



وزارة البحث العلمي والتعليم العالي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE
LA RECHERCHE SCIENTIFIQUE

جامعة عبد الحميد بن باديس مستغانم

Université Abdelhamid Ibn Badis Mostaganem

كلية العلوم والتكنولوجيا

Faculté des Sciences et de la Technologie

DÉPARTEMENT DE GÉNIE ÉLECTRIQUE ET ÉLECTRONIQUE
(GEE)



Filière : Électronique

**MEMOIRE DE FIN
D'ETUDES DE MASTER
ACADEMIQUE**

Thème

**Conception d'une application intelligente pour la pré-évaluation
environnementale des produits agroalimentaires algériens
destinés à l'exportation**

Présenté par :

Mlle. HANAFI Hanan

Soutenu en juin 2026, devant le jury composé de :

Mr BENSTALI Wissam

Pr

Univ. Mostaganem Président

Mr YAGOUBI Benabdellah

Pr

Univ. Mostaganem Encadreur

Mr MEHIDI Abdellah

MCB

Univ. Mostaganem Examineur

Année Universitaire 2025/2026

Remerciements

Je remercie tout d'abord ALLAH Tout-Puissant de m'avoir donné la force, la patience, la volonté, le courage et la connaissance afin d'accomplir ce modeste travail.

Mes premiers remerciements s'adressent, de tout cœur et tout naturellement, à mon encadrant, **Monsieur YAGOUBI Benabdellah**, Professeur à l'université Abdelhamid Ibn Badis de Mostaganem, pour la qualité exceptionnelle de son encadrement, son aide précieuse, ses compétences, ses conseils et ses orientations tout au long de ce mémoire. Ces derniers m'ont permis de mener à bien et de finaliser ce travail de Master en Systèmes Embarqués. Je lui suis également reconnaissante pour sa disponibilité et ses qualités humaines qui m'ont permis de travailler dans les meilleures conditions.

Mes sincères gratitudes s'adressent également aux membres du jury qui ont accepté d'évaluer ce travail et de l'enrichir par leurs critiques constructives.

Enfin, je ne saurais oublier de remercier toutes les personnes qui me sont chères, et particulièrement mes parents, mes frères et mes sœurs, pour leurs encouragements, leur confiance et le soutien moral dont ils ont fait preuve tout au long de ce travail. Je remercie également toutes les personnes qui ont contribué, de près ou de loin, au succès de ce mémoire de Master en Systèmes Embarqués.

Dédicace

Je dédie ce mémoire à mes chers parents,

À ma mère, **Fatiha**, mon pilier et ma lumière. Maman, merci d'être ce refuge de douceur et cette source inépuisable de bienveillance. Tes prières silencieuses, ton amour inconditionnel et ton soutien de chaque instant ont été le socle sur lequel j'ai bâti mes rêves. Ce travail est, avant tout, le reflet de ton dévouement.

À mon père, **Abdelali**, mon guide et ma force. Papa, merci pour tes sacrifices, ta droiture et cette confiance aveugle que tu m'as toujours témoignée. Ton courage est le moteur qui m'a poussée à aller toujours plus loin et à ne jamais baisser les bras. Merci d'avoir cru en moi, bien avant que je ne croie en moi-même.

À mon cher frère, **Mohammed**, ainsi qu'à mes chères sœurs, **Hadjira, Salima, Ilhem, Yasmin** et **Wafaa**, pour votre présence précieuse, vos encouragements constants et vos conseils qui ont été une véritable boussole dans les moments de doute.

À tous ceux qui ont contribué, de près ou de loin, à la réussite de ce parcours, et à toutes les âmes résilientes qui, malgré les épreuves, continuent de poursuivre leurs rêves avec passion et détermination.

HANAFI HANAN

Sommaires :

Introduction générale	1
Chapitre I : Contexte du Projet et Enjeux de la Transition Écologique	3
Chapitre II : État de l'Art sur les approches d'apprentissage automatique	12
Chapitre III : Conception et Développement de l'Application	29
Chapitre IV : Étude de faisabilité et modèle économique	56
Conclusion générale et Perspectives	65
Références bibliographiques	

Table des matières :

Introduction générale	1
-----------------------------	---

Chapitre I : Contexte du Projet et Enjeux de la Transition Écologique

I.1.Introduction.....	3
I.2. Introduction à la Transition Écologique	3
I.3. Problématique de l'Empreinte Carbone en Algérie	4
I.4. Analyse de cycle de vie (ACV).....	5
I.4.1. Phases et normes ISO 14040/14044.....	5
I.4.2 Indicateurs environnementaux : CO ₂ , eau et énergie	6
I.5.1. Vision et objectifs de la plateforme	8
I.5.2. Proposition de valeur : évaluation et classification de l'IA.....	8
I.5.3. Positionnement par rapport aux solutions existantes	9
I.6. Analyse SWOT de la Startup ÉcoTrace-IA.....	10
I.7.Conclusion	11

Chapitre II : État de l'Art sur les approches d'apprentissage automatique

II.1. Introduction	12
II.2. Panorama du Machine Learning	12
II.2.1. IA, ML (apprentissage automatique) et DL (apprentissage profond).....	12
II.2.2. Du ML classique au DL : évolution et ruptures	13
II.2.3. Les paradigmes d'apprentissage : supervisé, non supervisé, par renforcement	15
II.3. Le neurone artificiel _ le perceptron	16
II.3.1. Du neurone biologique au neurone artificiel	16
II.3.2. Rôle des poids, du biais et de la fonction d'activation	17
II.4. Fonctions d'activation.....	17
II.4.1. La fonction ReLU (Rectified Linear Unit).....	18
II.4.2. La fonction Sigmoidale.....	18
II.4.3. La fonction Softmax	18

II.4.4. La fonction Tangente Hyperbolique (Tanh).....	19
II.5. Réseaux de neurones artificiels (ANN / MLP).....	20
II.5.1. Architecture générale	20
II.5.2. Capacité d'approximation universelle	21
II.6. Processus d'apprentissage et optimisation	22
II.6.1. Propagation avant et rétropropagation	22
II.6.2. Fonctions de coût et optimiseurs	22
II.7. Apprentissage Non Supervisé : K-Means	23
II.7.1. Principe de l'algorithme	23
II.7.2. Initialisation et choix de k	24
II.8. Évaluation des modèles et régularisation	26
II.9. Revue des travaux connexes.....	27
II.10. Conclusion	27

Chapitre III : Conception et Développement de l'Application

III.1. Introduction	29
III.2. Architecture globale.....	29
III.2.1. Processus opérationnel.....	29
III.2.2. Flux de données.....	30
III.2.3. Justifications techniques	30
III.3. Modélisation du système de notation environnementale	30
III.3.1. Définition des critères environnementaux	31
III.3.2. Échelle d'évaluation et modalités de Notation.....	31
III.3.2.1. Logique de conception de l'échelle.....	31
III.4. Évolution méthodologique : Modélisation par Réseaux de Neurones Artificiels (ANN)	32
III.4.1. Architecture du Perceptron Multicouche (MLP).....	32
III.4.1.1. Méthode 1 : Approche par Pôles Extrêmes et Interpolation de Confiance	32
III.4.1.2. Méthode 2 : Approche par Intervalles Statistiques et Écart-Type	32

III.4.2. Justification du choix d'une approche inspirée du K-Means	33
III.5. Conception algorithmique et développement du modèle K-Means :	33
III.5.1. Application N°1 — LogiClass : phase d'Apprentissage	34
III.5.1.1. Description Fonctionnelle des Commandes de l'Interface Graphique	34
III.5.1.2. Déroulement du Processus en Trois Étapes Successives	35
III.5.1.3. Analyses Expérimentales et Dynamique de Convergence	38
III.5.2. Application N°2 — RT-KMC : Phase de test et Décision	41
III.5.2.1. Description de l'Interface de Test	41
III.5.2.2. Algorithme de Décision et Traitement Interne	42
III.5.2.3. Évaluation Expérimentale et Cas de Tests du Module de Décision	43
III.5.3. Architecture et Approche de l'Application N°3 — ÉcoTrace-IA	45
III.5.3.1. Interface Utilisateur (GUI) et Saisie des Données	46
III.5.3.2. Moteur de Calcul : Logique Algorithmique et Pondération	48
III.5.3.3. Résultats et Validation par cas pratiques	50
III.6. Déploiement et mise en place de l'application (Format .exe)	54
III.6.1. Unité d'apprentissage — LogiClass	54
III.6.2. Unité d'apprentissage (RT-KMC)	54
III.6.3. Unité d'évaluation (ÉcoTrace-IA)	55
III.7. Conclusion	55

Chapitre IV : Étude de faisabilité et modèle économique

Introduction	57
Business Model Canvas (BMC)	57
Business Plan	59
Conclusion	65
Conclusion générale	65
Perspectives	69

Références bibliographiques

Résumé

Face aux exigences croissantes du mécanisme d'ajustement carbone aux frontières (MACF) et aux défis de la transition écologique, ce mémoire présente ÉcoTrace-IA, une solution logicielle de pré-évaluation environnementale dédiée aux PME agroalimentaires algériennes. Le système repose sur un classificateur à distance euclidienne inspiré de l'algorithme K-Means, évaluant chaque produit selon 16 critères environnementaux pondérés et lui attribuant automatiquement un Éco-Grade de A à E. Trois applications complémentaires ont été développées sous Python/Tkinter : LogiClass, pour la convergence et la persistance des centroïdes ; RT-KMC, pour la classification en temps réel dans un espace bidimensionnel ; et ÉcoTrace-IA, pour la pré-évaluation multicritère et la génération de recommandations d'optimisation. L'étude de faisabilité économique confirme la viabilité commerciale du projet à travers un modèle d'abonnement annuel de 4 800 000 DA par entreprise, projeté sur la période 2027–2032, dégagant un cash flow réel positif dès la première année. Les résultats obtenus démontrent la pertinence de l'approche proposée comme levier concret au service de la durabilité industrielle algérienne.

Mots-clés : ÉcoTrace-IA, évaluation environnementale, PME agroalimentaires, K-Means, distance euclidienne, Éco-Grade, intelligence artificielle, MACF, Python/Tkinter.

Abstract

In response to the growing requirements of the Carbon Border Adjustment Mechanism (CBAM) and the challenges of ecological transition, this thesis presents **ÉcoTrace-IA**, a software solution for the pre-environmental assessment of agri-food products, designed specifically for Algerian small and medium-sized enterprises. The system is based on a Euclidean distance classifier inspired by the K-Means algorithm, evaluating each product across 16 weighted environmental criteria and automatically assigning an Eco-Grade ranging from A to E. Three complementary applications were developed using Python/Tkinter : **LogiClass**, for centroid convergence and persistence ; **RT-KMC**, for real-time classification in a two-dimensional space ; and **ÉcoTrace-IA**, for multicriteria pre-assessment and the generation of optimization recommendations. The economic feasibility study confirms the commercial viability of the project through an annual subscription model of 4,800,000 DZD per client company, projected over the 2027–2032 period, generating a positive real cash flow from the

first year of operation. The results obtained demonstrate the relevance of the proposed approach as a concrete lever in support of industrial sustainability in Algeria.

Keywords : ÉcoTrace-IA, environmental assessment, agri-food SMEs, K-Means, Euclidean distance, Eco-Grade, artificial intelligence, CBAM, Python/Tkinter.

الملخص

يقدم هذا العمل تصميم وتطوير حل برمجي مبتكر مخصص لتقييم وتصنيف الأثر البيئي للمنتجات الغذائية الزراعية في الجزائر. نظراً لتعقيد طرق تقييم دورة الحياة التقليدية، يقترح نهجنا **منظومة ثلاثية** منظمة حول ثلاث تطبيقات برمجية.

التطبيق الأول، **LogiClass**، هو وحدة التعلم الخوارزمي. يُنفذ هذا التطبيق خوارزمية K-Means في شكلها التكراري الكامل، ويضمن توليد مراكز التجمعات (Centroïdes) وحفظها.

التطبيق الثاني، **RT-KMC**، هو منصة تجريبية للتصنيف الفوري. يستغل مراكز التجمعات المُولدة بواسطة LogiClass للتحقق الرياضي من آلية القرار بالمسافة الإقليدية.

أما التطبيق الثالث، **ÉcoTrace-IA**، فهو أداة صناعية لاتخاذ القرار، مستقلة ومحددة. تم تطويره باستخدام لغة Python (مكتبتا Tkinter و NumPy)، ويوفر واجهة رسومية تفاعلية تمكّن المصنّعين من إدخال **16 معياراً** بيئياً لتحديد الدرجة البيئية فوراً من A إلى E.

تؤكد دراسة الجدوى الاقتصادية قابلية المشروع للتسويق من خلال نموذج اشتراك سنوي بقيمة 4.800.000 دج لكل مؤسسة عميلة، مُسقط على الفترة 2027-2032، مع تحقيق تدفق نقدي حقيقي إيجابي منذ السنة الأولى. تُثبت النتائج المحققة مدى ملائمة المقاربة المقترحة بوصفها رافعة فعلية لدعم الاستدامة الصناعية الجزائرية.

يمنح هذا النهج **الثلاثي**، الذي يجمع بين البحث الخوارزمي والتطبيق المهني، الشركات الجزائرية أداة ملموسة لاستباق المتطلبات البيئية وتحسين عمليات الصياغة البيئية.

الكلمات المفتاحية: الذكاء الاصطناعي، الدرجة البيئية (Éco-Grade)، خوارزمية K-Means، الشبكات العصبية، تقييم دورة الحياة (LCA)، الصناعات الغذائية الزراعية، **LogiClass**، **RT-KMC**، **ÉcoTrace-IA**، Python.

Liste des Figures

Chapitre I : Contexte du Projet et Enjeux de la Transition Écologique

FIGURE I.1 : APPROCHE MULTICRITÈRES ET IMPACTS ENVIRONNEMENTAUX DE L'ACV [15]	5
FIGURE I.2 : LES ÉTAPES DU CYCLE DE VIE D'UN PRODUIT [16]	6
FIGURE I.3 : RÉPARTITION DES ÉMISSIONS DE CO2 PAR ÉTAPE [18]	7

Chapitre II : État de l'Art sur les approches d'apprentissage automatique

FIGURE II.1 : RELATION D'INCLUSION ENTRE IA, ML ET DL [23]	13
FIGURE II.2 : COMPARAISON DU PROCESSUS ML CLASSIQUE ET DEEP LEARNING [25]	14
FIGURE II.3 : LES TROIS PARADIGMES DE L'APPRENTISSAGE AUTOMATIQUE [27]	16
FIGURE II.4 : ANALOGIE STRUCTURELLE ENTRE LE NEURONE BIOLOGIQUE ET LE NEURONE ARTIFICIEL [29]	17
FIGURE II.5 : REPRÉSENTATION GRAPHIQUE DES FONCTIONS D'ACTIVATION (SIGMOÏDE, TANH, RELU, SOFTMAX) [31]	19
FIGURE II.6 : ARCHITECTURE D'UN MLP À DEUX COUCHES CACHÉES [32]	21
FIGURE II.7 : ÉTAT INITIAL DES CENTROÏDES K-MEANS AVANT CONVERGENCE [39]	24
FIGURE II.8 : CONVERGENCE FINALE DE L'ALGORITHME K-MEANS — PARTITIONNEMENT EN QUATRE CLUSTERS DISTINCTS [39]	25

Chapitre III : Conception et Développement de l'Application

FIGURE III.1 : CAPTURE D'ÉCRAN DE L'INTERFACE TKINTER VIDE OU AVEC LES BOUTONS.....	36
FIGURE III.2 : NUAGE DE POINTS BRUTS AVEC UNIQUEMENT LES ÉTOILES NOIRES.....	37

