

La prédiction de la faillite : essai par le modèle d'Altman et par le modèle IA des machines à vecteur de support

The prediction of bankruptcy : trial by the Altman model and by the IA model of support vector machines

MRABET Najib

Enseignant chercheur

Ecole Supérieure de Technologie OUJDA

Université Mohamed Premier

Management et développement des entreprises et des organisations (MADEO)
Maroc

ADLI Narimane

Doctorante LURIGOR

Faculté de Droit OUJDA

Université Mohamed Premier

Laboratoire Universitaire de Recherche en Instrumentation et Gestion des Organisations
Maroc

Date de soumission : 24/02/2025

Date d'acceptation : 07/04/2025

Pour citer cet article :

MRABET N. & ADLI N. (2025) «La prédiction de la faillite : essai par le modèle d'Altman et par le modèle IA des machines à vecteur de support », Revue du contrôle, de la comptabilité et de l'audit « Volume 9 : numéro 1 » pp : 289- 309.

Résumé :

Les modèles de prévision de la faillite ont subi un grand changement du fait de l'intégration de l'intelligence artificielle aux méthodes de prédiction. En effet, doté d'une capacité d'apprentissage et d'une puissance de calcul, l'I.A a permis le passage des modèles classiques (Z-score d'Altman, Ohlson ...) utilisant des bases de données limitées à des modèles puissants et flexibles qui peuvent résoudre des problèmes complexes en apprentissage automatique. Ces modèles, en général issus de théorèmes mathématiques connus tels que les arbres de CART ou la théorie des ensembles, une fois combinés à des algorithmes d'apprentissage (supervisés ou non), ont permis d'améliorer la performance des prédictions. Bien qu'ils nécessitent de grandes quantités de données et une puissance de calcul importante, ces modèles sont de plus en plus utilisés dans la prédiction de séries temporelles et dans la construction des arbres de décision. Contrairement à la majorité de ces modèles qui utilisent des techniques comme le bagging, le boosting ou la régression, les SVM sont utilisées essentiellement dans le clustering à travers la recherche d'un hyper-plan optimal séparant un groupe de données ou dans notre cas séparant un groupe d'entreprises mais peuvent être aussi utilisées dans la régression.

Mots clés : Prédiction ; Faillite ; Z-score ; I.A ; SVM

Abstract:

Bankruptcy prediction models have undergone a major change as a result of the integration of artificial intelligence into prediction methods. With its learning capacity and computing power, A.I has enabled the transition from traditional models (Altman's Z-score, Ohlson, etc.) using limited databases to powerful and flexible models that can solve complex learning problems. These models, generally based on well-known mathematical theorems such as CART trees or set theory, when combined with learning algorithms (supervised or unsupervised), have improved the prediction performance. Although they require large quantities of data and significant computing power, these models are increasingly used to predict time series and to build decision trees. Unlike the majority of these models, which use techniques such as bagging, boosting or regression, SVMs are used primarily for clustering, through the search for an optimal hyper-plane separating a group of data or, in our case, separating a group of companies, but can also be used for regression.

Keywords: Prediction ; Bankruptcy ; Z-score ; AI ; SVM

Introduction

Le risque de faillite est un concept qui fait toujours l'objet de débats et de discussions à la fois théoriques et empiriques. En effet, les modèles qui ont essayé de prédire la faillite sont très abondants, et ce depuis la moitié du siècle précédent. Afin d'examiner le risque de faillite et ses déterminants, les indicateurs de performance de l'entreprise peuvent être utilisés dans le sens opposé. Ainsi, si la rentabilité, la solvabilité et la création de valeur sont parmi les critères essentiels de mesure de la performance, les modèles de prédiction de la faillite utilisent l'insuffisance de rentabilité où les résultats négatifs, l'insolvabilité et le déficit de création de valeur. Ainsi, des modèles comme le modèle d'Altman (Z-score), le modèle de Zmijewski, ou encore le modèle de Merton renvoient toujours à des indicateurs comme les ratios de rentabilité (ROE, ROI...), ceux de la solvabilité (Ratios d'endettement, Ratios de liquidité) ou les ratios d'activité (Rotation des actifs, Cash-flow opérationnel...).

L'introduction des modèles IA utilisant les mêmes principes et les mêmes indicateurs des modèles qu'on vient de citer a permis de performer l'ensemble des techniques économétriques les plus récurrentes telles que :

- La régression linéaire et logistique : Ces modèles sont souvent utilisés pour estimer des relations linéaires entre des variables dépendantes et indépendantes.
- Les arbres de décision et forêts aléatoires : Ces modèles sont utiles pour capturer des relations non linéaires et des interactions complexes entre variables.
- Les réseaux de neurones : ces modèles, en particulier les réseaux de neurones profonds, sont capables de modéliser des relations très complexes et non linéaires.
- Les modèles de séries temporelles : Des modèles comme les réseaux de neurones récurrents (RNN) ou les transformateurs sont utilisés pour analyser des données temporelles, comme les séries chronologiques économiques.
- Les modèles de clustering : Comme les K-means, ces modèles sont utilisés pour identifier des groupes ou des clusters dans les données, ce qui peut être utile pour segmenter des marchés ou identifier des comportements économiques similaires.

De ce fait, il est logique de s'interroger sur la réalité de l'amélioration de l'I.A. de cette capacité de prédiction. Les nombreuses études qui se sont concentrées sur la comparaison entre ces techniques et leurs homologues classiques confirment leur supériorité en termes de qualité de prédiction (Zammel & Khoufi, 2020). Les techniques d'intelligence artificielle ont l'avantage

de pouvoir être adaptées à n'importe quel problème, notamment la prédiction de la faillite d'une entreprise.

