in UK financial services: <https://www.bankofengland.co.uk/report/2019/machine-learning-in-uk-financial-services>
- BankOfEngland. (2019). *Machine learning in UK financial services*. London : Bank Of England.
- Bärtland, M., & Krummaker, S. (2020, Mars 1). Prediction of Claims in Export Credit Finance: A Comparison of Four Machine Learning Techniques. *Risks MDPI*, pp. 2-4.
- Basel Committee on Banking, S. (2022). CRE36: IRB approach: Minimum requirements to use IRB approach. Basel: Bank for International Settlements.
- Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 76-110.
- Berne Union. (s.d.). Récupéré sur Berne Union , About Export Credit and Investment Insurance: <https://www.berneunion.org/Stub/Display/17>
- Bishop, C. M. (2002). *Pattern Recognition and Machine Learning*. New York: Springer.
- Bpifrance. (s.d.). Récupéré sur Bpifrance: <https://www.bpifrance.com/>

- BREIMAN, L. (2001). Random Forests. *Machine Learning*, 11.
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2020). Explainable Machine Learning in Credit Risk Management. *Computational Economics*, 204-206.
- CAGEX . (s.d.). Récupéré sur CAGEX – Compagnie Algérienne d'Assurance et de Garantie des Exportations: <https://www.cagex.dz/fr/about>
- Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: a systematic literature review. *Artificial Intelligence Review*, 3-8.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *International Conference on Knowledge Discovery and Data Mining*.
- CNIL. (2016, Mai 24). Récupéré sur Le règlement général sur la protection des données - RGPD: <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>
- Coface . (2025, 03 26). Récupéré sur Coface :What Happens When You Request a New Credit Limit?: <https://www.coface.uk/news-economy-and-insights/what-happens-when-you-request-a-new-credit-limit>
- COFACE FOR TRADE . (2024, Avril 04). Récupéré sur COFACE FOR TRADE Qu'est-ce que le risque politique et quel est son impact sur le credit commercial ? : <https://www.coface.com/fr/actualites-economie-conseils-d-experts/qu-est-ce-que-le-risque-politique-et-quel-est-son-impact-sur-le-credit-commercial#comment-ces-risques-sont-ils-identifies>
- Colombier, A. (1958, Mai). L'assurance des risques a l'exportation. *Revue économique*, pp. 482-496.
- CRESWELL, J. W. (2009). *RESEARCH DESIGN :Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage Publications.
- de Lange, P. E., Melsom, B., Vennerød, C. B., & Westgaard, S. (2022). Explainable AI for Credit Assessment in Banks. *Risk and Financial*, 2-3.
- Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. *ArXiv*.
- ECIC. (s.d.). Récupéré sur ECIC EXPORT CREDIT INSURANCE: <https://www.ecic.co.za/products/export-credit-insurance/>
- EIOPA . (s.d.). Récupéré sur Solvency II: https://www.eiopa.europa.eu/browse/regulation-and-policy/solvency-ii_en?prefLang=fr
- EU Artificial Intelligence Act . (2021). Récupéré sur Article 14 : Surveillance humaine: <https://artificialintelligenceact.eu/article/14/?utm>
- European Commission. (2019, Avril 08). Récupéré sur Ethics guidelines for trustworthy AI: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- European Commission. (2021, Janvier 25). Récupéré sur Information request: Availability of short-term export-credit insurance for exports to all countries in the Annex of the Short-term export-credit

insurance Communication: https://competition-policy.ec.europa.eu/public-consultations/2021-short-term-export-credit-insurance_en

- European Banking Authority. (2023). Machine learning for IRB models: Follow-up report from the consultation on the discussion paper on machine learning for IRB models. Paris: European Banking Authority.
- EXIM Bank. (s.d.). Récupéré sur Protect Against Buyer Nonpayment: https://www.exim.gov/solutions/protect-against-buyer-nonpayment?utm_
- EXIM BANK. (s.d.). Récupéré sur Export Import BANK: <https://www.exim.gov/>
- Financial Stability Board. (2017). Artificial intelligence and machine learning in financial services: Market developments and financial stability implications. Basel: Financial Stability Board.
- Finlay, S. (2012). Credit Scoring, Response Modeling, and Insurance Rating: A Practical Guide to Forecasting Consumer Behavior.
- Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (December 2019). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. Basel: Bank for International Settlements.
- Grunert, J., Norden, L., & Weber, M. (2005). The role of non-financial factors in internal credit rating. *Journal of Banking & Finance*, 513-525.
- GUBA, E. G., & LINCOLN, Y. A. (1994). Competing Paradigms in Qualitative Research. Dans E. G. GUBA, & Y. A. LINCOLN, *Handbook of qualitative research* (p. 109). Thousand Oaks, CA: Sage Publications.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D., & Giannotti, F. (2018). A Survey Of Methods For Explaining Black Box Models. arXiv. Récupéré sur arXiv.
- Hand, D. J., & Henley, W. E. (1997). Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of the Royal Statistical Society*, 523-541.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. New York: Springer.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. John Wiley & Sons.
- International Finance Corporation. (s.d.). Récupéré sur International Finance Corporation : <https://www.ifc.org/en/what-we-do/sector-expertise/trade-and-supply-chain-finance/global-trade-finance-program?utm>
- JOURNAL OFFICIEL. (1996). l'Ordonnance n° 96/06 du 10 janvier 1996.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., . . . Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting. *Advances in Neural Information Processing Systems*. doi:https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