FIGURE III.3 : GRAPHIQUE DU NUAGE DE POINTS AVEC UNIQUEMENT LES CROIX 'X' AU MILIEU	38
FIGURE III.4 : GRAPHIQUE MONTRANT ITÉRATIF = 1, OÙ LES CERCLES TOUCHENT PRESQUE LES CROIX	39
FIGURE III.5 : GRAPHIQUE MONTRANT ITÉRATIF = 2, AVEC LES TRAJECTOIRES EN '+'.....	40
FIGURE III.6 : GRAPHIQUE FINAL MONTRANT ITÉRATIF = 3, STABILITÉ COMPLÈTE	41
FIGURE III.7 : CAPTURE D'ÉCRAN DE L'INTERFACE TKINTER DU MODULE.....	42
FIGURE III.8 : COMPORTEMENT DU MODULE DE TEST LORS DE L'INJECTION D'UN POINT HORS-LIMITES ($X = 0.1, Y = 0.1$).....	43
FIGURE III.9 : IDENTIFICATION ET CLASSIFICATION EN TEMPS RÉEL D'UN POINT DE TEST ($X = 0.32, Y = 0.18$) ASSOCIÉ À LA CLASSE 4.....	44
FIGURE III.10 : IDENTIFICATION ET CLASSIFICATION EN TEMPS RÉEL D'UN POINT DE TEST ($X = 0.42, Y = 0.40$) ASSOCIÉ À LA CLASSE 1	45
FIGURE III.11 : INTERFACE D'ACCUEIL DU SYSTÈME ÉCOTRACE-IA	47
FIGURE III.12 : INTERFACE PRINCIPALE D'ÉVALUATION : SAISIE DES CRITÈRES ET ZONE DE RÉSULTATS	48
FIGURE III.13 : ÉVALUATION DU PRODUIT-TYPE 3 (GRADE C)	50
FIGURE III.14 : RÉSULTAT DE L'ÉVALUATION DU PRODUIT-TYPE 3 : PRODUIT INDUSTRIEL À CHAÎNE LOGISTIQUE MODÉRÉE (ÉCO-GRADE C)	50
FIGURE III.15 : ÉVALUATION DU PRODUIT-TYPE 2 GRADE B)	51
FIGURE III.16 : RÉSULTAT DE L'ÉVALUATION DU PRODUIT-TYPE 2 : PRODUIT DE CONSOMMATION COURANTE (ÉCO-GRADE B)	51
FIGURE III.17 : ÉVALUATION DU PRODUIT-TYPE 1 (GRADE A)	52

FIGURE III.18 : RÉSULTAT DE L'ÉVALUATION DU PRODUIT-TYPE 1 : PRODUIT À FAIBLE TRANSFORMATION (ÉCO-GRADE A)	52
FIGURE III.19 : ÉVALUATION DU PRODUIT-TYPE 5 (GRADE E)	53
FIGURE III.20 : RÉSULTAT DE L'ÉVALUATION DU PRODUIT-TYPE 5 : PRODUIT IMPORTÉ À EMPREINTE ENVIRONNEMENTALE CRITIQUE (ÉCO-GRADE E)	53

Liste des tableaux :

Chapitre I : Contexte du Projet et Enjeux de la Transition Écologique

TABEAU I.1 : COMPARAISON ENTRE ÉCOTRACE-IA ET LES LOGICIELS D'ACV CONVENTIONNELS [19] [20] [21].10

TABEAU I.2 : ANALYSE SWOT DE ÉCOTRACE-IA11

Chapitre II : État de l'Art sur les approches d'apprentissage automatique

TABEAU II.1 : SYNTHÈSE DES AVANTAGES ET DES LIMITES DES RÉSEAUX DE NEURONES ARTIFICIELS [35] 22

TABEAU II.2 : SYNTHÈSE DES AVANTAGES ET DES LIMITES DE L'ALGORITHME K-MEANS [38] 26

Chapitre III : Conception et Développement de l'Application

TABEAU III.1 : LISTE DES CRITÈRES ENVIRONNEMENTAUX RETENUS PAR DIMENSION..... 31

Chapitre IV : Étude de faisabilité et modèle économique

TABEAU IV.1 : BUSINESS MODEL CANVAS D'ÉCOTRACE-IA 58

TABEAU IV.2 : PROJECTIONS FINANCIÈRES D'ÉCOTRACE-IA (2027–2032) 63

Liste des Symboles et abréviations :

CO ₂	Dioxyde de carbone
CO ₂ eq	Équivalent dioxyde de carbone
kWh	Kilowattheure (unité d'énergie)
kWh/kg	Kilowattheure par kilogramme
MJ	Mégajoule
l/kg	Litre par kilogramme
m ³	Mètre cube
%	Pourcentage
$\vec{X}$	Vecteur d'entrée (scores bruts, dim. 16)
$\vec{W}$	Vecteur des poids normalisés
$\vec{Y}$	Vecteur pondéré ($\vec{Y} = \vec{X} \odot \vec{W}$)
$\odot$	Produit de Hadamard
$\ \cdot\ $	Norme euclidienne (distance L ₂)
μ_k	Centroïde du cluster k
J	Fonction objectif (inertie intra-cluster)
k	Nombre de clusters

K	Nombre total de classes ($K = 5$)
ε	Seuil de convergence
$^{\circ}\text{C}$	Degré Celsius
IA	Intelligence Artificielle
ML	Machine Learning
DL	Deep Learning
ANN	Artificial Neural Network
MLP	Multilayer Perceptron
K-Means	K-Moyennes (algorithme de clustering)
LogiClass	Application d'apprentissage K-Means
RT-KMC	Real-Time K-Means Classifier
ACV	Analyse du Cycle de Vie
LCA	Life Cycle Assessment
MACF	Mécanisme d'Ajustement Carbone aux Frontières
CBAM	Carbon Border Adjustment Mechanism
SEQUE-UE	Système d'Échange de Quotas d'Émission de l'UE
CSRD	Corporate Sustainability Reporting Directive

CDN	Contributions Déterminées au niveau National
PME	Petites et Moyennes Entreprises
ETI	Entreprises de Taille Intermédiaire
GUI	Graphical User Interface
ISO	Organisation Internationale de Normalisation
FAO	Organisation des Nations Unies pour l'Alimentation
GIEC	Groupe d'Experts Intergouvernemental sur le Climat
IPCC	Intergovernmental Panel on Climate Change
CCNUCC	Convention-Cadre des Nations Unies sur les Changements Climatiques
ALGEX	Agence Algérienne de Promotion du Commerce Extérieur
ANADE	Agence Nationale d'Appui au Développement de l'Entreprise
ANSEJ	Agence Nationale de Soutien à l'Emploi des Jeunes
MEER	Ministère de l'Environnement et des Energies Renouvelables
ReLU	Rectified Linear Unit
Tanh	Tangente Hyperbolique
SWOT	Strengths, Weaknesses, Opportunities, Threats
PV	Photovoltaïque

PET	Polyéthylène Téréphtalate
TER	Train Express Régional
TGV	Train à Grande Vitesse
pkl	Python Pickle (format de sérialisation)
npy	NumPy Array (format de données)
exe	Fichier exécutable Windows

Introduction générale :

Introduction générale

Le secteur agroalimentaire occupe une place centrale dans les problématiques environnementales actuelles en raison de sa dépendance aux ressources naturelles et de son empreinte carbone significative [1]. En Algérie, la modernisation des processus de production industrielle et agricole répond à une nécessité d'alignement avec les standards internationaux de durabilité [2], notamment les normes ISO 14040 et 14044 relatives à l'Analyse du Cycle de Vie (ACV) [3],[4].

L'évaluation de l'impact environnemental des produits reste cependant une procédure complexe et coûteuse, souvent inaccessible aux petites et moyennes entreprises (PME) ainsi qu'aux startups locales [5]. Le projet ÉcoTrace-IA a été initié pour répondre à ces contraintes en proposant une solution logicielle destinée à automatiser et fiabiliser la pré-évaluation environnementale des produits agroalimentaires par l'intégration de techniques d'intelligence artificielle.

La présente recherche s'articule autour de trois axes de développement :

- D'abord, la modélisation d'un système d'évaluation multicritère intégrant des indicateurs environnementaux tels que la nature de l'emballage, les sources d'énergie consommées, l'utilisation des ressources hydriques et les flux logistiques amont et aval.
- Ensuite, l'application de méthodes d'apprentissage automatique pour le traitement et la classification des données, en s'appuyant sur l'algorithme K-Means pour la segmentation non supervisée et les réseaux de neurones artificiels (MLP) pour la modélisation prédictive.
- Enfin, la conception de trois outils logiciels complémentaires : LogiClass, dédiée à l'apprentissage et à la persistance des centroïdes ; RT-KMC, assurant la classification en temps réel dans un espace bidimensionnel ; et ÉcoTrace-IA, permettant la pré-évaluation environnementale, l'attribution d'un Éco-Grade et la génération de recommandations d'optimisation à destination des industriels.

Le présent mémoire est structuré en quatre chapitres :

Introduction générale

- Le premier chapitre définit le cadre contextuel de l'étude, examine les enjeux de la transition écologique en Algérie et détaille les principes méthodologiques de l'ACV ainsi que les caractéristiques de la solution proposée.
- Le deuxième chapitre présente l'état de l'art relatif au machine learning, en abordant l'architecture des réseaux de neurones artificiels, les fonctions d'activation et les fondements mathématiques de la classification par K-Means.
- Le troisième chapitre expose la contribution technique et méthodologique, incluant la conception des trois applications logicielles, la formalisation des approches algorithmiques, le développement des interfaces graphiques sous Python/Tkinter et la validation expérimentale des modèles obtenus.
- Le quatrième chapitre présente l'étude de faisabilité et le modèle économique d'ÉcoTrace-IA, à travers l'élaboration du Business Model Canvas et des projections financières pluriannuelles sur la période 2027–2032.

Chapitre I : Contexte du Projet et Enjeux de la Transition Écologique

I.1.Introduction

Le monde est actuellement confronté à une transformation structurelle et irréversible des fondements de l'économie mondiale. Avec la fin de l'époque où seul le critère financier dictait la compétitivité à l'échelle internationale, une nouvelle ère s'est ouverte. Dans celle-ci, l'efficacité écologique devient un facteur crucial pour déterminer la participation sur le marché mondial. Le problème algérien se pose maintenant dans ce nouveau contexte marqué par une transformation radicale du système économique.

En effet, la mise en œuvre du Mécanisme d'Ajustement Carbone aux Frontières (MACF) par l'Union européenne constitue une menace directe et tangible pour la compétitivité des exportations algériennes. Les produits nationaux, insuffisamment décarbonés, risquent de se voir confrontés à une « barrière verte » susceptible d'éroder durablement leur position sur le marché européen — premier partenaire commercial de l'Algérie.

Face à l'urgence de cette échéance réglementaire, les plus hautes autorités nationales ont mandaté les centres de recherche afin d'analyser les implications de ce mécanisme et d'identifier les voies d'adaptation possibles [6]. C'est dans cette perspective que s'inscrit la présente étude, dont la problématique centrale repose sur l'articulation entre les contraintes imposées par ce nouveau cadre législatif et les solutions d'adaptation envisageables pour le tissu industriel algérien.

Le présent chapitre ne se limite pas à une simple contextualisation : il postule la nécessité d'une rupture épistémologique avec les approches conventionnelles. Après avoir analysé les mécanismes de la transition écologique et les fondements méthodologiques de l'Analyse du Cycle de Vie (ACV), nous introduisons notre projet innovant. Celui-ci prend la forme d'une startup application Windows s'appuyant sur l'intelligence artificielle pour transformer la menace du MACF en levier de souveraineté économique au service des produits algériens — convertissant ainsi une contrainte réglementaire en avantage concurrentiel durable.

I.2. Introduction à la Transition Écologique

Le passage vers une économie durable à faibles émissions qui respecte les besoins de la nature, depuis une économie à forte consommation de ressources, est appelé la transition écologique [7]. En réponse à l'urgence climatique, marquée par une hausse de la température mondiale

dépassant +1 °C, et confrontés au risque de franchir le seuil de +1,5 °C d'ici 2030 [8], une série de modifications radicales a été initiée.

Pour superviser cette dynamique à l'échelle mondiale, L'Accord de Paris (2015) représente le premier mécanisme international ayant force obligatoire visant à lutter contre le réchauffement climatique [9]. L'accord, qui compte 196 signataires, a pour objectif de contenir la hausse du réchauffement climatique à 2 °C par rapport aux niveaux préindustriels, avec une préférence pour le maintenir à 1,5 °C. Cet objectif sera réalisé en soumettant des contributions déterminées au niveau national (CDN) de plus en plus ambitieuses tous les cinq ans [9].

S'inscrivant dans cette lignée mondiale, l'Union européenne a renforcé son dispositif à travers le Pacte vert pour l'Europe (2019) [10], une des dispositions établit les obligations concernant la lutte contre le changement climatique. Parmi les outils législatifs mis en place pour réaliser l'objectif de neutralité carbone en Europe d'ici 2050 figurent la loi européenne sur le climat, le paquet « Fit for 55 ». Au cœur de cette stratégie se trouve le mécanisme d'ajustement carbone aux frontières (CBAM) [11]. Le CBAM est un dispositif mis en place pour intégrer le coût du carbone dans les importations, contribuant ainsi à éviter l'évasion du carbone et permettant aux producteurs européens de rester compétitifs tout en respectant le SEQE-UE [11].

Cependant, cette mutation ne se limite pas aux frontières européennes. En complément des initiatives de l'UE, d'autres dispositifs fiscaux tels que la taxonomie écologique [12], les systèmes d'étiquetage (éco score, étiquetage obligatoire) et les obligations de rapport non financier (CSRD) [13] émergent à l'échelle internationale. Ainsi, le respect des normes environnementales est devenu une obligation pour les exportateurs à l'échelle mondiale, y compris en Algérie [14]. Il s'agit désormais d'un élément essentiel de la gouvernance internationale du commerce et des exportations.

I.3. Problématique de l'Empreinte Carbone en Algérie

L'Algérie, en tant que puissance industrielle de la région, est aujourd'hui confrontée à un enjeu stratégique avec l'entrée en vigueur, à partir de janvier 2026, du Mécanisme d'ajustement carbone aux frontières (MACF/CBAM) imposé par l'Union européenne [6]. Il s'agit en effet d'une restriction importante pour les exportations algériennes, d'autant qu'elle consiste à imposer un système de tarification du carbone sur les marchandises entrant dans l'UE. Pourtant, les autorités algériennes, à travers leurs nouvelles orientations économiques, ambitionnent de transformer cet impératif environnemental en un véritable levier de souveraineté économique. Pour que cette transition économique vers une économie bas carbone soit réalisée, la traçabilité de l'empreinte carbone est un impératif absolu. Aujourd'hui, il est nécessaire pour les

entreprises algériennes de s'armer d'outils numériques innovants (comme le nôtre) pour évaluer, classer et certifier la durabilité environnementale de leurs produits, afin de conserver leur accès privilégié aux marchés internationaux.

I.4. Analyse de cycle de vie (ACV)

L'ACV est une méthode permettant de mesurer les effets environnementaux d'un produit tout au long de son cycle de vie (tels que le transport, l'emballage, l'agriculture, etc.). Non seulement c'est une méthode qui prend en compte chaque étape du cycle de vie d'un produit, mais elle considère également un certain nombre de questions environnementales majeures (comme le changement climatique, la qualité de l'eau et de l'air, l'effet sur les sols, etc.) plutôt que le seul climat [5]. Pour standardiser cette mesure, on s'appuie sur des normes internationales.

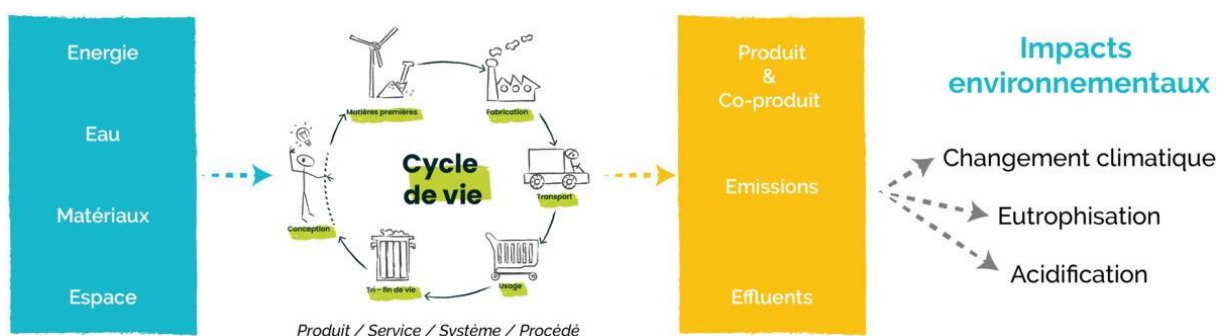


Figure I.1 : Approche multicritères et impacts environnementaux de l'ACV [15].