Dans cet article nous allons, essayer d'appliquer un modèle classique de prédiction de la faillite à un ensemble d'entreprises marocaines cotées en bourse à savoir le modèle d'Altman, ensuite, et à l'aide d'un modèle d'IA, nous allons essayer d'améliorer cette prédiction. Pour cela, notre travail sera scindé en deux parties. Dans la première partie, nous allons présenter quelques modèles classiques de prédiction de la faillite ainsi que de nouveaux modèles d'intelligence artificielle utilisés pour le même objectif. Dans la deuxième partie, nous allons essayer de tester des modèles sur des entreprises marocaines cotées en bourse à savoir le modèle Z-score d'Altman et le modèle MSV.

1. Les modèles de prévision de la faillite

Afin de prédire la faillite des entreprises, le choix du modèle dépend d'un ensemble de facteurs, notamment la taille de l'échantillon (nombre d'entreprises et nombre d'années de données disponibles), la complexité des données (nombre de variables), les objectifs de précision et d'interprétabilité et surtout la puissance de calcul à la disposition. Ce dernier facteur nous pousse à diviser cette première partie en deux points distincts à savoir les modèles classiques simples d'utilisation et les modèles à intelligence artificielle.

1.1. Les modèles classiques de prévision de la faillite

La majorité des modèles de prévision de la faillite utilisent généralement des indicateurs financiers et, dans une moindre mesure, des indicateurs opérationnels pour évaluer la performance d'une entreprise. Parmi les modèles les plus récurrents dans la littérature nous allons expliquer le modèle d'Altman (Z-score) et le modèle d'Ohlson.

1.1.1. Le modèle d'Altman (Z-score)

Avec un travail célèbre et précurseur, Edward Altman a développé le modèle de Z-score en 1968, en utilisant une combinaison de 22 ratios financiers (Altman, 1968), ramené à 5 ratios dans la version ZETA, (Altman, Haldeman, & Narayan, 1977), pour prédire la probabilité de faillite sur un échantillon constitué de 66 firmes américaines entre 1946 et 1965. Le Z-score d'Altman a conçu sa fonction prédictive en synthétisant cinq ratios financiers, dont le score pivot est 2.99.

$$Z = 1,2 V_1 + 1,4 V_2 + 3,3 V_3 + 0,6 V_4 + 0,995 V_5$$

Avec 5 ratios financiers (V_1, V_2, V_3, V_4 et V_5) :

- V_1 : Bénéfice non réparti / Total actif : Mesure le renforcement de la solidité financière, c'est-à-dire la part des capitaux propres supplémentaire dans le total bilan qui lui permet d'autofinancer ces investissements et de renforcer sa capacité d'endettement. Plus cet indicateur est important, plus l'entreprise s'éloigne de la faillite ;
- V_2 : EBIT / Total actif : Mesure de la rentabilité opérationnelle. Une entreprise dont le résultat d'exploitation faible voit sa solvabilité diminuée et son risque de faillite augmenté ;
- V_3 : Chiffre d'affaires / Total actif : Mesure de la productivité des actifs. C'est-à-dire plus ce ratio augmente plus l'entreprise réalise des opportunités avec un minimum d'actifs et plus sa solvabilité augmente ;
- V_4 : Valeur de marché des fonds propres / Total actif : Indicateur de solvabilité défini par la confiance des investisseurs dans les actions de l'entreprise, dans le cas d'une entreprise cotée il est égal à la capitalisation boursière sur le total actif ;
- V_5 : BFR / Total actif : Mesure de la liquidité. En effet, les entreprises qui ont des problèmes de liquidité, même les plus solvables, courent un risque de faillite important.

Le score Z obtenu à partir de la formule permet de classer les entreprises dans l'une des trois catégories suivantes :

- $Z > 2,99$: L'entreprise est financièrement saine, avec une faible probabilité de faillite.
- $1,81 < Z < 2,99$: L'entreprise se trouve dans une zone d'alerte (zone grise), ce qui signifie qu'elle est à risque mais que son avenir reste incertain.
- $Z < 1,81$: L'entreprise est en zone de danger et court un risque élevé de faillite.

Le modèle d'Altman est devenu un outil puissant pour évaluer le risque de faillite d'une entreprise en raison de sa simplicité et de sa commodité. Toutefois, s'il a été conçu pour les grandes entreprises manufacturières cotées en bourse, il a connu plusieurs versions adaptées aux besoins spécifiques de différents types d'entreprises :

- Entreprises non cotées (privées) : Altman a ajusté la formule pour les entreprises non cotées, car elles n'ont pas de valeur de marché pour leurs capitaux propres.

La formule devient alors :

$$Z = 0,717V_1 + 0,847V_2 + 3,107V_3 + 0,420V_4 + 0,998V_5$$

- Entreprises du secteur des services ou entreprises non manufacturières : La formule a été modifiée car certaines industries n'ont pas les mêmes caractéristiques que les entreprises manufacturières, notamment en termes d'utilisation des actifs et de répartition des dettes et des capitaux.

$$Z = 6,56V_1 + 3,2V_2 + 6,72V_3 + 1,05V_4$$

1.1.2. Le modèle d'Ohlson

Le modèle d'Ohlson est un modèle prédictif des difficultés des entreprises développé par James A. Ohlson en 1980. Il repose sur une approche statistique, utilisant la régression logistique pour estimer la probabilité qu'une entreprise fasse faillite dans l'année qui suit. Contrairement au modèle précédent (Z-score d'Altman), qui utilisait des seuils fixes, le modèle d'Ohlson fournissait une probabilité continue, ce qui le rend plus flexible et adapté à une variété d'entreprises.