- Kramar, L. (1979, juillet). Quelques aspects de l'assurance dans les relations économiques internationales: l'Assurance-Crédit à l'exportation, l'Assurance-Cautionnement à l'étranger et l'Assurance des investissements privés à l'étranger. *The Geneva Papers on Risk and Insurance*, pp. 66-82.
- Lessmann, S., Baesens, B., Seow, H.-v., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*.
- Liao, M., Jiao, W., & Zhang, J. (2025, septembre 26). Research on Trade Credit Risk Assessment for Foreign Trade Enterprises Based on Explainable Machine Learning. *Information MDPI*, pp. 2-4.
- Lundberg, S., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *ArXiv*, 4.
- Mitchell, T. M. (1997). *Machine Learning*. Ohio: McGraw-Hill Science/Engineering/Math.
- Molnar, C. (2025). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Auto-édition.
- NEXI -Nippon Export and Investment Insurance -. (s.d.). Récupéré sur Export Credit Insurance: <https://www.nexi.go.jp/en/products/types/export.html>
- Noriega, J. P., & Rivera, L. A. (2023, NOVEMBRE 7). Machine Learning for Credit Risk Prediction: A Systematic Literature Review. *Data MDPI*, pp. 1-3.
- OCDE. (s.d.). Récupéré sur Export credits: <https://www.oecd.org/en/topics/policy-issues/export-credits.html>
- OCDE. (s.d.). Récupéré sur Arrangement et accords sectoriels: <https://www.oecd.org/en/topics/sub-issues/arrangement-and-sector-understandings.html>
- OCDE . (2026, Avril 28). Récupéré sur Export restrictions on critical materials reached an all-time high: <https://www.oecd.org/en.html?utm>
- OCDE. (2023, mai). Récupéré sur OCDE AI principles: <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>
- OECD. (s.d.). Récupéré sur OECD: https://www.oecd.org/fr/publications/l-arrangement-relatif-a-des-lignes-directrices-pour-les-credits-a-l-exportation-beneficiant-d-un-soutien-public_9789264289871-fr.html
- OECD . (s.d.). Récupéré sur OECD Classification des risques pays des Participants: <https://www.oecd.org/fr/themes/classification-des-risques-pays.html>
- OECD Export credits. (2026, Avril). Récupéré sur OECD: <https://www.oecd.org/en/topics/export-credits.html>
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 109-131.

- Petersen, M. A., & Rajan, R. G. (1997). Trade credit: Theories and evidence. *The Review of Financial Studies*, 678-686.
- Petrová, M., Krügerová, M., & Kozieł, M. (2020). Export Credit Insurance, Risk classification and Current Development: Application in the Czech Republic. Grant: Hodnocení rizik v mezinárodním obchodě na příkladu vybraných zemí OECD a minimalizace rizik v kontextu českého, pp. 54-56.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?" Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 206-207.
- SAUNDERS, A., & ALLEN, L. (2002). *Credit Risk Measurement: New approaches to value at risk and other paradigms*. New York: John Wiley & Sons.
- Tarde, G. d. (1929). LES DÉBUTS DE L'ASSURANCE-CRÉDIT A L'EXPORTATION. *Revue d'économie politique*, pp. 269-279.
- (2021). *The Charter of the Export-Import Bank of the United States*. Washington, D.C.: EXIM Bank.
- TheWorldBankGroup. (2019). *Credit Scoring Approaches guidelines*. Washington, DC: The World Bank Group.
- Uddin, M. S., Chi, G., & Al Janabi, M. A. (2022). Modeling credit risk with a multi-stage hybrid model: An alternative statistical approach. *Journal of Forecasting*, 1387-1392.
- UK Export Finance. (2025). Récupéré sur UK Export Finance: <https://www.ukexportfinance.gov.uk/>
- UnionAfricaine. (2024). *STRATÉGIE CONTINENTALE DE L'INTELLIGENCE ARTIFICIELLE: Mettre l'IA au service du Développement et de la Prospérité de l'Afrique*. Addis-Abeba: Union Africaine.
- Wang, K., Zhao, R., & Peng, J. (2017, Octobre 06). Trade Credit Contracting under Asymmetric Credit Default Risk: Screening, Checking or Insurance. *European Journal of Operational Research*, p. 6.
- Yang, A., Bakshi, N., & Chen, C. J. (2020, juillet 01). Trade Credit Insurance: Operational Value and Contract choice. *Management Science*, pp. 4-5.
- Yazdi, A. K., Hanne, T., & Wang, Y. J.-M. (2019, JUILLET 09). A Credit Rating Model in a Fuzzy Inference System Environment. *Algorithms MDPI*, pp. 3-4.

ANNEXE UNIQUE
ORGANIGRAMME DE LA CAGEX