I.4.1. Phases et normes ISO 14040/14044

ISO 14040 :2006 → définit les principes généraux et le cadre méthodologique de l'ACV [3].

ISO 14044 :2006 → précise les exigences et lignes directrices pour la mise en œuvre concrète de l'ACV, notamment en matière de collecte de données, d'interprétation des résultats et de communication [4].

Conformément à ces critères, une évaluation environnementale complète s'effectue en quatre phases consécutives et récurrentes :

Phase 1 → Définition du périmètre et des buts : Cette phase déterminera l'étendue de l'évaluation (unité fonctionnelle, frontières du système) ainsi que ses objectifs [3],[4].

Phase 2 → Inventaire du cycle de vie : Cette phase implique une collecte minutieuse des données relatives aux flux de matières et d'énergie, tant en entrée qu'en sortie, dans le système analysé durant tout son cycle de vie : extraction des matières premières, fabrication, transport, distribution, utilisation et disposition finale [3],[4].

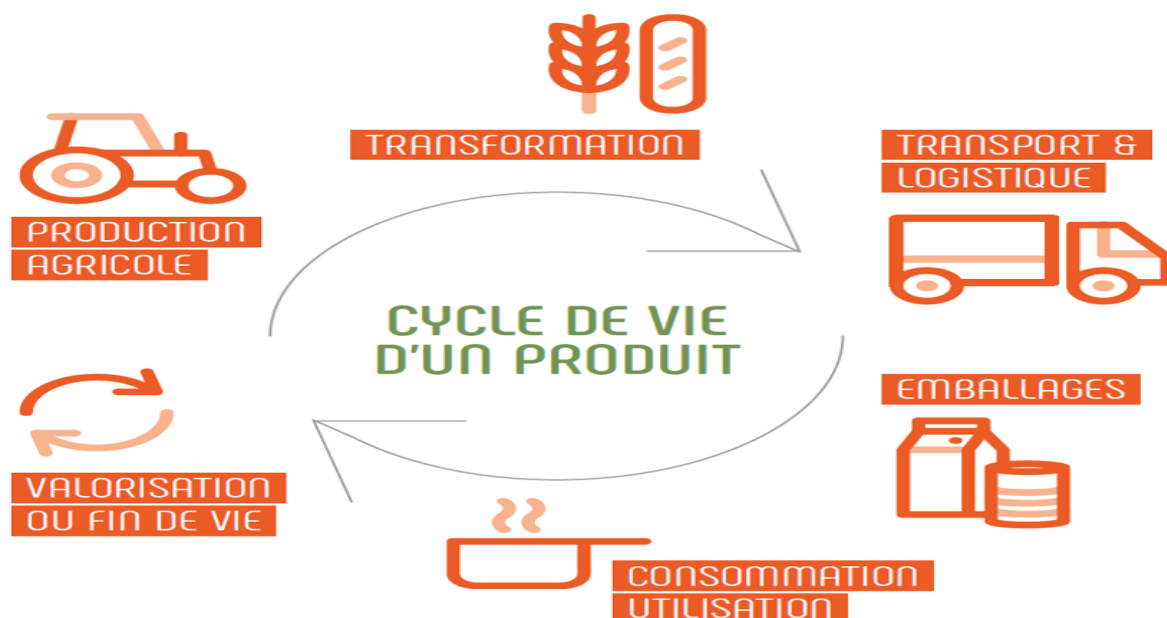


Figure I.2 : Les étapes du cycle de vie d'un produit [16].

Phase 3 → Évaluation de l'impact tout au long du cycle de vie : Les données d'inventaire sont exprimées en termes d'impacts environnementaux, permettant d'estimer la contribution du produit à des phénomènes tels que le réchauffement climatique, l'acidification, l'eutrophisation [3],[4].

Phase 4 → Interprétation : Les résultats seront analysés, les points environnementaux problématiques identifiés et des suggestions d'amélioration seront soumises aux décideurs [3],[4].

I.4.2 Indicateurs environnementaux : CO₂, eau et énergie

Une ACV n'est pas seulement fondée sur la simple mesure de l'émission de dioxyde de carbone. Elle recourt à une variété d'autres indicateurs environnementaux, permettant ainsi une approche plus globale de la portée écologique d'un produit. Parmi ceux-ci figurent notamment les suivants :

Empreinte carbone (kg CO₂eq) : Central à toute analyse d'ACV, cet indicateur représente la contribution d'un produit au changement climatique, à travers la mesure de tous les gaz à effets de serre émis tout au long de sa trajectoire de vie, équivalents à une certaine quantité de CO₂. C'est précisément ce dernier qui est visé par le concept de MACF [17].

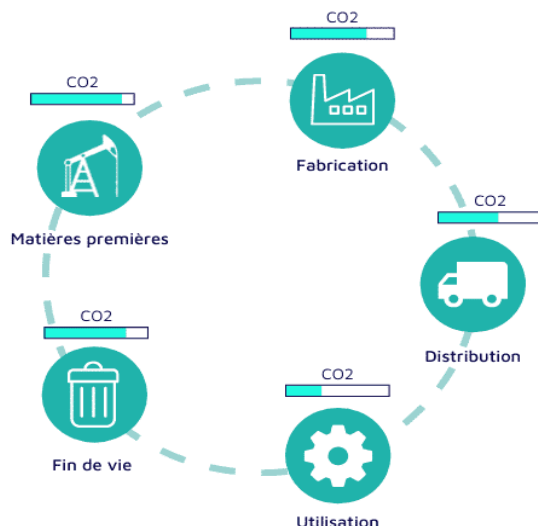


Figure I.3 : Répartition des émissions de CO2 par étape [18].

Empreinte eau (m³ ou litres) : L'impact eau correspond à l'utilisation totale d'eau douce tout au long de la vie du produit. Ce dernier est particulièrement sensible dans le cas de produits agroalimentaires, où la partie agro-environnementale est très consommatrice de ressources en eau [17].

Empreinte énergétique (MJ ou kWh) : Elle quantifie la demande cumulée en énergie primaire, en distinguant l'énergie renouvelable de l'énergie fossile. Cet indicateur est déterminant pour évaluer la dépendance d'un procédé de production aux combustibles fossiles et son potentiel de décarbonation [17].

Ces indicateurs constituent le socle de calcul de notre application **ÉcoTrace-IA**, qui les intègre comme variables structurantes. Il convient toutefois de préciser que notre plateforme ne sollicite pas directement la quantité de CO₂ émise en entrée — une donnée souvent complexe à obtenir pour un industriel. Elle repose sur une approche indirecte et prédictive basée sur un système de pondération multicritère : l'utilisateur renseigne les ressources consommées, classées par ordre d'importance selon leur contribution à l'impact environnemental (du pilier majeur comme l'Agriculture et l'Énergie, aux facteurs opérationnels tels que le Transport et l'Emballage). Le

moteur d'intelligence artificielle se charge ensuite de déduire l'empreinte carbone globale en croisant ces données pondérées avec les facteurs d'émission intégrés à sa logique algorithmique. C'est précisément cette capacité de traitement intelligent de la donnée, alliée à une pondération rigoureuse des impacts, qui distingue ÉcoTrace-IA des outils conventionnels et en fait un levier d'accessibilité pour les PME.

I.5. Présentation de la Solution Proposée (La Startup)

I.5.1. Vision et objectifs de la plateforme

La solution que nous proposons consiste en le développement d'une application desktop intelligente, spécifiquement conçue pour accompagner les entreprises algériennes dans leur transition écologique. Notre objectif principal est de mettre à leur disposition un outil capable de simuler, évaluer et classifier l'empreinte carbone de leurs produits, en amont de toute certification officielle. La plateforme les accompagne tout au long de ce processus, de l'évaluation initiale jusqu'à l'obtention d'une classification environnementale fiable et exploitable.

I.5.2. Proposition de valeur : évaluation et classification de l'IA

Notre proposition de valeur repose sur l'exploitation de l'Intelligence Artificielle pour offrir un service à forte valeur ajoutée, articulé autour de quatre axes fondamentaux :

Diagnostic préventif et amélioration en temps réel La plateforme identifie les « points noirs » environnementaux au sein de la chaîne de production, permettant à l'entreprise de comprendre précisément quels processus génèrent le plus d'émissions carbone, et d'agir en conséquence avant même de solliciter une certification externe. Cela transforme la contrainte réglementaire en un levier d'amélioration continue de la performance industrielle.

Substitution intelligente à l'expertise étrangère Plutôt que d'engager des frais élevés auprès de consultants internationaux spécialisés en bilan carbone, les industriels algériens disposeront d'un outil autonome leur fournissant des résultats d'une précision comparable. La plateforme démocratise l'accès à une expertise jusqu'ici réservée aux grandes entreprises disposant de ressources financières importantes.

Prévention de la taxe carbone (MACF) L'un des apports les plus stratégiques de la solution réside dans sa capacité à simuler, avant l'exportation, si un produit sera soumis à la taxe carbone européenne et à quel niveau. L'entreprise peut ainsi prendre des décisions éclairées : améliorer son processus, reformuler son produit, ou adapter sa stratégie commerciale — sans subir de mauvaises surprises aux frontières de l'Union européenne.

Accessibilité et souveraineté économique La plateforme est conçue pour être intuitive et simple d'utilisation, même pour des industriels ne disposant d'aucun expert interne en environnement. En accompagnant les produits « Made in Algérie » vers la conformité aux normes mondiales au moindre coût, elle contribue directement au renforcement de la compétitivité nationale et à la consolidation de la souveraineté économique du pays.

I.5.3. Positionnement par rapport aux solutions existantes

Le marché des logiciels d'ACV est aujourd'hui dominé par des solutions étrangères développées principalement en Europe et en Amérique du Nord. Les outils les plus répandus sont SimaPro (Pays-Bas), OpenLCA (Allemagne) et GaBi (Allemagne). Si ces solutions offrent une rigueur scientifique reconnue, elles présentent des limites structurelles majeures pour les entreprises algériennes.

Le tableau suivant établit une comparaison objective entre ces outils et notre solution ÉcoTrace-IA :

Critère	SimaPro	OpenLCA	GaBi	ÉcoTrace-IA
Coût	3 000–15 000 €/an	Gratuit (données payantes)	5 000–20 000 €/an	Abonnement accessible (PME)
Langue	Anglais/Français	Anglais	Anglais	Français
Adaptation au contexte algérien	Aucune	Aucune	Aucune	Native
Classification automatique	Non	Non	Non	Oui

Adaptabilité au contexte local	Non	Non	Non	Oui
Public cible	Experts ACV	Chercheurs	Grands groupes	PME/ETI algériennes
Pré-audit MACF	Non	Non	Non	Oui
Facilité d'utilisation	Complexe	Technique	Très complexe	Facilité d'utilisation

Tableau I.1 : Comparaison entre ÉcoTrace-IA et les logiciels d'ACV conventionnels [19][20][21].

I.6. Analyse SWOT de la Startup ÉcoTrace-IA

L'analyse SWOT permet d'évaluer de manière structurée le positionnement stratégique de ÉcoTrace-IA, en croisant ses dimensions internes (forces et faiblesses) avec les dynamiques de son environnement externe (opportunités et menaces).

	Positif	Négatif
Interne	· Solution IA innovante sans équivalent local · Coût d'accès réduit par rapport aux outils étrangers · Adaptation native au contexte algérien · Interface (français) · Approche pré-certification unique	· Dataset partiellement synthétique (données réelles limitées) · Équipe technique en phase de structuration · Notoriété à construire · Dépendance à la maintenance manuelle (déploiement et mises à jour sur site).
Externe	· Entrée en vigueur du MACF en janvier 2026 · Politique nationale de diversification des exportations	· Résistance au changement des industriels · Arrivée potentielle de concurrents étrangers localisés · Instabilité réglementaire

	· Soutien institutionnel (ALGEX, Ministères) · Marché B2B local peu concurrencé · Dispositifs de financement startups (ANADE, ANSEJ)	· Faible culture ACV au sein des entreprises algériennes
--	--	--

Tableau I.2 : Analyse SWOT de ÉcoTrace-IA.

Cette analyse confirme que la fenêtre d'opportunité est à la fois étroite et stratégique : l'entrée en vigueur du MACF génère une demande urgente à laquelle aucune solution locale ne répond à ce jour. ÉcoTrace-IA se positionne ainsi comme un pionnier sur le marché algérien. La stratégie de la startup consistera à transformer progressivement ses faiblesses en forces, notamment par l'acquisition de données industrielles réelles, et à neutraliser les menaces identifiées grâce à l'établissement de partenariats institutionnels solides et d'une présence commerciale précoce.

I.7.Conclusion

Ce premier chapitre a démontré que la pression réglementaire du MACF impose le développement d'une solution technologique locale adaptée. L'étude de la transition écologique, de l'Analyse du Cycle de Vie et du positionnement stratégique de ÉcoTrace-IA ont confirmé la pertinence et l'urgence du projet.

La valeur centrale d'ÉcoTrace-IA repose sur son moteur algorithmique d'intelligence artificielle, capable de classer automatiquement l'impact environnemental d'un produit agroalimentaire sur une échelle de A à E. Le Chapitre II sera consacré à la présentation détaillée de l'architecture retenue et de ses performances.

Chapitre II :

État de l'Art sur les approches d'apprentissage automatique.

II.1. Introduction

Le premier chapitre a présenté le contexte économique et les règles réglementaires qui justifient la création de ÉcoTrace-IA. Le présent chapitre en constitue le socle scientifique. La force d'un tel système dépend en effet bien plus de la pertinence de ses fondements théoriques que de ses performances brutes ; ce sont d'ailleurs ces bases qui ont guidé l'ensemble des choix algorithmiques réalisés dans ce travail.

L'apprentissage automatique rassemble aujourd'hui de nombreux paradigmes dont les propriétés varient selon la nature des données et la structure du problème. Loin de se résumer à la sélection d'un algorithme par défaut, une analyse précise de l'état de l'art montre que chaque choix découle d'une correspondance rigoureuse entre les exigences du problème et les caractéristiques de l'approche retenue.

Ce chapitre commence par un aperçu général de l'apprentissage automatique et de son évolution historique. Les concepts fondamentaux y sont développés de manière progressive, chacun préparant le suivant, jusqu'à la présentation des travaux existants dans le domaine de l'évaluation environnementale par intelligence artificielle. Enfin, la conclusion résume les décisions techniques retenues à travers un tableau comparatif argumenté.

II.2. Panorama du Machine Learning

II.2.1. IA, ML (apprentissage automatique) et DL (apprentissage profond)

Avant d'examiner les algorithmes retenus pour ÉcoTrace-IA, il convient de situer ces derniers dans leur cadre conceptuel global. Les termes intelligence artificielle, apprentissage automatique et apprentissage profond sont souvent employés de manière interchangeable dans le discours courant, alors qu'ils entretiennent en réalité une relation d'inclusion stricte, chacun constituant un sous-ensemble du précédent.

L'intelligence artificielle (IA) désigne, dans son sens le plus large, l'ensemble des techniques visant à doter les machines de capacités cognitives habituellement associées à l'intelligence humaine : raisonnement, perception, compréhension du langage, prise de décision. [22] Il s'agit d'un champ disciplinaire vaste, dont les premières formalisations remontent aux travaux fondateurs [22].

L'apprentissage automatique (*Machine Learning*, ML) constitue une branche de l'IA qui se distingue par son approche : plutôt que de programmer explicitement des règles, on fournit au système des données à partir desquelles il apprend à identifier des patterns et à généraliser. Le ML a connu un essor considérable à partir des années 1990, porté par la disponibilité croissante de données et l'augmentation de la puissance de calcul [22].

L'apprentissage profond (*Deep Learning*, DL) représente quant à lui une sous-famille du ML, fondée sur des architectures de réseaux de neurones à multiples couches. Sa particularité réside dans sa capacité à apprendre automatiquement des représentations hiérarchiques à partir de données brutes, sans ingénierie manuelle des caractéristiques [22]. C'est à partir de 2012, avec les travaux de Krizhevsky et al. Sur la reconnaissance d'images, que le Deep Learning s'est imposé comme paradigme dominant pour les tâches impliquant des données non structurées [22].