Ce modèle permet non seulement de classer les firmes en défaillantes /non-défaillantes, mais il permet aussi de déterminer la probabilité des entreprises qu'elles fassent faillite (Friedman, Hastie & Tibshirani, 2000). Pour ce faire, Ohlson intègre combinaison de neuf ratios pour confectionner le O-Score :

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot \text{SIZE} + \beta_2 \cdot \text{TL/TA} + \beta_3 \cdot \text{NI/TA} + \beta_4 \cdot \text{WC/TA} + \beta_5 \cdot \text{FFO/TL} + \beta_6 \cdot \text{ROA} + \beta_7 \cdot \text{RET} + \beta_8 \cdot \text{GTA} + \beta_9 \cdot \text{CL})}}$$

Avec :

- $P(Y=1)$: Probabilité que l'entreprise fasse faillite (1= Faillite, 0 = non-faillite)
- β_0 : Constante.
- $\beta_1, \beta_2, \dots, \beta_9$: Coefficients des variables explicatives.

Les neuf variables sont définies comme ci-dessous :

- Taille de l'entreprise (SIZE) : Mesurée par le logarithme du total des actifs. Les grandes entreprises ont généralement plus de ressources pour éviter la faillite, donc une taille plus grande réduit la probabilité de faillite.
- Structure financière (TL/TA) : Ratio de la dette totale (Total Liabilities) sur le total des actifs (Total Assets). Un ratio élevé indique un endettement important, ce qui augmente le risque de faillite.
- Rentabilité (NI/TA) : Ratio du résultat net (Net Income) sur le total des actifs (Total Assets). Une rentabilité faible ou négative augmente la probabilité de faillite.
- Liquidité (WC/TA) : Ratio du fonds de roulement (Working Capital) sur le total des actifs (Total Assets). Une liquidité insuffisante peut indiquer des difficultés à payer les dettes à court terme, ce qui augmente le risque de faillite.
- Flux de trésorerie (FFO/TL) : Ratio des flux de trésorerie d'exploitation (Funds From Operations) sur la dette totale (Total Liabilities). Un ratio faible indique que l'entreprise génère peu de liquidités pour couvrir ses dettes, ce qui augmente le risque de faillite.
- Rendement des actifs (ROA) : Ratio du résultat net (Net Income) sur le total des actifs (Total Assets). Un rendement faible ou négatif indique une mauvaise utilisation des actifs, ce qui augmente le risque de faillite.
- Variabilité des résultats (RET) : Mesure de la volatilité des résultats nets sur plusieurs périodes. Une variabilité élevée indique une instabilité financière, ce qui augmente le risque de faillite.
- Croissance des actifs (GTA) : Taux de croissance du total des actifs (Growth Total Assets). Une croissance trop rapide peut indiquer des problèmes de gestion ou de surinvestissement, ce qui augmente le risque de faillite.
- Performance boursière (CL) : Mesure de la performance boursière de l'entreprise par rapport au marché. Une performance médiocre peut refléter des problèmes sous-jacents, ce qui augmente le risque de faillite.

La probabilité de faillite $P(Y=1)$ est comprise entre 0 et 1. Une probabilité proche de 1 indique un risque élevé de faillite. En revanche, une probabilité proche de 0 indique un risque faible de faillite. En pratique, un seuil (par exemple, 0,5) est souvent utilisé pour classer les entreprises comme "à risque" ou "saines".

1.2. Les modèles d'intelligence artificielle pour la prédiction de la faillite

Les modèles d'intelligence artificielle pour la prédiction de la faillite sont tous des modèles qui basent sur des algorithmes à itération, ils sont capables de réaliser ce qu'on appelle un apprentissage supervisé pour arriver à l'objectif assigné. Les modèles de prédiction de la faillite sont parmi de multiples modèles IA utilisant les réseaux de neurones, la régression logistique, les arbres de décision ou les K-means, dans la prédiction. Leurs applications sont nombreuses tel que le comportement du consommateur dans les études marketing, la simulation des crises en macroéconomie et l'évaluation du risque de défaut de paiement par les banques.

Dans ce qui suit nous allons présenter deux des modèles d'intelligence artificielle les plus utilisés dans la prédiction de la faillite.

1.2.1. La machine à vecteurs de support

Les machines à vecteurs de support (SVM) sont une méthode de classification supervisée, (Boser, Guyon & Vapnik, 1992). Elles sont largement utilisées pour résoudre des problèmes de classification binaire. La méthode SVM (Cortes & Vapnik, 1995 ; Cristianini & Shawe-Taylor, 2000), est un algorithme puissant pour la classification et aussi pour la régression, il est utilisé sur un échantillon d'observations décrites par des variables qualitatives ou quantitatives et surtout pour prédire des valeurs continues.

1.2.1.1. Principe général

La SVM repose sur la recherche d'un hyperplan optimal qui maximise la marge entre les classes, en utilisant des vecteurs de support. Grâce au kernel trick, elle peut également gérer des problèmes non linéaires. L'entraînement des SVM implique la résolution d'un problème d'optimisation quadratique de la forme suivante (Fan, Chen, & Lin, 2005) :

$$\min_{\alpha} f(\alpha) = \frac{1}{2} \alpha^T Q \alpha - e^T \alpha$$

Où :

- α , est un vecteur de variables duales.
- Q , est une matrice symétrique dense, souvent trop grande pour être stockée en mémoire.