La figure suivante illustre cette relation d'imbrication :

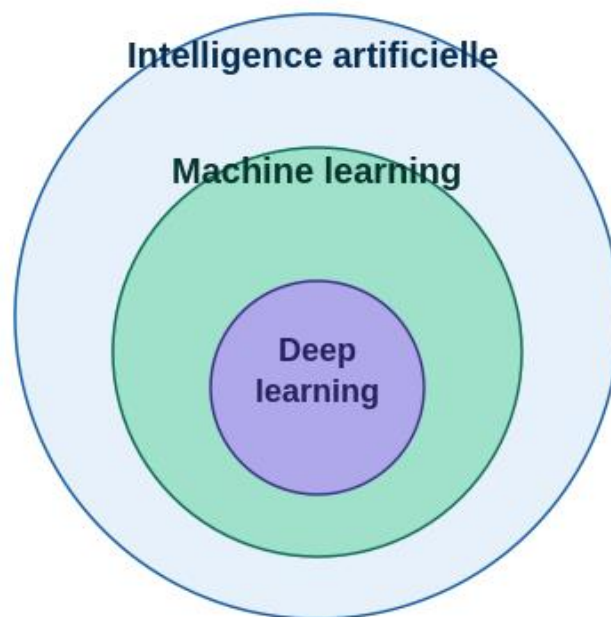


Figure II.1 : Relation d'inclusion entre IA, ML et DL [23].

Cette hiérarchie établie, il convient à présent de comprendre comment le ML a évolué vers le DL, et pourquoi cette évolution ne signifie pas que le DL est systématiquement préférable.

II.2.2. Du ML classique au DL : évolution et ruptures

L'évolution de l'apprentissage automatique est marquée par des transitions méthodologiques majeures. Les approches dites « classiques » — régression logistique, arbres de décision, forêts aléatoires, machines à vecteurs de support — ont longtemps dominé le domaine. [24] Efficaces sur des jeux de données de taille modérée aux variables bien définies, elles se distinguent par leur interprétabilité et leur faible coût de calcul.

La montée en puissance du Deep Learning s'est opérée sous l'effet conjugué de trois facteurs : l'explosion des volumes de données, l'essor des processeurs graphiques (GPU) et la maturité croissante des architectures de réseaux de neurones [24]. La rupture essentielle tient au traitement des caractéristiques : là où le ML classique requiert une intervention humaine experte pour les extraire et les sélectionner, les réseaux profonds les construisent d'eux-mêmes, directement à partir des données brutes [24].

Cette puissance a cependant un prix : une complexité computationnelle élevée et une forte dépendance aux grandes masses de données. Dans des contextes où les données sont structurées et de dimension modérée, des architectures plus légères restent non seulement suffisantes, mais souvent mieux adaptées ce que la section suivante illustrera à travers l'examen des paradigmes d'apprentissage retenus.

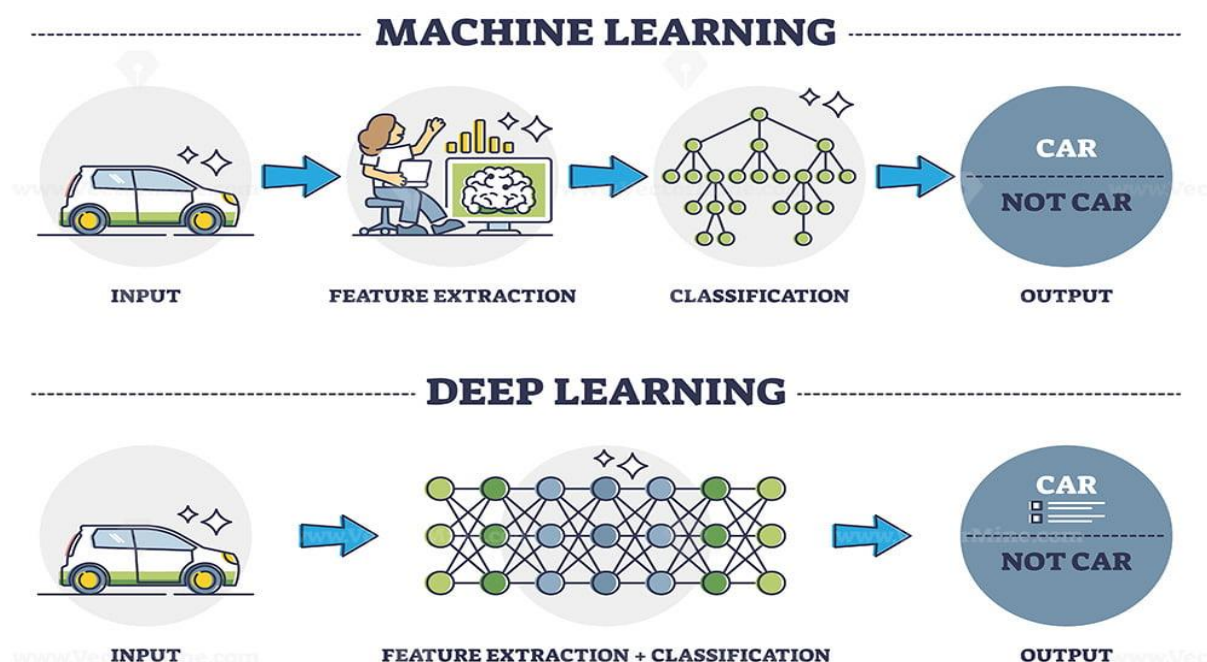


Figure II.2 : Comparaison du processus ML classique et Deep Learning [25].

II.2.3. Les paradigmes d'apprentissage : supervisé, non supervisé, par renforcement

L'apprentissage automatique se structure autour de trois grands paradigmes de modélisation. Le choix de l'une de ces approches est fondamentalement dicté par la nature des données disponibles à savoir la présence ou l'absence d'étiquettes ainsi que par l'objectif applicatif visé.

L'apprentissage supervisé s'appuie sur un ensemble de données préalablement étiquetées par un expert ou via un processus d'annotation. Chaque exemple en entrée étant associé à une cible connue, l'algorithme cherche à approximer la fonction de correspondance sous-jacente. Pour ce faire, il ajuste ses paramètres internes de façon à minimiser l'écart entre ses prédictions et les étiquettes réelles [26]. Ce paradigme englobe deux tâches majeures : la classification, qui consiste à attribuer une catégorie discrète à une observation, et la régression, qui vise à prédire une valeur numérique continue. Les arbres de décision, les machines à vecteurs de support (SVM) et les réseaux de neurones multicouches (MLP) en sont les représentants les plus répandus.

L'apprentissage non supervisé, à l'inverse, traite des données brutes ne disposant d'aucune étiquette prédéfinie. L'algorithme explore de manière autonome l'espace de description afin d'y découvrir des structures cachées, des corrélations ou des régularités intrinsèques, le plus souvent en mesurant des distances mathématiques de similarité entre les exemples [26]. Sa principale application est le partitionnement (regroupement homogène), qui se divise en deux approches : le partitionnement strict, où chaque donnée appartient strictement à un seul groupe, et le partitionnement flou, où chaque donnée est associée à une probabilité d'appartenance à plusieurs grappes. Ce paradigme s'avère indispensable lorsque les données ne disposent d'aucune classification officielle préalable.

L'apprentissage par renforcement repose sur un principe d'apprentissage dynamique et itératif. Un agent autonome interagit avec un environnement changeant : chaque action entreprise produit une rétroaction sous forme de signal de récompense ou de pénalité, ce qui guide et optimise progressivement son comportement au fil du temps [26]. Bien que performante en robotique ou pour les systèmes de décision séquentielle, cette logique n'est pas adaptée aux problématiques de classification statique à partir de données tabulaires structurées.

La figure suivante synthétise l'articulation de ces trois paradigmes ainsi que leurs domaines d'application respectifs :

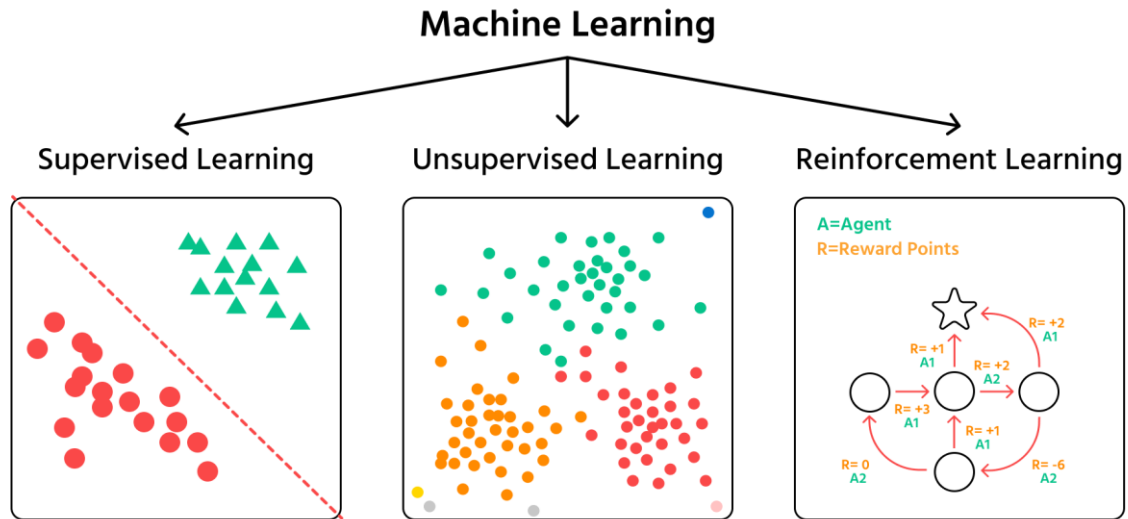


Figure II.3 : Les trois paradigmes de l'apprentissage automatique [27].

II.3. Le neurone artificiel _ le perceptron

II.3.1. Du neurone biologique au neurone artificiel

L'idée fondatrice des réseaux de neurones artificiels est née d'une observation simple : le cerveau humain, composé de milliards de neurones interconnectés, est capable d'apprentissage, de reconnaissance et de décision [28]. Les chercheurs ont donc cherché à reproduire ce mécanisme de manière mathématique.

Le neurone biologique reçoit des signaux via ses dendrites, les intègre dans le corps cellulaire, puis transmet une réponse via son axone si le signal cumulé dépasse un certain seuil [28]. C'est ce principe de seuil d'activation qui a inspiré la conception du neurone artificiel.

Le neurone artificiel, introduit par McCulloch et Pitts en 1943, reproduit ce mécanisme sous forme d'une opération mathématique simple [28]. Il reçoit un vecteur d'entrées $x_1, x_2, \dots, x_n$, associe à chacune un poids w_i traduisant son importance relative, calcule leur somme pondérée, y ajoute un terme de biais b , puis applique une fonction d'activation f pour produire une sortie y [22] :

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

Ce qui peut s'écrire sous sa forme développée [22] :

$$y = f(w_1x_1 + w_2x_2 + \dots + w_nx_n + b)$$

La figure suivante illustre cette analogie :

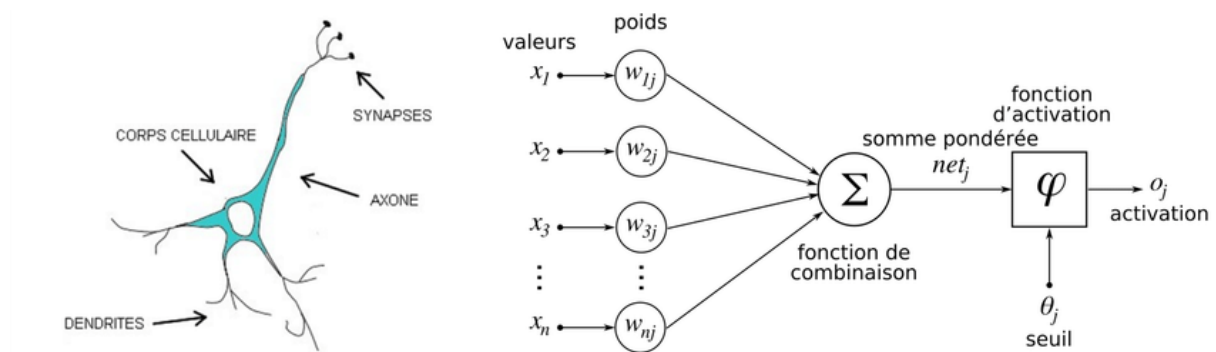


Figure II.4 : Analogie structurelle entre le neurone biologique et le neurone artificiel [29].

II.3.2. Rôle des poids, du biais et de la fonction d'activation

Les poids w_i traduisent l'importance de chaque entrée dans la décision du neurone. Plus un poids est élevé, plus l'entrée associée pèse dans le calcul ; à l'inverse, un poids proche de zéro la rend quasiment invisible. C'est en ajustant ces poids progressivement, à chaque itération d'entraînement, que le réseau apprend [22].

Le biais b joue un rôle différent : il décale le seuil de décision du neurone indépendamment de ce qu'il reçoit en entrée. Concrètement, il permet au neurone de s'activer même si toutes ses entrées sont nulles ce qui donne au modèle une marge d'adaptation bien plus large face à des données variées [22].

La fonction d'activation f est ce qui donne au réseau sa véritable puissance. Sans elle, empiler plusieurs couches de neurones ne changerait rien : le résultat resterait une simple combinaison linéaire des entrées, incapable de capturer des relations complexes. C'est elle qui introduit la non-linéarité nécessaire pour modéliser des phénomènes réels [22]. Ses différentes formes et leurs cas d'usage font l'objet de la section suivante.

II.4. Fonctions d'activation

La fonction d'activation détermine si un neurone doit transmettre un signal ou non, et sous quelle forme. Son choix est loin d'être anodin : il conditionne directement la capacité du réseau

à apprendre, sa vitesse de convergence et la qualité de ses prédictions. Trois fonctions sont particulièrement répandues dans les architectures modernes : ReLU, Sigmoidé et Softmax.

II.4.1. La fonction ReLU (Rectified Linear Unit)

La fonction ReLU est aujourd'hui la plus utilisée dans les réseaux de neurones profonds. Elle est définie par :

$$f(x) = \max(0, x)$$

Elle laisse passer les valeurs positives telles quelles et annule les valeurs négatives [30]. Cette simplicité lui confère un faible coût de calcul, une convergence rapide et une absence de saturation pour les valeurs positives — contrairement à la sigmoïde qui souffre du problème de disparition du gradient. Sa principale limite est le phénomène *dying ReLU* : un neurone recevant systématiquement des valeurs négatives cesse d'apprendre. Des variantes comme *Leaky ReLU* ont été proposées pour y remédier [30].

II.4.2. La fonction Sigmoidé

La fonction Sigmoidé est définie par :

$$f(x) = \frac{1}{1 + e^{-x}}$$

Elle transforme toute valeur réelle en un nombre compris entre 0 et 1, ce qui la rend naturellement adaptée aux problèmes de classification binaire [22]. Son interprétation est probabiliste : plus l'entrée est élevée, plus la sortie se rapproche de 1.

Sa limite principale apparaît dans les réseaux profonds : ses gradients deviennent très faibles aux extrémités de la courbe, ce qui ralentit considérablement l'apprentissage par rétropropagation — c'est le problème du *Vanishing Gradient* [22]. Elle est donc aujourd'hui réservée aux couches de sortie des problèmes binaires.

II.4.3. La fonction Softmax

La fonction Softmax est définie par :

$$f(z_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$$

Elle convertit un vecteur de scores bruts en une distribution de probabilités : chaque sortie est comprise entre 0 et 1, et leur somme est égale à 1 [22]. La classe prédite correspond à la probabilité maximale.

II.4.4. La fonction Tangente Hyperbolique (Tanh)

La fonction tangente hyperbolique est définie par :

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Elle transforme toute valeur réelle en un nombre compris dans l'intervalle $[-1, 1]$, avec une sortie nulle pour une entrée nulle [22]. Contrairement à la sigmoïde, sa symétrie par rapport à l'origine lui confère un avantage lors de l'entraînement : les gradients sont mieux centrés, ce qui accélère la convergence.

Elle reste cependant sujette au problème du Vanishing gradient pour des valeurs d'entrée très élevées ou très faibles, où la courbe sature et les gradients deviennent quasi nuls [22]. Elle est néanmoins préférée à la sigmoïde pour les couches cachées, notamment lorsque les données d'entrée sont normalisées autour de zéro.

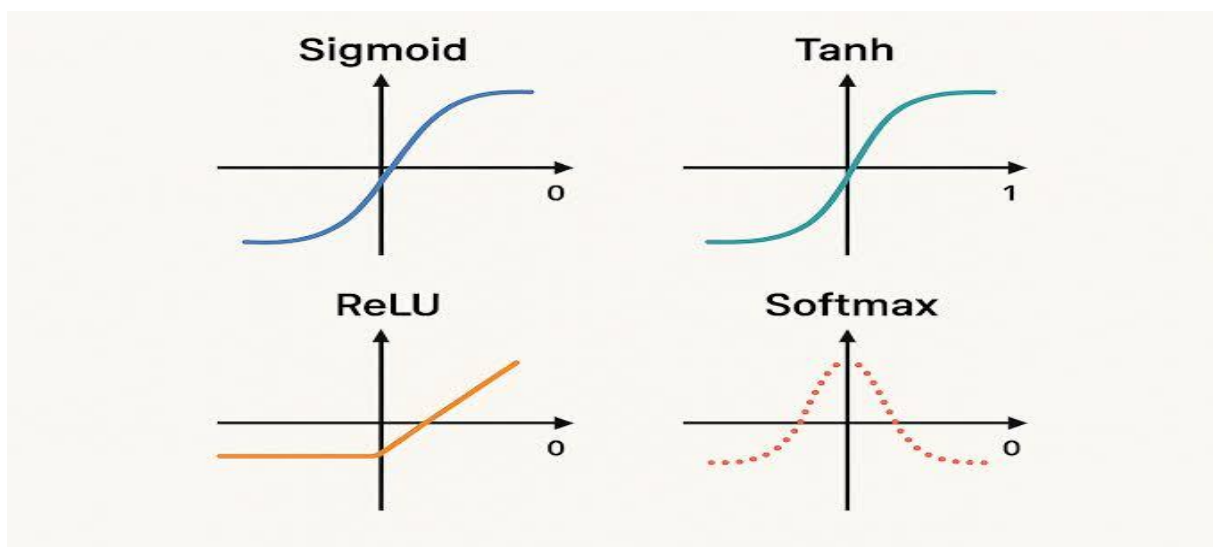


Figure II.5 : Représentation graphique des fonctions d'activation

(Sigmoïde, Tanh, ReLU, Softmax) [31].

II.5. Réseaux de neurones artificiels (ANN / MLP)

II.5.1. Architecture générale

Un réseau de neurones artificiels (Artificial Neural Network, ANN), également appelé perceptron multicouche (Multilayer Perceptron, MLP), est une extension directe du neurone isolé présenté à la section précédente. L'idée fondatrice est simple : en interconnectant un grand nombre de neurones organisés en couches successives, il devient possible de modéliser des relations d'une complexité bien supérieure à ce qu'un neurone unique ou une fonction linéaire pourrait capturer [22].

Un MLP est structuré en trois types de couches :

La couche d'entrée reçoit les données brutes. Chaque nœud y correspond à une variable du vecteur de caractéristiques. Elle ne réalise aucun calcul ; son rôle est purement de distribuer les valeurs aux couches suivantes [22].

Les couches cachées (hidden layers) constituent le cœur du réseau. Chaque neurone y calcule une combinaison pondérée de toutes les sorties de la couche précédente, puis applique une fonction d'activation non linéaire. C'est à ce niveau que s'opèrent la transformation et l'abstraction progressive des représentations. Le nombre de couches cachées et le nombre de neurones par couche sont des hyperparamètres déterminant la capacité expressive du modèle [26].

La couche de sortie produit la prédiction finale. Sa structure dépend de la tâche : un seul neurone à activation sigmoïde pour la classification binaire, ou k neurones à activation Softmax pour la classification multiclassée à k catégories [26].

La figure suivante illustre cette architecture :

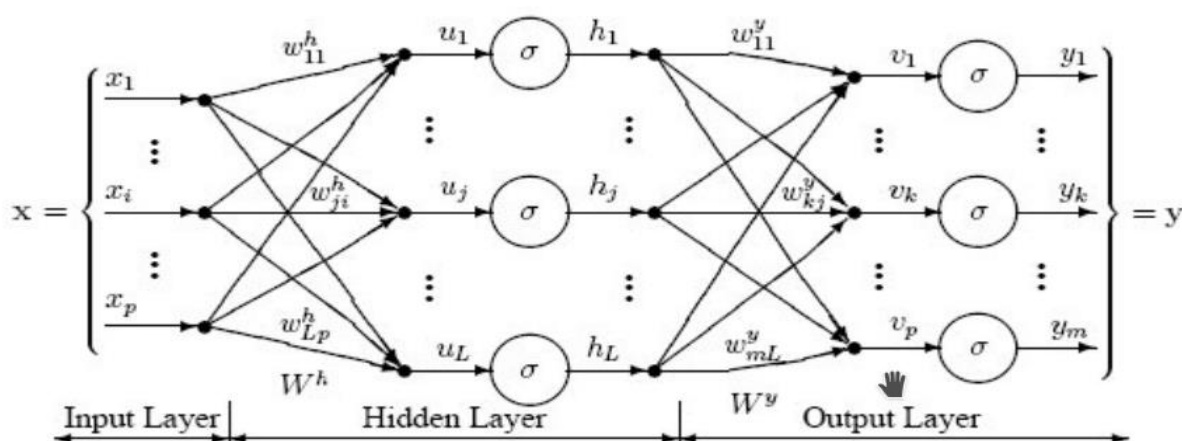


Figure II.6 : Architecture d'un MLP à deux couches cachées [32].

II.5.2. Capacité d'approximation universelle

Le théorème d'approximation universelle (Cybenko, 1989 ; Hornik, 1991) établit qu'un MLP comportant une seule couche cachée de taille suffisante peut approximer n'importe quelle fonction continue sur un domaine compact, avec une précision arbitrairement élevée [33]. Ce résultat théorique justifie l'utilisation des réseaux de neurones pour des tâches de régression ou de classification complexes, même en l'absence de connaissance a priori sur la forme de la fonction cible [34]. En pratique, des réseaux plus profonds mais moins larges permettent souvent d'atteindre le même niveau d'approximation avec un nombre de paramètres moindre [34].

Le tableau suivant présente une synthèse comparative des principaux avantages et limites des réseaux de neurones artificiels :

Avantages	Limites
Capacité à modéliser des relations non linéaires complexes	Nécessite de grands volumes de données d'entraînement
Apprentissage automatique des représentations pertinentes	Temps d'entraînement élevé selon la profondeur du réseau

Avantages	Limites
Applicable à de nombreux domaines (vision, NLP, classification)	Modèle de type "boîte noire", faible interprétabilité
Haute performance sur des données de grande dimension	Sensible au choix des hyperparamètres (η , couches, neurones)
Extensible aux architectures profondes (Deep Learning)	Risque de surapprentissage (Overfitting) sans régularisation

Tableau II.1: Synthèse des avantages et des limites des réseaux de neurones artificiels [35].

II.6. Processus d'apprentissage et optimisation

II.6.1. Propagation avant et rétropropagation

L'apprentissage d'un réseau de neurones se déroule en deux phases alternées. La propagation avant (forward propagation) fait traverser les données de la couche d'entrée jusqu'à la couche de sortie, produisant une prédiction $\hat{y}$ comparée à la cible réelle y via une fonction de coût $L(y, \hat{y})$ [34]. La rétropropagation (Backpropagation), formalisée par Rumelhart, Hinton et Williams (1986), calcule le gradient de cette fonction de coût par rapport à chaque paramètre via la règle de dérivation en chaîne, permettant la mise à jour des poids [36] :

$$w_i \leftarrow w_i - \eta \cdot \frac{\partial L}{\partial w_i}$$

Où η est le taux d'apprentissage contrôlant l'amplitude des mises à jour [36].

II.6.2. Fonctions de coût et optimiseurs

Le choix de la fonction de coût dépend de la tâche : l'entropie croisée pour la classification (binaire ou multiclasse), et l'erreur quadratique moyenne (MSE) pour la régression [37].

Concernant l'optimisation, la descente de gradient stochastique (SGD) met à jour les paramètres par mini-batch, offrant une convergence rapide mais sensible au taux d'apprentissage [37]. L'optimiseur Adam (Kingma & Ba, 2015) lui est généralement préféré : il adapte individuellement le taux d'apprentissage de chaque paramètre via les estimations des premier et seconds moments du gradient, assurant une convergence plus robuste [37].

II.7. Apprentissage Non Supervisé : K-Means

II.7.1. Principe de l'algorithme

L'algorithme K-Means, formalisé par Lloyd en 1957 (publié en 1982), est la méthode de partitionnement la plus répandue en apprentissage non supervisé [38]. Son objectif est de partitionner un ensemble de n observations en k groupes homogènes en minimisant l'inertie intra-classe :

$$\mathcal{J} = \sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - \mu_j\|^2$$

Où μ_j est le centroïde de la classe C_j . L'algorithme alterne deux étapes jusqu'à convergence [38] :

- Affectation : chaque observation est assignée au centroïde le plus proche.
- Mise à jour : chaque centroïde est recalculé comme la moyenne de toutes les observations du cluster.

La convergence est garantie en un nombre fini d'itérations, bien que la solution obtenue puisse correspondre à un minimum local [38].

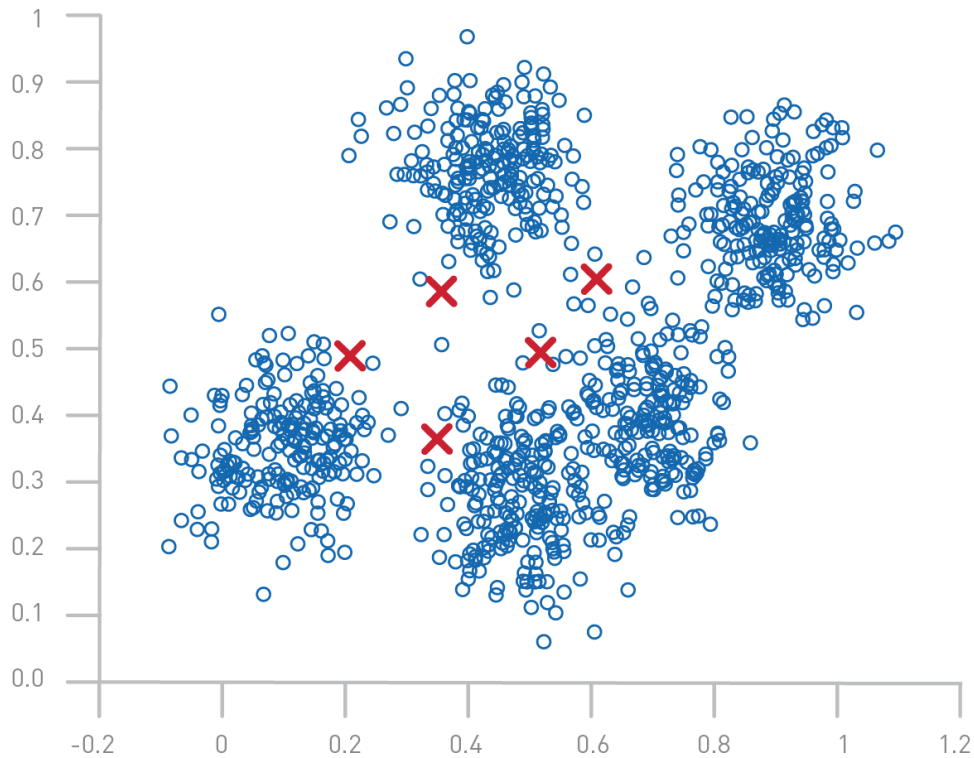


Figure II.7 : État initial des centroïdes K-Means avant convergence [39].

II.7.2. Initialisation et choix de k

La qualité du résultat dépend fortement de l'initialisation. Une initialisation aléatoire peut conduire à des solutions sous-optimales ; la méthode K-Means++ (Arthur & Vassilvitskii, 2007) y remédie en espaçant les centroïdes initiaux de manière à couvrir l'espace des données, réduisant ainsi l'inertie finale de façon significative [40].

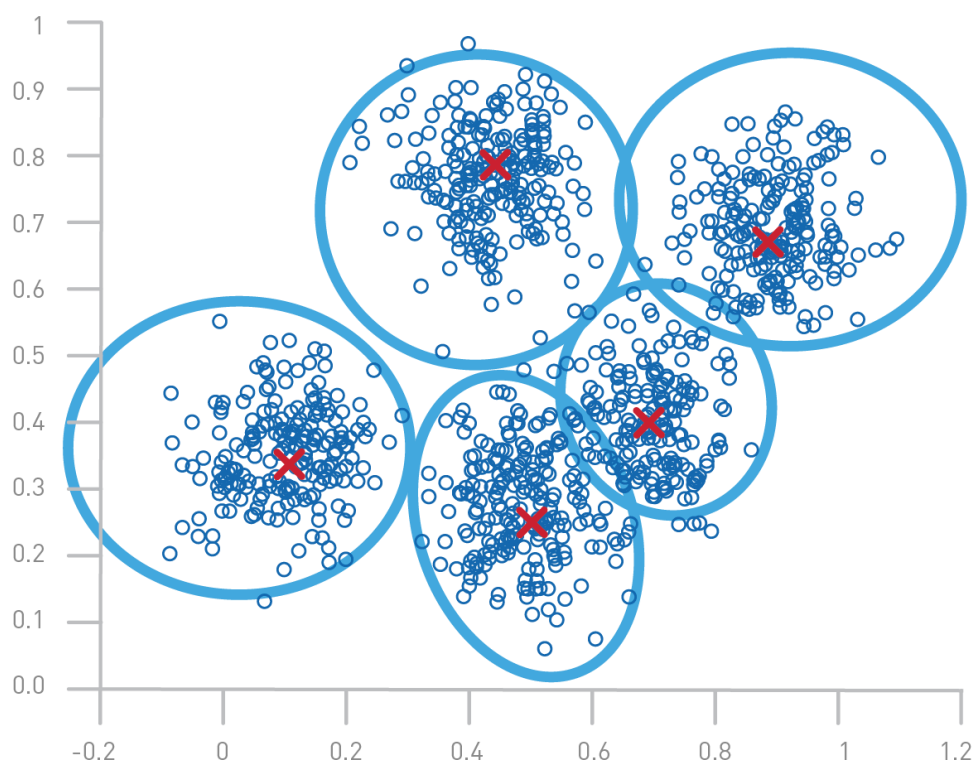


Figure II.8 : Convergence finale de l'algorithme K-Means — partitionnement en quatre clusters distincts [39].

Le nombre k n'étant pas connu a priori, deux critères permettent de le déterminer : la méthode du coude (elbow method), qui identifie le point à partir duquel l'ajout d'une classe n'apporte qu'un gain marginal sur l'inertie, et l'indice de silhouette, qui mesure la cohésion intra-classe et la séparation inter-clusters, fournissant un critère de sélection plus rigoureux [40].

Le tableau suivant synthétise les principaux avantages et limites de l'algorithme K-Means :

Avantages	Limites
Simple à implémenter et à interpréter	Sensible à l'initialisation des centroïdes
Complexité linéaire, rapide à l'exécution	Nécessite de fixer k à l'avance

Avantages	Limites
Résultats facilement explicables	Convergence vers un minimum local possible
Adapté aux données de grande dimension	Sensible aux valeurs aberrantes (outliers)
Performant sur de grands volumes de données	Suppose des classes de forme sphérique

Tableau II.2: Synthèse des avantages et des limites de l'algorithme K-Means [38].