- C , est une constante qui contrôle le compromis entre la maximisation de la marge et la minimisation des erreurs de classification.

Les contraintes du problème sont :

$$0 \leq \alpha_i \leq C \quad \text{pour } i = 1, \dots, l,$$

$$\mathbf{y}^T \boldsymbol{\alpha} = 0,$$

où $\mathbf{y}$ est un vecteur contenant les étiquettes des données

$$(y_i \in \{1, -1\}).$$

Cet outil nécessite un choix minutieux des hyperparamètres et peut être coûteux en calcul pour de grands ensembles de donnée (Chang & Lin, 2011). Il est décomposé en deux méthodes :

- Classification (SVM) : L'objectif est de séparer les données en classes distinctes en trouvant un hyperplan optimal qui maximise la marge entre les classes. Les SVM sont principalement utilisées pour la classification binaire, mais peuvent être étendues à la classification multi-classes.
- Régression (SVR) : L'objectif est de prédire une valeur continue en trouvant une fonction qui approxime au mieux la relation entre les variables d'entrée et la variable de sortie. La SVR tolère une certaine marge d'erreur ϵ autour de la prédiction.

1.2.1.2. Fonctionnement de l'algorithme

L'algorithme de SVM a pour objectif de trouver la séparation entre deux classes d'objets avec l'idée que plus la séparation est large, plus la classification est robuste. Dans sa forme la plus simple, celle d'une séparation linéaire et de classes séparables, l'algorithme sélectionne l'hyperplan qui sépare le jeu d'observations en deux classes distinctes de façon à maximiser la distance entre l'hyperplan et les observations les plus proches de l'échantillon d'apprentissage.

Cela se fait en introduisant un paramètre ϵ (epsilon) qui définit une marge d'erreur autour de la fonction de régression.

- Les erreurs de prédiction qui se situent à l'intérieur de cette marge sont considérées comme acceptables et n'affectent pas la fonction de régression.
- Seules les erreurs qui dépassent la marge ϵ sont pénalisées.

Les vecteurs de support sont les points de données qui se trouvent à la limite de la marge ϵ ou qui dépassent cette marge. Ce sont ces vecteurs qui déterminent la fonction de régression. Les autres points de données n'ont pas d'influence sur la fonction de régression.

Pour la méthode SVM, l'utilisation des noyaux (kernels) permet d'agrandir très largement l'espace des données. Les noyaux permettent de projeter les données dans un espace de dimension supérieure où une relation linéaire peut être trouvée. Ainsi, des structures plus compliquées peuvent être détectées. Les noyaux couramment utilisés incluent le noyau linéaire, le noyau polynomial et le noyau gaussien (RBF).

1.2.2. Les modèles à K-means

K-means est un algorithme de clustering qui permet de regrouper les données en différents groupes (ou clusters) en fonction de leurs similitudes. Il est utilisé pour déterminer les groupes de clients présentant des risques similaires, ce qui permet aux entreprises de cibler efficacement leurs efforts de gestion des risques.

Le principe de fonctionnement de l'algorithme K-means consiste à sélectionner aléatoirement un certain nombre de points de données comme centres de clusters, puis à assigner chaque point de données au cluster dont le centre est le plus proche. Ensuite, les centres des clusters sont recalculés en fonction des points de données qui leur sont assignés. Ce processus est répété jusqu'à ce que les clusters ne changent plus ou que la convergence soit atteinte. Cette répétition permet à peaufiner l'algorithme et chaque fois que de nouvelles données sur les entreprises sont introduites ce dernier devient plus précis.

L'algorithme K-means présente des avantages tels que sa rapidité, sa simplicité et sa capacité à gérer de grandes quantités de données. Cependant, il présente également des inconvénients tels que la sensibilité aux données bruyantes et la nécessité de connaître à l'avance le nombre de clusters souhaité.

1.3. Présentation de la base de données

Notre base de données se compose au départ de 43 entreprises marocaines cotées à la SBVC qui ont été cotées durant la période allant de 2010 à 2020. Dans cette base de données nous avons exclu les entreprises financières et les holdings et celles dont les données ne sont pas disponibles durant cette période et cela pour des raisons d'homogénéité dans les indicateurs financiers. Le tableau suivant présente ces entreprises numérotées :

Tableau N°1 : Entreprises de la base de données

N°	Entreprise	N°	Entreprise
1	AFRIC INDUSTRIES SA	23	LAFARGEHOLCIM MAR
2	AFRIQUIA GAZ	24	LESIEUR CRISTAL
3	ALUMINIUM DU MAROC	25	M2M GROUP
4	AUTO-HALL	26	MAGHREB OXYGENE
5	AUTO NEJMA	27	MANAGEM
6	BRASSERIES DU MAROC	28	MED PAPER
7	CARTIER SAADA	29	MICRODATA
8	CENTRALE DANONE	30	MINIERE TOUISSIT
9	CIMENTS DU MAROC	31	NEXANS MAROC
10	COLORADO	32	OULMES
11	CTM	33	PROMOPHARM
12	COSUMAR	34	SAMIR
13	DARI COUSPATE	35	SNEP
14	DELATTRE LEVIVIER	36	SOCIÉTÉ DE REALISATIONS MECANIQUES
15	DISWAY	37	SOCIÉTÉ MAGHREBINE DE MONETIQUE
16	FENIE BROSSETTE	38	SOCIÉTÉ METALLURGIQUE D'IMITER
17	HIGHTECH PAYMENT SYSTEMS	39	SONASID
18	IB MAROC.COM SA	40	SOCIÉTÉ DE THERAPEUTIQUE MAROCAINE
19	INVOLYS	41	STOKVIS NORD AFRIQUE
20	ITISSALAT AL-MAGHRIB	42	STROC INDUSTRIE
21	Jet CONTRACTOR	43	UNIMER
22	LABEL VIE		