II.8. Évaluation des modèles et régularisation

L'entraînement d'un modèle d'apprentissage automatique ne se limite pas à minimiser les erreurs sur les données d'entraînement. Un modèle performant doit être capable de généraliser sur des données inédites. Il existe deux phénomènes qui peuvent empêcher cette généralisation : le sous-apprentissage et le surapprentissage. Le sous-apprentissage se produit lorsque le modèle est trop simple pour comprendre la structure des données. Cela entraîne des erreurs élevées sur les données d'entraînement et sur les données de test. Le surapprentissage est le phénomène inverse : le modèle mémorise les données d'entraînement, mais perd la capacité de généraliser. Cela signifie que le modèle a une erreur faible lors de l'entraînement, mais une erreur élevée lors du test. Pour évaluer les performances d'un modèle, on utilise plusieurs métriques, selon la nature de la tâche. La précision mesure la proportion d'observations correctement classifiées. Elle est utilisée pour les problèmes de classification équilibrés. La matrice de confusion fournit une vision détaillée des erreurs de classification. Elle distingue les vrais positifs, les vrais négatifs, les faux positifs et les faux négatifs pour chaque classe. La validation croisée permet d'estimer la performance réelle du modèle en le testant sur différentes partitions des données. Pour éviter le surapprentissage, on utilise plusieurs techniques de régularisation. La régularisation L₁ pénalise les poids élevés dans le modèle. Le Dropout désactive aléatoirement une fraction des neurones lors de l'entraînement, ce qui force le modèle à apprendre des représentations robustes. L'arrêt précoce interrompt l'entraînement lorsque la performance sur un ensemble de validation commence à se dégrader. Ces techniques aident à améliorer la généralisation des modèles d'apprentissage automatique [22],[23]. Ces principes

d'évaluation, bien que centraux pour les approches d'apprentissage supervisé, ne s'appliquent pas directement à l'approche déterministe finalement retenue (cf. III.4.2), mais ont guidé l'analyse comparative des méthodes envisagées.

II.9. Revue des travaux connexes

L'utilisation de l'intelligence artificielle pour évaluer l'impact environnemental des produits agroalimentaires est un domaine de recherche qui évolue rapidement. De nombreux travaux ont exploré des approches algorithmiques visant à automatiser la classification ou la prédiction de l'impact écologique. Des études récentes ont montré qu'il est possible d'utiliser des réseaux de neurones artificiels pour estimer l'empreinte carbone de produits industriels [41] en utilisant des données sur les processus de production, les modes de transport et les types d'emballage. Ces travaux démontrent l'importance d'utiliser des critères quantifiés pour alimenter un modèle de classification automatique. D'autres recherches ont utilisé l'apprentissage automatique pour analyser le cycle de vie des produits. Les méthodes de regroupement, comme K-Means, permettent de regrouper des produits en fonction de leur profil environnemental [41] sans nécessiter de données préalablement étiquetées. Cela est particulièrement utile dans des contextes industriels où les bases de données certifiées sont rares, comme c'est le cas pour le secteur agroalimentaire algérien. Les outils commerciaux existants, tels que SimaPro ou OpenLCA, reposent sur des approches manuelles et ne proposent pas d'automatisation par intelligence artificielle. Ils sont destinés aux experts en analyse du cycle de vie qui disposent de données d'inventaire détaillées, ce qui les rend inaccessibles aux petites et moyennes entreprises. ÉcoTrace-IA se distingue de ces outils par son interface simplifiée, son moteur de classification automatique et son adaptation au contexte algérien. Ces travaux confirment deux orientations clés d'ÉcoTrace-IA : l'évaluation environnementale multicritère et l'utilisation d'un algorithme de regroupement pour la classification automatique [41][5] des produits en fonction de leur impact écologique.

II.10. Conclusion

Ce chapitre a établi les fondements théoriques sur lesquels repose le moteur algorithmique de l'application de classification des données. Après avoir situé l'apprentissage automatique dans son cadre conceptuel global, les mécanismes essentiels du neurone artificiel, des fonctions d'activation et de l'architecture MLP ont été présentés. Le processus d'apprentissage par

rétropropagation, les critères d'évaluation des modèles et les principales techniques de régularisation ont ensuite été exposés.

Bien que les architectures ANN/MLP soient présentées dans ce chapitre à des fins d'état de l'art comparatif, leur complexité opérationnelle (besoin de données labellisées, coût computationnel) a orienté le choix méthodologique vers l'algorithme K-Means. Ce dernier est d'abord exploité sous sa forme itérative complète, avant d'être adapté sous une forme déterministe par centroïdes analytiques pour les besoins spécifiques du système final, comme le détaille le Chapitre III.

Chapitre III : Conception et Développement de l'Application

Chapitre III : Conception et Développement de l'Application

III.1. Introduction

Ce chapitre détaille les étapes de conception du système ÉcoTrace-IA, de la définition de son architecture initiale jusqu'à l'implémentation d'un algorithme de classification environnementale. Le système vise à fournir aux petites et moyennes entreprises du secteur agroalimentaire algérien un outil d'évaluation structuré permettant l'attribution automatisée d'un Éco-Grade, sur une échelle de A à E, à partir de 16 critères environnementaux quantifiés.

La structure de ce chapitre suit le développement technique du projet. Après une présentation de l'architecture globale, le système de notation est exposé. La démarche méthodologique ayant conduit au choix de l'algorithme K-Means, à la suite de phases expérimentales, est ensuite explicitée. Enfin, les résultats de la validation ainsi que les caractéristiques de l'interface utilisateur développée sous Tkinter sont présentés.

Cette organisation rend compte de la nature itérative du travail de recherche, marqué par des ajustements dictés par les contraintes techniques et les orientations de la supervision. L'objectif demeure la proposition d'une solution technique adaptée aux spécificités opérationnelles des PME agroalimentaires dans le contexte algérien.

III.2. Architecture globale

Le système ÉcoTrace-IA repose sur une architecture logicielle locale et autonome, fonctionnant sans connexion réseau. La structure comprend trois modules principaux : une unité de saisie de données environnementales, un moteur de classification fondé sur l'algorithme K-Means et une interface graphique réalisée avec Tkinter. L'ensemble est implémenté en Python, choisi pour ses bibliothèques scientifiques et sa compatibilité avec les environnements de recherche.

III.2.1. Processus opérationnel

Le fonctionnement d'ÉcoTrace-IA suit un enchaînement linéaire. L'utilisateur soumet les caractéristiques environnementales d'un produit via un formulaire structuré. Chaque critère est converti en valeur numérique selon une grille de notation prédéfinie. Ces scores sont ensuite traités par le moteur de classification qui calcule un score moyen et le compare aux centroïdes des cinq classes environnementales. Le système génère en sortie un grade environnemental et un indice de confiance exprimé en pourcentage.

III.2.2. Flux de données

Le flux de données suit un protocole de traitement rigoureux, conçu pour garantir la transparence de l'évaluation environnementale. Les entrées se composent de 16 critères environnementaux renseignés par l'utilisateur via une interface structurée.

Chaque sélection est convertie en un score numérique (échelle de 1 à 5). Pour assurer une normalisation adaptative, les critères non applicables sont affectés d'une valeur nulle et exclus du calcul, permettant ainsi une évaluation juste quelle que soit la nature du produit.

Le moteur de classification procède ensuite par une inférence vectorielle : il calcule le score moyen pondéré du produit, lequel est comparé aux centroïdes théoriques (calculés mathématiquement pour chaque Éco-Grade de A à E). Le système identifie alors la classe la plus proche par le calcul de la distance euclidienne minimale. Enfin, l'application génère le grade final accompagné d'un indice de confiance, qui quantifie la proximité du produit par rapport au centre de gravité de la classe identifiée.

III.2.3. Justifications techniques

Les choix technologiques ont été dictés par des impératifs de déploiement, de transparence algorithmique et de portabilité.

Python a été retenu comme langage principal pour ses capacités de traitement de données.

NumPy assure les calculs matriciels nécessaires à l'algorithme K-Means.

Tkinter est utilisé pour l'interface graphique en raison de son intégration native au langage, évitant ainsi des dépendances additionnelles.

Matplotlib permet la génération des visualisations.

L'absence de frameworks externes complexes est une décision méthodologique visant à garantir la compatibilité du système avec les ressources informatiques limitées de certaines structures industrielles, notamment les PME.

III.3. Modélisation du système de notation environnementale

Le cadre analytique d'ÉcoTrace-IA repose sur un système de notation structuré à partir de critères quantifiés. Cette méthodologie permet d'évaluer les dimensions de l'impact

environnemental d'un produit agroalimentaire, couvrant l'ensemble de son cycle, de la phase de production au traitement final de l'emballage.

III.3.1. Définition des critères environnementaux

L'évaluation s'appuie sur 16 critères sélectionnés pour leur représentativité dans le secteur agroalimentaire et leur conformité aux référentiels d'analyse environnementale. Ces indicateurs sont regroupés en six catégories principales : emballage, énergie, transport, déchets, eau et production agricole. Les critères retenus pour chaque dimension sont les suivants :

Dimension	Critères retenus
Emballage	Type d'emballage, Poids de l'emballage, Fin de vie de l'emballage
Énergie	Type d'énergie, Consommation d'énergie
Transport	Type de transport, Distance de transport (amont, aval)
Déchets	Type de déchets, Taux de mise en décharge
Eau	Source d'eau, Consommation d'eau
Production	Type d'agriculture, Classe alimentaire, Type de sol

Tableau III.1 : Liste des critères environnementaux retenus par dimension.

III.3.2. Échelle d'évaluation et modalités de Notation

Chaque indicateur est transcrit sur une échelle ordinale allant de 1 (impact minimal) à 5 (impact maximal). Les grilles de notation, élaborées à partir de données sectorielles, distinguent deux types de variables : les variables qualitatives, où chaque modalité correspond à une valeur fixe, et les variables quantitatives, dont le score dépend de seuils prédéfinis. Par exemple, le critère relatif au type d'énergie attribue la valeur 1 aux sources renouvelables, telles que l'éolien ou l'hydraulique, et la valeur 5 aux combustibles fossiles à forte empreinte carbone. Pour la consommation d'eau, le score 1 correspond à un volume inférieur à 100 l/kg, tandis que le score 5 est atteint au-delà de 8000 l/kg.

III.3.2.1. Logique de conception de l'échelle

Le choix d'une échelle de notation de 1 à 5 n'est pas arbitraire. Il est rigoureusement aligné sur les 5 niveaux de performance environnementale (Éco-Grades A à E) que nous avons définis.

Cette symétrie entre l'échelle de notation et les grades de sortie garantit la cohérence globale du modèle et facilite l'interprétation des résultats par l'utilisateur final. Nous avons classé chaque modalité d'entrée (type de transport, énergie, etc.) selon cette logique de gradation, allant du meilleur impact (1) au plus mauvais (5), en nous appuyant sur des références sectorielles pour assurer la justesse de chaque valeur.

III.4. Évolution méthodologique : Modélisation par Réseaux de Neurones Artificiels (ANN)

III.4.1. Architecture du Perceptron Multicouche (MLP)

La première itération du système de classification environnementale utilisait un perceptron multicouche. L'implémentation a été réalisée manuellement avec la bibliothèque NumPy, sans recourir à des cadres de travail tels que TensorFlow ou PyTorch, afin de garantir une maîtrise totale de la logique algorithmique et du processus d'apprentissage.

Dans le cadre de l'étude théorique du moteur de classification, deux méthodes d'apprentissage distinctes ont été modélisées pour structurer la reconnaissance des profils environnementaux via un réseau de neurones :

III.4.1.1. Méthode 1 : Approche par Pôles Extrêmes et Interpolation de Confiance

Cette première méthode repose sur la définition de deux vecteurs prototypes absolus représentant les limites du système : le pôle idéal (Classe A, associé à une cible $Y = 1$ soit 100% de confiance) et le pôle critique (Classe E, associé à une cible $Y = 0$ soit 0% de confiance). L'apprentissage consiste à entraîner le réseau à interpoler le comportement des produits intermédiaires à partir de ces deux extrêmes. En injectant des valeurs de test de plus en plus proches du pôle A, le réseau apprend à cartographier les transitions pour délimiter les frontières de chaque Éco-Grade (B, C, D) en fonction du pourcentage de sortie obtenu.

III.4.1.2. Méthode 2 : Approche par Intervalles Statistiques et Écart-Type

La seconde méthode structure l'apprentissage en introduisant une variabilité statistique par classe. Pour chaque Éco-Grade (A et E), un intervalle de scores est défini. Le système calcule d'abord la moyenne arithmétique (μ) de chaque intervalle, puis détermine l'écart-type (σ) modélisé selon les bornes (soit $Max - \sigma$, soit $\sigma - Min$). L'apprentissage du ANN est

configurée pour associer les caractéristiques de ces intervalles aux cibles binaires 1 (100% de confiance pour le comportement de type A) et 0 (0% de confiance pour le comportement de type E). À partir de ces références probabilistes, le réseau est exposé à des valeurs glissantes à l'intérieur de chaque intervalle afin de formaliser mathématiquement les zones d'influence et de transition de chaque catégorie écologique.

III.4.2. Justification du choix d'une approche inspirée du K-Means

Bien que le système ne reproduise pas la phase itérative complète, l'approche retenue s'inspire directement du principe de classification par distance aux centroïdes propre à K-Means, pour les raisons suivantes :

Premièrement, K-Means ne requiert aucune phase d'entraînement itératif sur un jeu de données labellisé. Dans le contexte d'ÉcoTrace-IA, les centroïdes sont déterminés analytiquement à partir des intervalles de grades prédéfinis, ce qui élimine toute dépendance à des données réelles certifiées — ressource inexistante pour le secteur agroalimentaire algérien à ce jour.

Deuxièmement, la structure ordinale et bornée de l'échelle de notation, comprise entre 1 et 5, confère aux cinq classes environnementales une disposition linéaire et régulière dans l'espace des données. Cette propriété rend le partitionnement par centroïdes particulièrement adapté, les frontières de décision étant définies de manière déterministe sans recours à une optimisation stochastique.

Troisièmement, l'algorithme présente une interprétabilité totale. Chaque décision de classification est directement explicable par le calcul d'une distance absolue entre le score moyen pondéré du produit et les centroïdes des cinq classes. Cette transparence algorithmique facilite la validation du modèle et renforce la confiance de l'utilisateur dans les résultats produits.

Il convient de préciser que l'implémentation retenue ne reproduit pas la phase itérative classique de K-Means. Seule la phase de décision est utilisée : les centroïdes étant fixés analytiquement, le système procède directement à la classification par distance minimale, conformément au principe fondamental de l'algorithme.

III.5. Conception algorithmique et développement du modèle K-Means :

L'intégration de l'algorithme K-Means au sein du système se déploie à travers deux applications logicielles distinctes et complémentaires, structurant de manière progressive le passage d'un cadre théorique général à un outil décisionnel industriel spécifique.

Cette approche méthodologique bicéphale a été adoptée afin de dissocier, dans un premier temps, l'étude comportementale et la validation mathématique de l'algorithme de partitionnement dans un espace géométrique standard, de son application concrète aux contraintes environnementales du projet.

Ainsi, le système passe d'une phase d'apprentissage générique, dédiée à la manipulation des clusters et à la persistance des structures de données, à une phase d'inférence vectorielle hautement optimisée, calibrée pour répondre aux exigences opérationnelles de l'application ÉcoTrace-IA sans surcharger les ressources calculatoires du système.

III.5.1. Application N°1 — LogiClass : phase d'Apprentissage

Dans le cadre de ce projet, deux applications logicielles complémentaires ont été développées et déployées. La première, LogiClass, constitue l'unité d'apprentissage du système : il s'agit d'une application autonome et interactive, développée sous Python avec l'interface graphique Tkinter, dédiée exclusivement à la phase d'entraînement de l'algorithme K-Means. Elle ne se limite pas à un simple script de calcul ; elle intègre une interface graphique complète permettant à l'utilisateur de piloter l'intégralité du processus d'apprentissage, de la saisie des données jusqu'à la persistance des centroïdes stabilisés via la bibliothèque pickle.

LogiClass opère sur un jeu de données bidimensionnel dont les coordonnées (x, y) représentent les observations à partitionner. Son rôle est fondamental dans l'architecture globale du système : elle génère le fichier de configuration R.pkl contenant les centroïdes finaux, qui sera ensuite chargé et exploité par la seconde application RT-KMC lors de la phase de classification en temps réel.

III.5.1.1. Description Fonctionnelle des Commandes de l'Interface Graphique

Pour piloter l'exécution de l'algorithme K-Means depuis l'interface de LogiClass, trois boutons d'action principaux ont été implémentés (illustrés sur la **Figure III.1**) :

Bouton odat (Chargement des Données) : Ce bouton initialise l'application en chargeant les fichiers de données simulées. Lors de son activation, le système lit les vecteurs de coordonnées

stockés (fichiers x.npy et y.npy) pour générer et afficher le nuage initial de 100 points (représentés par des étoiles noires). C'est l'étape préliminaire obligatoire avant tout calcul.