Source : élaboré par les auteurs

1.4. Les variables explicatives

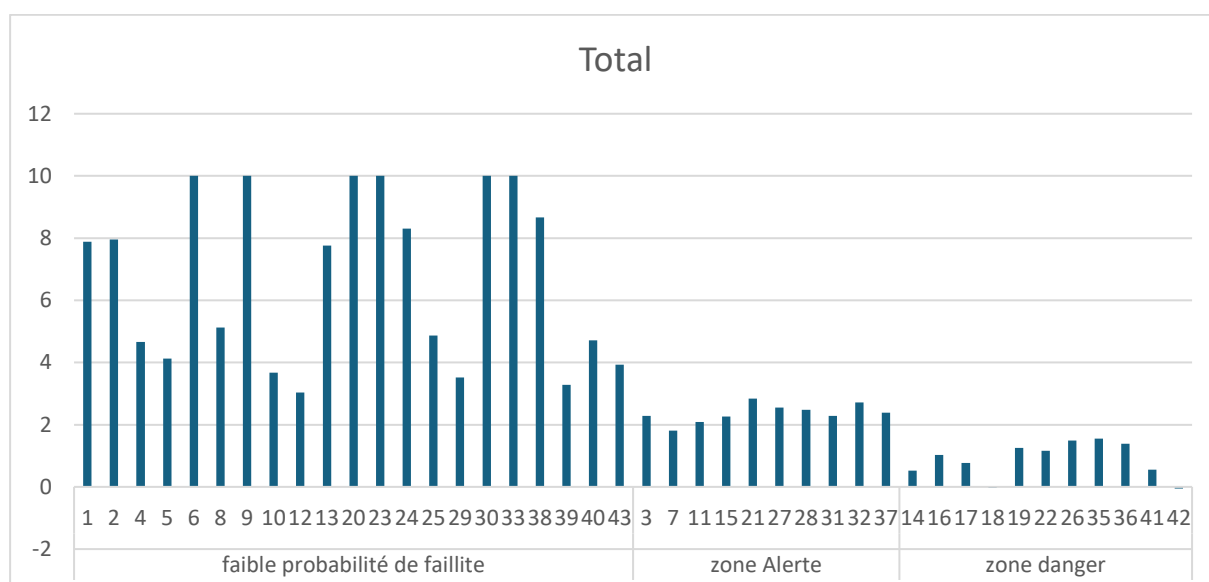
Les variables qu'on va utiliser dans cette étude de cas sont les mêmes proposées par Altman dans son modèle de prédiction de la faillite (Voir la page 5). Ces variables vont nous permettre de tester le fameux Z-score d'Altman, la régression logistique parmi les modèles classiques de prédiction et les machines à vecteurs support parmi les modèles d'intelligence artificielle.

1.5. Le modèle Z-score

La fonction Z-score, cité plus haut, permet d'obtenir un indicateur de faillite de court terme puisqu'il est calculé annuellement. La vérification de l'efficacité de ses prédictions ne peut se faire que durant une à trois années au plus tard. Le deuxième inconvénient à ce score est qu'il peut être dû à un déséquilibre momentané qui peut être inversé facilement ce qui nous a poussé à le calculer sur une période de 11 ans et de classer les entreprises sur la base de la moyenne de cette période.

Le graphe suivant montre le classement obtenu sur la base des trois catégories nommées dans le modèle d'Altman à savoir : Zone de danger, zone d'alerte et faible probabilité de faillite :

Figure N°1 : Répartition des entreprises selon le Z-score d'Altman



Source : élaboré par les auteurs

Afin de bien comprendre le résultat de ce modèle, il est important de faire quelques précisions :

- Le Z-score a été limité à 10 pour des raisons d'affichage (les scores exacts sont dans les annexes) ;
- L'entreprise numéro 34 (La Samir) ne se trouve pas sur le graphique puisqu'elle a déclaré sa faillite en 2016 ;
- Le classement des entreprises (42, 18, 14, 41, 17, 16, 22, 19, 36, 26, 35) en zone danger ne signifie pas que la faillite est imminente ou que le processus ne peut être inversé grâce aux diverses opérations de sauvetage, de redressement ou de restructuration financière ;
- Pour vérifier l'efficacité de prédiction de ce modèle, il a fallu chercher les entreprises qui ont fait faillite après 2020 parmi celles qui ont été étudiées, spécialement parmi celles classées dans la zone danger. En effet, le tableau suivant nous explique les situations de ces entreprises entre 2020 et 2024 :

Tableau N°2 : Situation des entreprises classées en zone danger après l'étude

Zone danger		Situation entre 2020 et 2024
34	SAMIR	Faillite en 2016
42	STROC INDUSTRIE	Demande de procédure de redressement judiciaire le 22 décembre 2022
18	IB MAROC.COM SA	Demande de redressement judiciaire le 30 décembre 2022
14	DELATTRE LEVIVIER	Procédure de sauvegarde en 2022 et liquidation en janvier 2023
41	STOKVIS NORD AFRIQUE	Double aug de K pour éponger les pertes au profit des principaux créanciers
17	HPS	Indicateurs financiers redressés (Problèmes financiers dû à l'internalisation)
16	FENIE BROSSETTE	Indicateurs financiers redressés à l'aide de la société mère SOMED
22	LABEL VIE MAROC	Indicateurs financiers redressés à l'aide de la société mère Label vie
19	INVOLYS	Indicateurs financiers redressés à l'aide de la société mère AM INVEST MOROCCO
36	SRM	Indicateurs financiers redressés à l'aide de la société mère GROUPE PREMIUM
26	MAGHREB OXYGENE	Indicateurs financiers redressés à l'aide de la société mère AKWA GROUP
35	SNEP	Indicateurs financiers redressés à l'aide de la société mère YNNA HOLDING

Source : élaboré par les auteurs

La capacité prédictive limitée au court terme du modèle d'Altman nous a amené à utiliser une matrice Z-score de 11 ans pour classer ces entreprises. Quatre entreprises parmi les douze classées en zone danger ont fait faillite réellement ce qui nous pousse à dire que le modèle Z-score d'Altman reste efficace malgré son ancienneté et sa simplicité. La majorité des autres

entreprises classées en zone danger mais qui n'ont pas fait faillite partagent une caractéristique importante, celle de l'appartenance à une société mère puissante.

1.6. La modélisation par machine à vecteurs de support

Nous allons dans ce dernier point utiliser le modèle d'intelligence artificielle avec un outil d'apprentissage supervisé, « les machines à vecteurs de support », expliqué plus haut pour améliorer cette capacité prédictive en se basant toujours sur les mêmes variables explicatives du modèle d'Altman. Pour prédire la faillite, nous allons utiliser la classification binaire (Zone de sécurité contre Zone danger) basé sur les résultats du modèle d'Altman. Nous avons aussi réintégré les deux entreprises éliminées de la régression logistique pour avoir 42 entreprises sur 11 années c'est-à-dire 462 observations :

2.6.1. Statistiques de l'évaluation

Tableau N°3 : Statistiques descriptives (Données qualitatives) :

Variable	Modalités	Comptages	Effectifs	%
Prédiction	Zone de sécurité	341	341	73,81
	Zone danger	121	121	26,19

Classifieur binaire (Zone de sécurité contre Zone danger) :

Classe positive	Zone de sécurité
Nombre d'observations dans le jeu d'apprentissage	462
Biais	2,375
Nombre de vecteurs supports	96

Source : élaboré par les auteurs

Tableau N°4 : Matrice de confusion pour l'échantillon d'apprentissage (Zone de sécurité contre Zone danger) :

De \ Vers	Zone de sécurité	Zone danger	Total	% correct
Zone danger	0	121	121	100
Zone de sécurité	330	11	341	96,774

Total	330	132	462	97,619
-------	-----	-----	-----	--------

Source : élaboré par les auteurs

- **De \ Vers** : Indique les classes réelles (de) et les classes prédites (vers).
- **Zone danger** : Représente la classe positive ou d'intérêt.
- **Zone de sécurité** : Représente toutes les autres classes combinées.
- **Total** : Indique le nombre total d'observations pour chaque classe réelle et prédite.
- **% correct** : Indique le pourcentage d'observations correctement classées pour chaque classe réelle et pour l'ensemble du modèle.

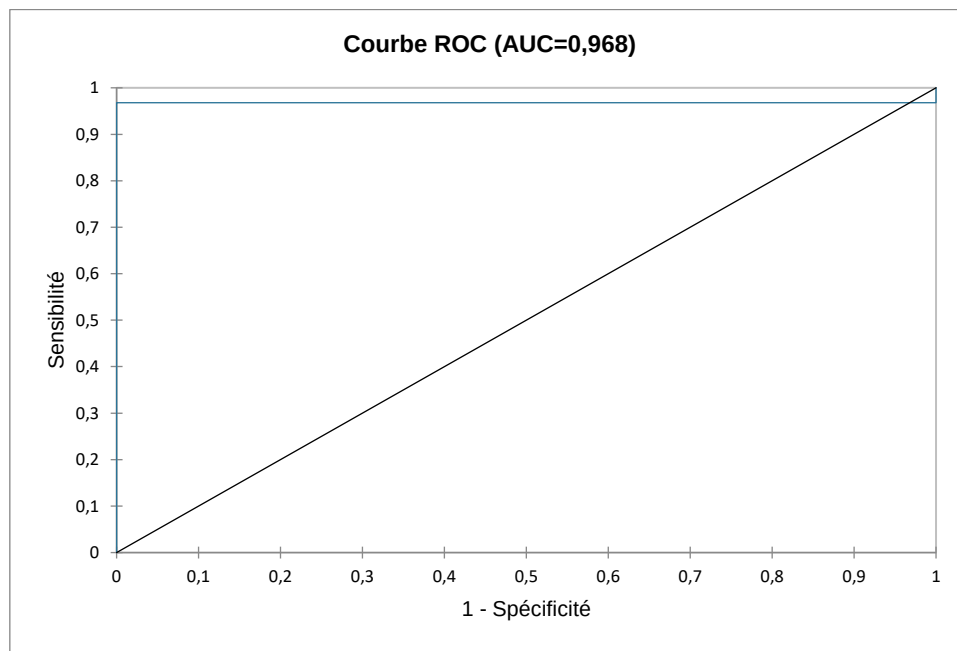
Tableau N°5 : Courbe ROC : Indicateurs de performance (Zone de sécurité contre Zone danger)

Statistique	Echantillon d'apprentissage
Exactitude	0,976
Précision	1,000
Sensibilité	0,968
F-mesure	0,984
Spécificité	1,000
TFP	0,000
Prévalence	0,738
Kappa de Cohen	0,940
TEN	0,738
AUC	0,968

Source : élaboré par les auteurs

- Exactitude (Accuracy) : 0,976 (97,6%)
 - Proportion d'observations correctement classées.
 - Très élevée, indiquant une excellente performance globale du modèle.
- Précision (Precision) : 1,000 (100%)
 - Proportion d'observations prédites comme positives qui sont réellement positives.
 - Maximale, indiquant que le modèle ne fait aucune erreur lorsqu'il prédit la classe positive.
- Sensibilité (Recall/Sensitivity) : 0,968 (96,8%)
 - Proportion d'observations positives qui sont correctement identifiées.