Bouton irf (Initialisation des Référents) : Ce bouton est responsable de l'injection des conditions initiales dans l'espace de travail. En cliquant dessus, l'application capture la matrice saisie dans le champ initref (contenant les 4 coordonnées initiales R_0) et positionne ces 4 référents sous forme de croix (×) colorées au centre du nuage de points, préparant ainsi l'algorithme à l'affectation.

Bouton trkm (Exécution du K-Means Itératif) : C'est le moteur de calcul principal de LogiClass. Lors de son activation, le système récupère la valeur entière saisie dans le champ itirat (le nombre d'itérations : 1, 2 ou 3) et lance la boucle d'entraînement itérative. Il exécute successivement le calcul des distances euclidiennes, l'affectation des points, et la mise à jour des barycentres. À la convergence, il sérialise les centroïdes finaux dans le fichier **R.pkl** via pickle, assurant ainsi leur persistance pour une utilisation ultérieure par RT-KMC.



Figure III.1 : Capture d'écran de l'interface Tkinter vide ou avec les boutons.

III.5.1.2. Déroulement du Processus en Trois Étapes Successives

Étape 1 — Visualisation des données brutes

Les coordonnées x et y sont chargées depuis des fichiers .npy et affichées sous forme de nuage de points. La Figure III.2 présente la distribution des données, révélant visuellement quatre regroupements naturels distincts dans l'espace bidimensionnel.

La **figure III.2** illustre les quatre regroupements.

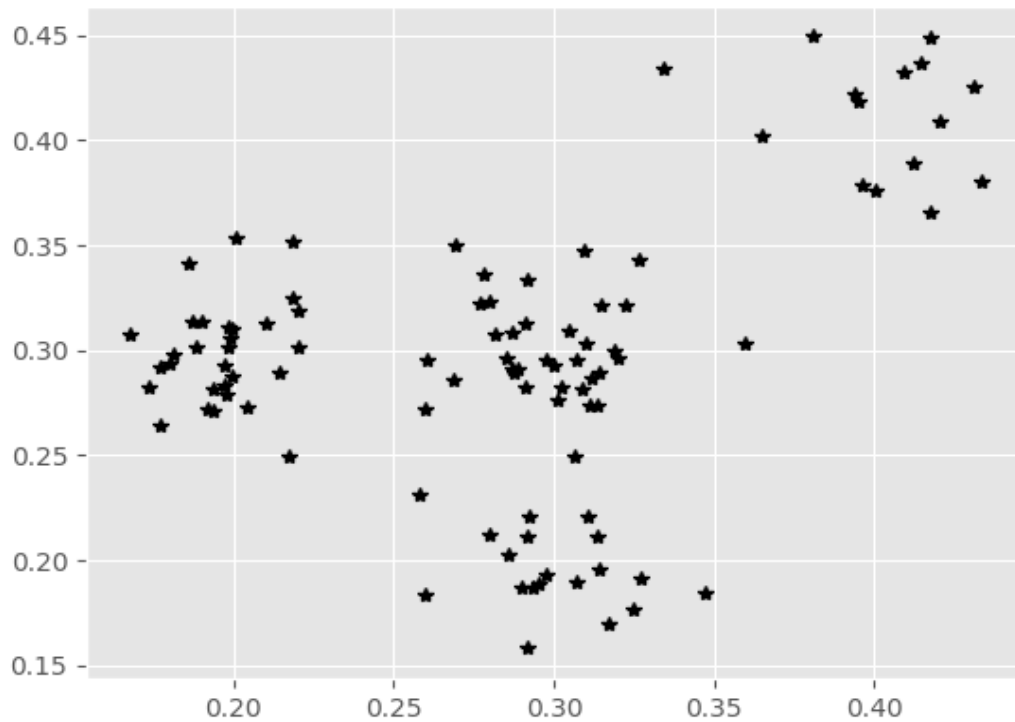


Figure III.2: Nuage de points bruts avec uniquement les étoiles noires.

Étape 2 — Visualisation des référents initiaux

Quatre référents initiaux (R_0) sont définis manuellement par l'utilisateur sous forme de vecteurs de coordonnées à deux dimensions. La matrice de positionnement initial s'établit ainsi :

$$R_0 = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{pmatrix} = \begin{pmatrix} 0.40 & 0.27 \\ 0.30 & 0.25 \\ 0.30 & 0.35 \\ 0.35 & 0.20 \end{pmatrix}$$

Ce qui correspond rigoureusement au cas de test suivant pour chaque pôle initial :

$$\text{Pôle Rouge } (C_1) : \begin{cases} x_1 = 0.40 \\ y_1 = 0.27 \end{cases}$$

$$\text{Pôle Vert } (C_2) : \begin{cases} x_2 = 0.30 \\ y_2 = 0.25 \end{cases}$$

$$\text{Pôle Cyan } (C_3) : \begin{cases} x_3 = 0.30 \\ y_3 = 0.35 \end{cases}$$

$$\text{Pôle Bleu } (C_4) : \begin{cases} x_4 = 0.35 \\ y_4 = 0.20 \end{cases}$$

Ces valeurs sont choisies au centre de l'espace des données afin d'illustrer la capacité de l'algorithme à converger de manière itérative depuis une position centrale vers les barycentres réels des clusters. La **Figure III.3** montre la projection graphique de ces référents sur le nuage de points.

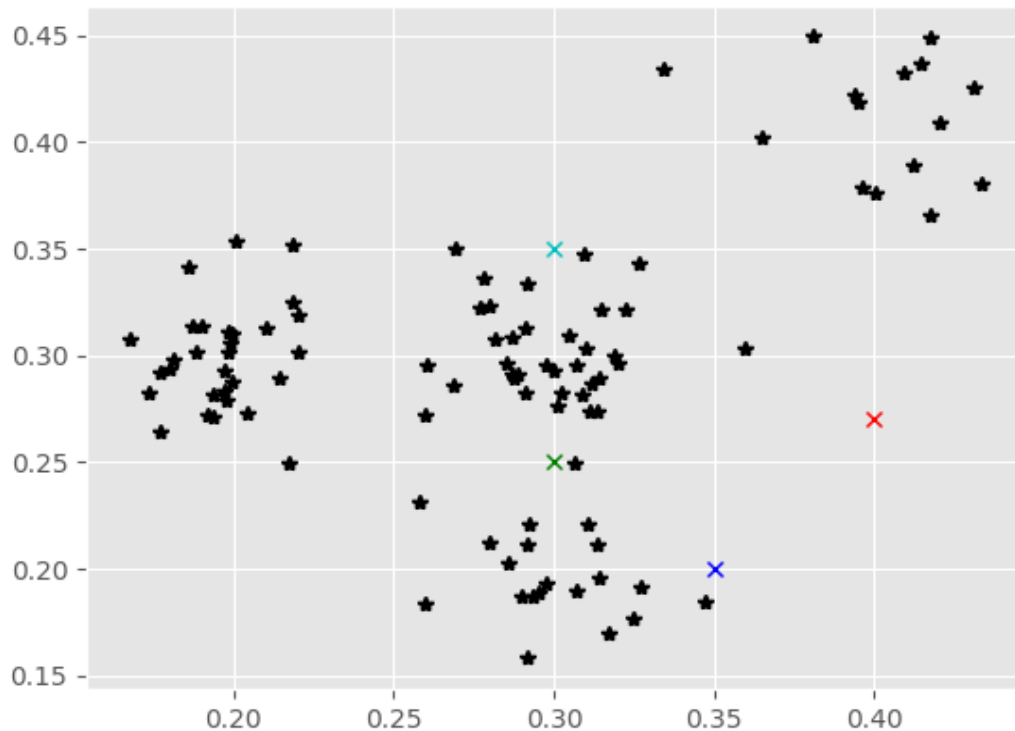


Figure III.3 : Graphique du nuage de points avec uniquement les croix 'x' au milieu.

Étape 3 — Entraînement itératif : L'algorithme procède par itérations successives. À chaque itération, le processus se divise en deux phases distinctes :

Affectation des observations : Chaque observation x_n est affectée au cluster C_i du référent R_i le plus proche, selon la distance euclidienne (Norme L_2) [38].

$$d(x_n, R_i) = \|x_n - R_i\|_2$$

Mise à jour des référents : Les positions des référents pour l'itération suivante ($k + 1$) se recalculent en déterminant la moyenne arithmétique (barycentre) des coordonnées de toutes les observations qui leur ont été assignées [38] :

$$R_i^{(k+1)} = \frac{1}{|C_i|} \sum_{x_n \in C_i} x_n$$

Où $|C_i|$ représente le cardinal (le nombre total d'observations) du cluster C_i .

L'exécution de ce processus itératif est indispensable pour permettre aux coordonnées initiales du vecteur de converger progressivement vers leurs positions barycentriques optimales.

Comme le montrent les figures suivantes, la répétition itérative des calculs est la condition nécessaire pour que chaque point pivot (x, y) atteigne sa position finale de stabilisation, assurant ainsi la convergence absolue de l'algorithme.

III.5.1.3. Analyses Expérimentales et Dynamique de Convergence

Cas de 1 itération (itirat = 1) :

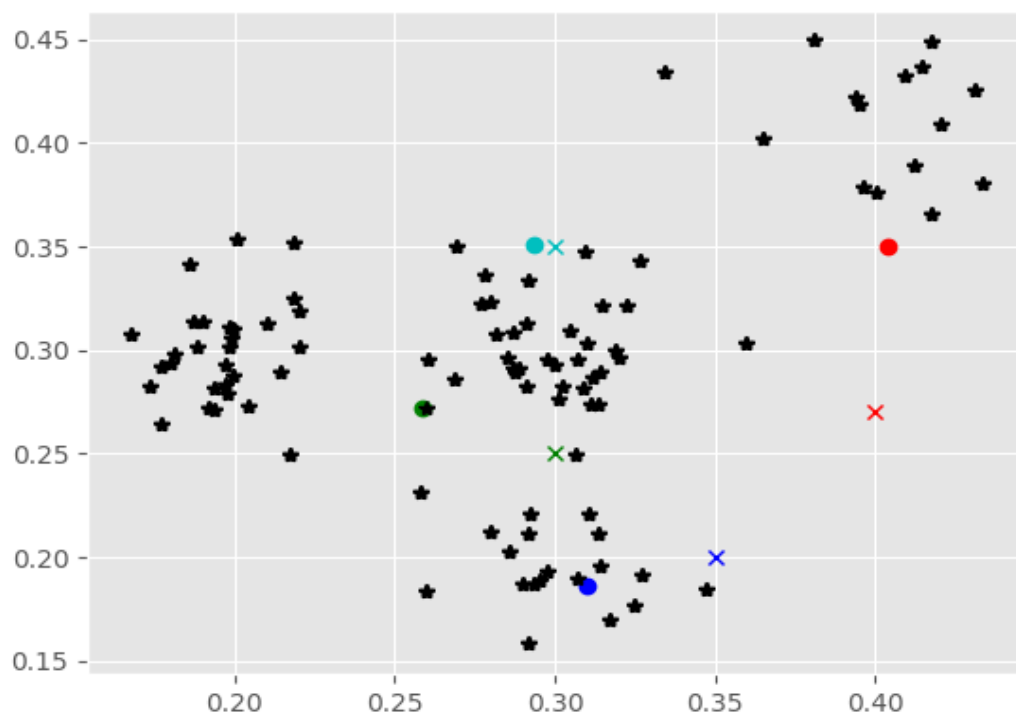


Figure III.4 : Graphique montrant itératif = 1, où les cercles touchent presque les croix.

Comme on le remarque dans la **Figure III.4**, à la première itération, les nouveaux centroïdes restent proches des positions initiales. Un seul passage de calcul ne suffit pas à briser la contrainte du regroupement central initial. Les frontières de décision restent biaisées et insuffisantes pour classer les groupes périphériques.

Cas de 2 itérations (itirat = 2)

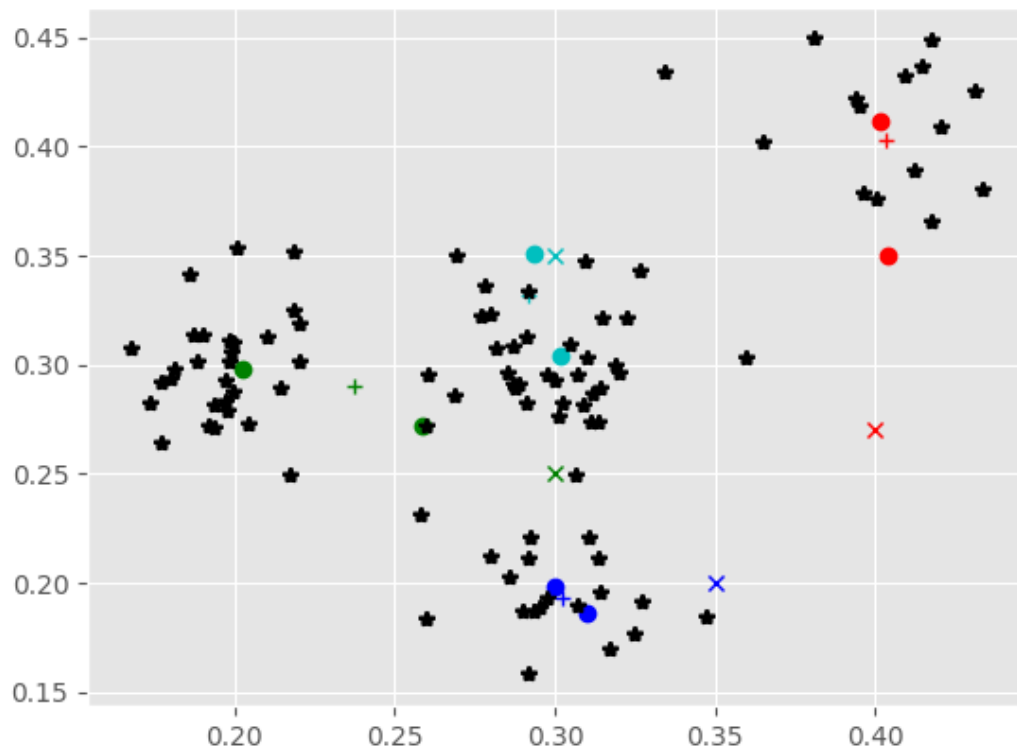


Figure III.5 : Graphique montrant itératif = 2, avec les trajectoires en '+'.

La deuxième itération révèle une dynamique d'expansion claire des pôles à travers l'espace. Les trajectoires transitoires (+) prouvent la migration des centroïdes vers les zones de densité réelle. Le recalcul des barycentres commence à réduire efficacement la variance intra-classe (Figure III.5)

Cas de 3 itérations (itirat = 3)

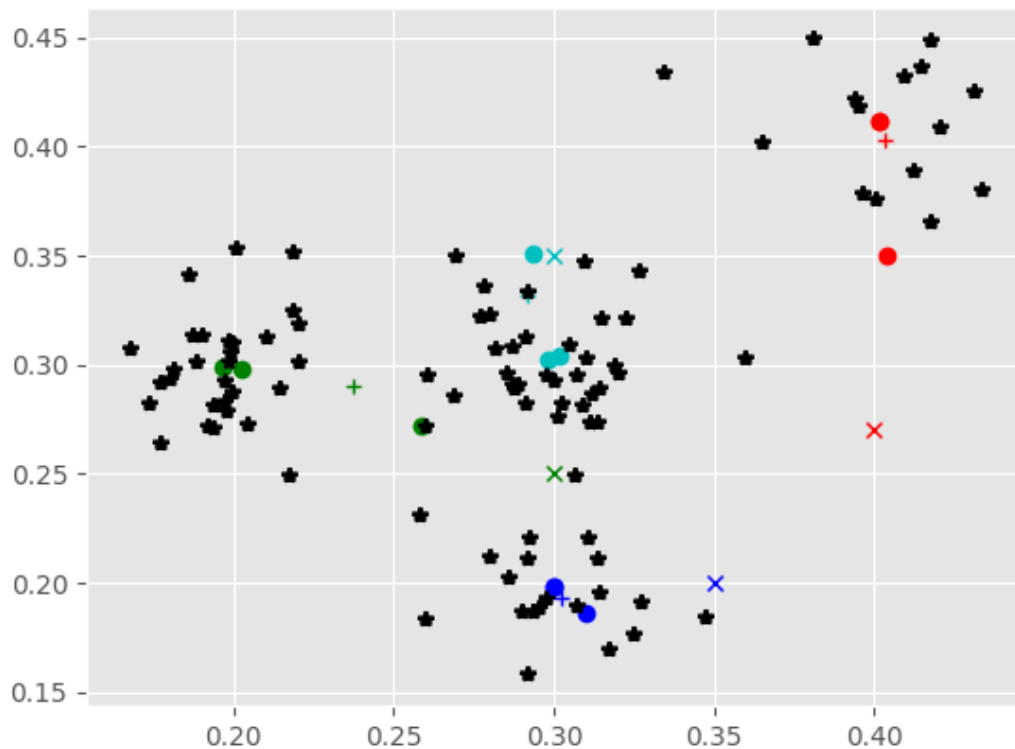


Figure III.6 : Graphique final montrant itératif = 3, stabilité complète.