- Très élevée, indiquant que le modèle détecte presque tous les cas positifs.
- F-mesure (F1-score) : 0,984
 - Moyenne harmonique de la précision et de la sensibilité.
 - Très élevée, confirmant une bonne performance globale, équilibrant précision et sensibilité.
- Spécificité (Specificity) : 1,000 (100%)
 - Proportion d'observations négatives qui sont correctement identifiées.
 - Maximale, indiquant que le modèle ne fait aucune erreur lorsqu'il prédit la classe négative.
- TFP (False Positive Rate) : 0,000 (0%)
 - Proportion d'observations négatives incorrectement prédites comme positives.
 - Nulle, confirmant une spécificité parfaite.
- Prévalence (Prevalence) : 0,738 (73,8%)
 - Proportion d'observations positives dans l'ensemble de données.
 - Indique un déséquilibre des classes, avec une prévalence élevée de la classe positive.
- Kappa de Cohen (Cohen's Kappa) : 0,940
 - Mesure l'accord entre les prédictions du modèle et les observations réelles, en tenant compte de l'accord aléatoire.
 - Très élevé, indiquant un excellent accord.
- TEN (Threat Score/Critical Success Index) : 0,738
 - Mesure la performance du modèle en tenant compte des vrais positifs, des faux positifs et des faux négatifs.
 - Bon score, indiquant une bonne performance.

Figure N°2 : Courbe ROC sensibilité/spécificité

Source : élaboré par les auteurs

- AUC (Area Under the ROC Curve) : 0,968
 - Mesure la capacité du modèle à distinguer les classes positives et négatives.
 - Très élevé, indiquant une excellente capacité de discrimination.

2.6.2. Interprétation Générale

Toutes les métriques indiquent une excellente performance du modèle SVM. En effet, le modèle ne commet aucune erreur lorsqu'il prédit les classes positives et négatives. Avec une précision et une spécificité parfaite, ce qui confirme les résultats de départ. Le modèle détecte aussi tous les cas positifs ce qui témoigne d'une sensibilité élevée. Avec un taux de classification globale correct de 97,619%, les prédictions du modèle sont en excellent accord avec les observations des variables du modèle d'Altman et de la régression logistique réalisée après.

En comparaison avec les résultats du modèle d'Altman, bien que le modèle SVM ait été entraîné sur un nombre d'observations limité dans notre situation (462 observations est un nombre trop petit par rapport aux capacités de l'algorithme), il a classé 11 observations de la "Zone de sécurité" dans la "Zone danger". Cette marge d'erreur de 2,3% n'est considérée comme une erreur de prédiction que dans l'hypothèse où les résultats du modèle d'Altman

soient corrects, dans le cas contraire le taux de classification globale correct passe de 97,619% à 100%.

Enfin, il faut rappeler que la majorité des autres entreprises classées en zone danger mais qui n'ont pas fait faillite après 2020 possèdent un facteur en commun très important qui réside dans leur appartenance à des groupes de sociétés puissants capables de les soutenir dans les moments délicats.

Conclusion

Dans notre étude on a pu avoir une idée sur la capacité de l'IA à améliorer les prédictions des modèles financiers classiques tel le modèle d'Altman. L'utilisation de l'IA pour estimer, prédire et analyser des relations économiques à partir de bases de données de plus en plus grandes et complexes est devenue incontournable actuellement. Cette capacité peut être utilisée dans de multiples domaines financiers comme la prédiction des mouvements des marchés financiers, l'évaluation des risques, ou l'optimisation des portefeuilles.

Des techniques comme XG-Boost, SMV et les forêts aléatoires peuvent gérer, analyser et prédire des données aussi complexes et volumineuses que celles de la comptabilité nationale, les stress-tests de prédiction des crises et des chocs économiques ou boursiers. Au niveau des variables qualitatives et comportementales, leur difficulté de prédiction liée au facteur humain ou sociale non quantifiable a beaucoup diminué par l'IA, grâce aux techniques de simulation, de minimisation des erreurs et de clustering.

Notant que malgré cette amélioration beaucoup de limites restent à franchir. En effet, ce processus de prédiction nécessite encore une compréhension approfondie des données et du domaine d'application qu'on veut prédire. Autrement dit, si l'IA a certainement permis de performer les modèles économétriques de prédiction, la compréhension du facteur théorique et de ses implications réelles sur le terrain reste la base de toute prédiction économique ou financière efficace.

Enfin, ce renforcement de la capacité de prédiction par l'IA en gérant plus de variables dans un espace temporelle plus grand ouvre des possibilités faramineuses à la planification (stratégique, budgétaire ou autre) pour toutes les entités économiques que ce soit au niveau micro ou macroéconomique. Cela pose obligatoirement la question de l'urgence de l'adoption

de ces techniques par ces entités car selon nous la capacité de prédiction signifie une rapidité d'adaptation, une prise de décision éclairée et un pouvoir renforcé d'optimisation sinon un plus de pouvoir tout court.

ANNEXES

Annexe 1 : Calcul du Z-score d'Altman (moyenne 2010-2020)

Faible probabilité de faillite			Zone alerte		Zone danger		
Entreprise	Z-score		Entreprise	Z-score	Entreprise	Z-score	
12 COSUMAR	3,031	7	CARTIER SAADA	1,811	42 STROC INDUSTRIE	-0,050	
39 SONASID	3,279	11	CTM	2,083	18 IB MAROC.COM SA	-0,032	
29 MICRODATA	3,522	15	DISWAY	2,260	14 DELATTRE LEVIVIER	0,522	
10 COLORADO	3,671	31	NEXANS MAROC	2,279	41 STOKVIS N AFRIQUE	0,558	
43 UNIMER	3,931	3	ALUMINIUM DU MAROC	2,280	17 HPS	0,768	
5 AUTO NEJMA	4,126	37	SMM	2,389	16 FENIE BROSSETTE	1,025	
4 AUTO-HALL	4,662	28	MED PAPER	2,474	22 LABEL VIE	1,163	
40 SOTHEMA	4,713	27	MANAGEM	2,553	19 INVOLYS	1,250	
25 M2M GROUP	4,869	32	OULMES	2,719	36 SRM	1,393	
8 CENTRALE DANONE	5,122	21	JET CONTRACTOR	2,839	26 MAGHREB OXYGENE	1,495	
13 DARI COUSPATE	7,756				35 SNEP	1,548	
1 AFRIC INDUSTRIES SA	7,887						
2 AFRIQUIA GAZ	7,957						
24 LESIEUR CRISTAL	8,310						
38 SMI	8,665						
6 MINIERE TOUISSIT	10,543						
9 CEMENTS DU MAROC	11,186						
20 SMI	11,277						
23 LAFARGEHOLCIM	12,252						
30 PROMOPHARM	13,935						
33 BRASSERIES DU MAROC	16,258						
20 ITISSALAT AL-MAGHRIB	2706,15						

Source : élaboré par les auteurs

BIBLIOGRAPHIE

Livres

- Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press.
- Ronald. S King (2015). Cluster Analysis and Data Mining » Mercury Learning and Information.

Articles de revues

- Altman E.I, (1968) Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, Journal of Finance, Vol. 23, No. 4., p594
- Altman E.I., Haldeman R.G. et Narayan P. (1977). Zeta Analysis: A new model to identify bankruptcy risk or corporations.
- Chang, C.-C., & Lin, C.-J. (2011) LIBSVM: A Library for Support Vector Machines. ACM Transactions on Intelligent Systems and Technology (TIST), 2(3), 1–27.
- Cortes, C., & Vapnik, V. (1995) Support-Vector Networks. Machine Learning, 20 (3), 273–297.
- Fan, R., Chen, P. & Lin, C. (2005) Working Set Selection Using Second Order Information for Training Support Vector Machines, Journal of Machine Learning Research (6), 1889–1918.
- Friedman, T. Hastie, T. and Tibshirani, R. Additive Logistic Regression: A Statistical View Of Boosting, The Annals of Statistics 2000, Vol. 28, (2), 337–407.
- Noble, W. S. (2006) What is a Support Vector Machine? Nature Biotechnology, 24(12), 1565–1567.
- Zammel, M. & Khoufi, W. (2020). The use of Bayesian Networks in the prediction of bankruptcy», Revue Française d'Economie et de Gestion «Volume 1 : Numéro 1» pp : 60- 88.

Rapports d'activités :

- Rapports financiers et sociaux des années 2010 à 2016 tirés du site : www.ammc.ma, concernant la SAMIR ;
- Rapports financiers et sociaux des années 2010 à 2020 tirés du site : www.ammc.ma, concernant les entreprises suivantes :

AFRIC INDUSTRIES SA, AFRIQUIA GAZ, ALUMINIUM DU MAROC, AUTO HALL, AUTO NEJMA, CARTIER SAADA, CENTRALE DANON, CEMENTS DU MAROC, COLORADO, COMPAGNIE DE TRANSPORT AU MAROC, COMPAGNIE MINIERE DE TOUISSIT, COSUMAR, DARI COUSPATE, DELATTRE LEVIVIER MAROC, DISWAY, FENIE BROSSETTE, HIGHTECH PAYMENT SYSTEMS, IB MAROC.COM SA, INVOLYS, ITISSALAT AL-MAGHRIB, JET CONTRACTORS, LABEL VIE, LAFARGEHOLCIM MAROC, LESIEUR CRISTAL, M2M GROUP, MAGHREB OXYGENE, MANAGEM, MED PAPER, MICRODATA, MUTANDIS SCA, SOCIETE DE PROMOTION PHARMACEUTIQUE DU MAGHREB S.A, SOCIETE DE REALISATIONS MECANQUES, SOCIETE DE THERAPEUTIQUE MAROCAINE, SOCIETE DES INDUSTRIE DU MAROC, SOCIETE LES EAUX MINERALES D'OULMES, SOCIETE MAGHREBINE DE MONETIQUE, SOCIETE METALLURGIQUE D'IMITER, SOCIETE NATIONALE DE SIDERURGIE SA, SOCIETE NATIONALE D'ELECTROLYSE ET DE PETROINDUSTRIE, STOKVIS NORD AFRIQUE, STROC INDUSTRIE, UNIMER.