À la troisième itération, LogiClass atteint la convergence absolue et les positions des centroïdes deviennent stationnaires. Chaque pôle est parfaitement ancré au centre de gravité réel de son cluster. Les frontières de décision sont optimales et le fichier R.pkl est généré avec les centroïdes définitifs (**Figure III.6**).

III.5.2. Application N°2 — RT-KMC : Phase de test et Décision

III.5.2.1. Description de l'Interface de Test

En plus de l'interface d'entraînement, un second module graphique développé sous Tkinter (**Figure III.7**) a été mis en place pour valider et exploiter les résultats de l'apprentissage. Ce module permet à l'utilisateur de tester le modèle en introduisant les coordonnées (x, y) d'une nouvelle observation brute afin d'identifier sa classe d'appartenance de manière instantanée.

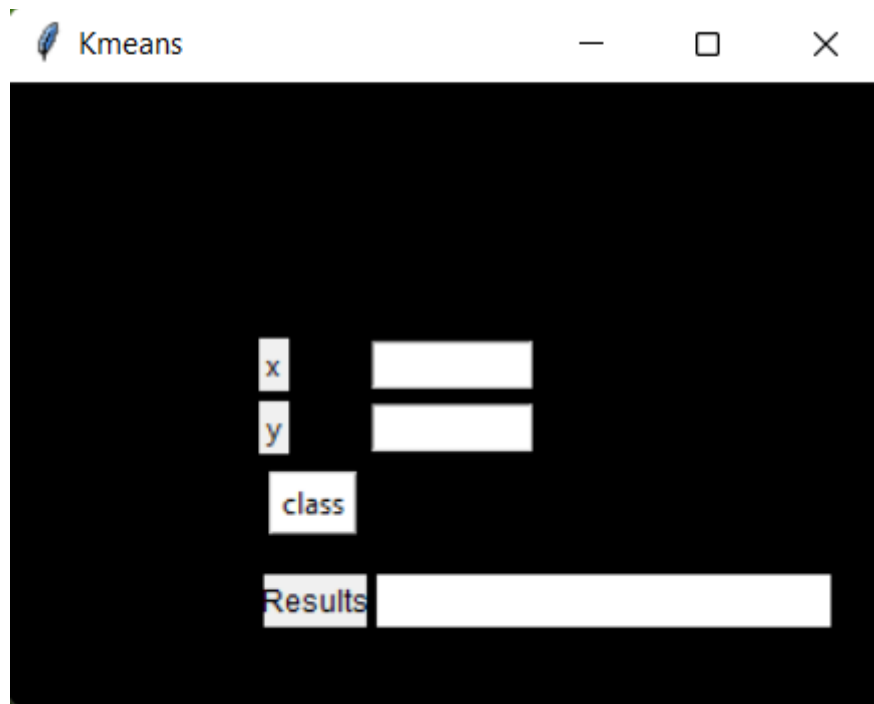


Figure III.7 : Capture d'écran de l'interface Tkinter du module.

III.5.2.2. Algorithme de Décision et Traitement Interne

Le fonctionnement de ce module de test repose sur la restauration du fichier de configuration persistant R (généré via la bibliothèque pickle lors de la phase d'entraînement). Ce fichier contient :

$R[0]$: La matrice des coordonnées des 4 centroïdes finaux stabilisés.

$R[1], R[2], R[3], R[4]$: Les valeurs limites (bornes maximales et minimales) de l'espace de données d'apprentissage.

Lors du clic sur le bouton class, le script exécute deux étapes logiques successives :

Vérification des Bornes (Contrainte d'Espace) : Le système valide d'abord que le point soumis appartient bien au domaine géométrique connu en évaluant la condition logique :

$$\text{Condition} = (x \leq R[1] \wedge x \geq R[2]) \wedge (y \leq R[3] \wedge y \geq R[4])$$

Si la condition n'est pas vérifiée, le système rejette l'entrée et affiche : "the point [x, y] has no-class".

Calcul de Proximité et Assignation :

Si le point est valide, le programme calcule sa distance euclidienne (Norme L_2) par rapport aux 4 centroïdes stockés [38] :

$$d_i = \sqrt{(R[0][i]_x - x)^2 + (R[0][i]_y - y)^2} \quad \text{pour } i \in \{0, 1, 2, 3\}$$

Le label graphique L1 est ensuite mis à jour pour afficher la classe correspondant à la distance minimale ($\min(d_i)$) parmi les quatre classes possibles (class1, class2, class3 ou class4).

Enfin, les champs de saisie sont réinitialisés via `x1.set("")` et `x2.set("")`.

III.5.2.3. Évaluation Expérimentale et Cas de Tests du Module de Décision

Afin de valider le comportement dynamique, l'exactitude mathématique ainsi que la robustesse du module de décision développé, trois scénarios de test critiques ont été injectés manuellement via l'interface graphique dédiée au test. Les résultats obtenus confirment le bon fonctionnement des filtres logiques et du calcul de proximité euclidienne.

Premier cas de test : Point hors-limites et rejet du système (No-Class)

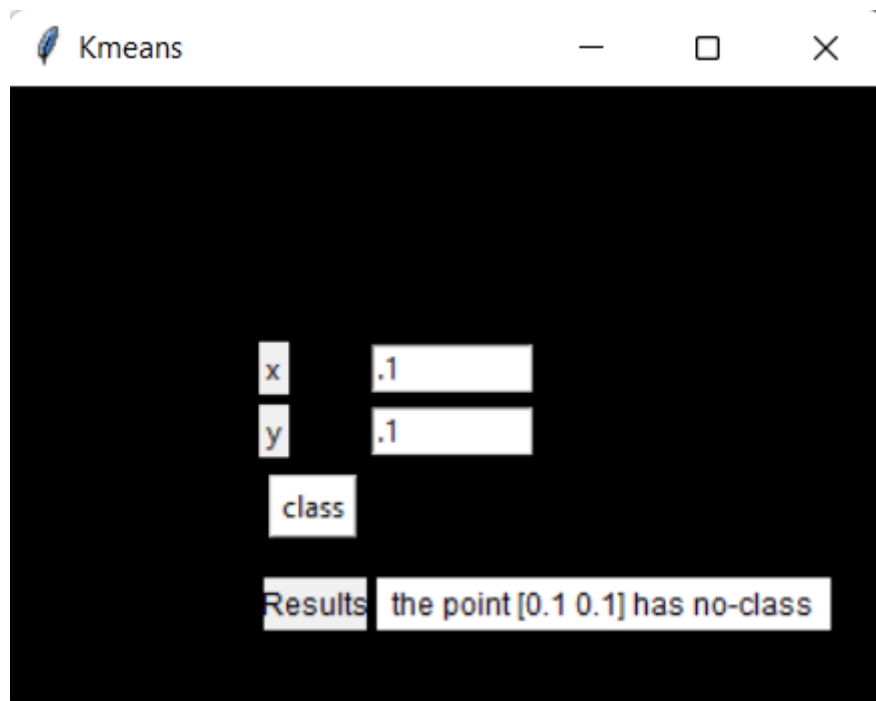


Figure III.8 : Comportement du module de test lors de l'injection d'un point hors-limites ($x = 0.1, y = 0.1$).

Analyse du comportement : Comme le met en évidence la **Figure III.8**, lorsque l'utilisateur introduit ces coordonnées, le point se situe géographiquement en dehors de la boîte englobante (Bounding Box) définie par les frontières $[R[1], R[2], R[3], R[4]]$ calculées lors de la phase d'apprentissage. Le filtre logique de sécurité intercepte la donnée avant l'activation du calcul de distance. Le système rejette instantanément l'entrée et affiche le message d'avertissement : "the point $[0.1, 0.1]$ has no-class". Cela prouve la fiabilité du système face à des données aberrantes ou inconnues.

Deuxième cas de test : Assignment optimale à la Class 4

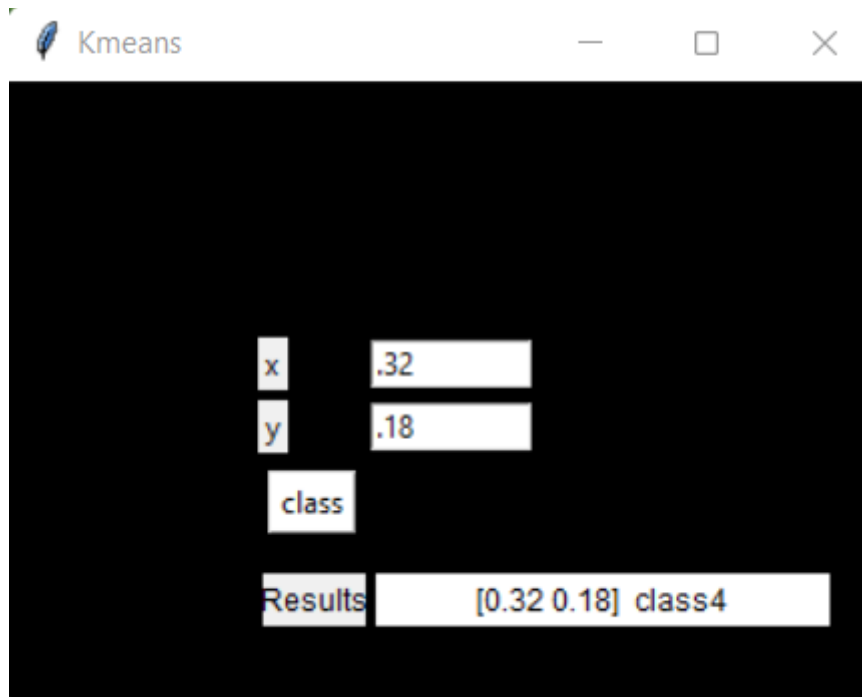


Figure III.9 : Identification et classification en temps réel d'un point de test ($x = 0.32$, $y = 0.18$) associé à la Classe 4.

Analyse du comportement : Dans ce deuxième scénario illustré par la **Figure III.9**, les coordonnées pénètrent le domaine de validité géométrique. L'algorithme calcule alors les distances euclidiennes par rapport aux quatre centroïdes stables récupérés depuis le fichier persistant R. C'est la distance par rapport au quatrième centroïde $R[0][3]$ (pôle bleu) qui s'avère minimale. Le système valide l'affectation et actualise l'affichage en temps réel : $[0.32 \ 0.18]$ class 4.

Troisième cas de test : Assignment optimale à la Class 1

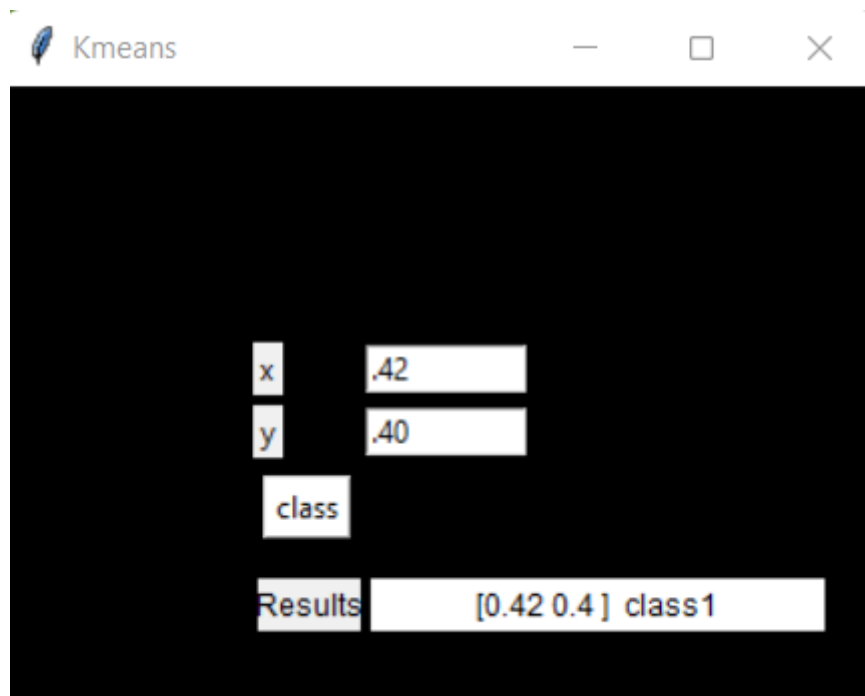


Figure III.10 : Identification et classification en temps réel d'un point de test ($x = 0.42$, $y = 0.40$) associé à la Classe 1.

Ce dernier cas de test, présenté sur la **Figure III.10**, cible la région supérieure droite de l'espace échantillonné. Après vérification positive des contraintes de bornes, le vecteur de test se révèle être mathématiquement plus proche du premier centroïde stabilisé $R[0][0]$ (pôle rouge). La fonction d'assignation isole cette distance minimale avec succès, et la fenêtre Tkinter confirme immédiatement le diagnostic en affichant : `[0.42 0.4] class1`.

III.5.3. Architecture et Approche de l'Application N°3 — ÉcoTrace-IA

Pour évaluer la performance environnementale, nous avons opté pour une approche déterministe basée sur une échelle de notation experte de 1 à 5. Contrairement aux méthodes d'apprentissage automatique qui nécessitent une phase d'entraînement, notre modèle calcule directement le score à partir de la pondération des 16 critères définis.

Calcul des Centroïdes (Centrage des classes) :

Pour assurer une distribution équitable des produits dans les catégories (A à E), nous avons défini les intervalles de classification de manière mathématique :

Borne inférieure (Grade A) : $16 \times 1 = 16$

Borne supérieure (Grade E) : $16 \times 5 = 80$

Intervalles (Pas) : En divisant cette amplitude par les 5 catégories, nous obtenons un pas de 12.3.

En normalisant ces valeurs par le nombre de critères (16), nous avons extrait les **centroïdes optimaux** pour chaque grade. Cette méthode garantit une classification transparente, logique et reproductible, éliminant ainsi le besoin de modèles complexes d'apprentissage.

Il est crucial de noter que si l'échelle de saisie est discrète (1 à 5), les centres de classification (1.4, 2.2, ...) sont des barycentres statistiques. Ce choix d'utiliser des centres décalés par rapport aux entiers (1, 2, 3...) permet de créer des zones de tolérance autour de chaque grade, évitant ainsi une instabilité de la classification et assurant que chaque Éco-Grade représente une plage de performance réelle plutôt qu'un point isolé.

III.5.3.1. Interface Utilisateur (GUI) et Saisie des Données

L'application ÉcoTrace-IA a été conçue pour offrir une expérience utilisateur fluide, ergonomique et visuellement intuitive. L'interface graphique (GUI) a été développée en utilisant la bibliothèque **Tkinter**, choisie pour sa légèreté et son intégration native avec Python, garantissant ainsi une exécution autonome (offline) sur les postes de travail des PME.

Écran de Bienvenue : Comme illustré dans la **Figure III.11**, l'écran d'accueil joue un rôle de guide. Il présente l'identité visuelle du projet à travers un logo épuré symbolisant la durabilité. Ce point d'entrée permet à l'utilisateur de s'approprier l'outil via un message d'accueil clair et un bouton d'action unique "Commencer l'évaluation", simplifiant ainsi la navigation dès les premières secondes.

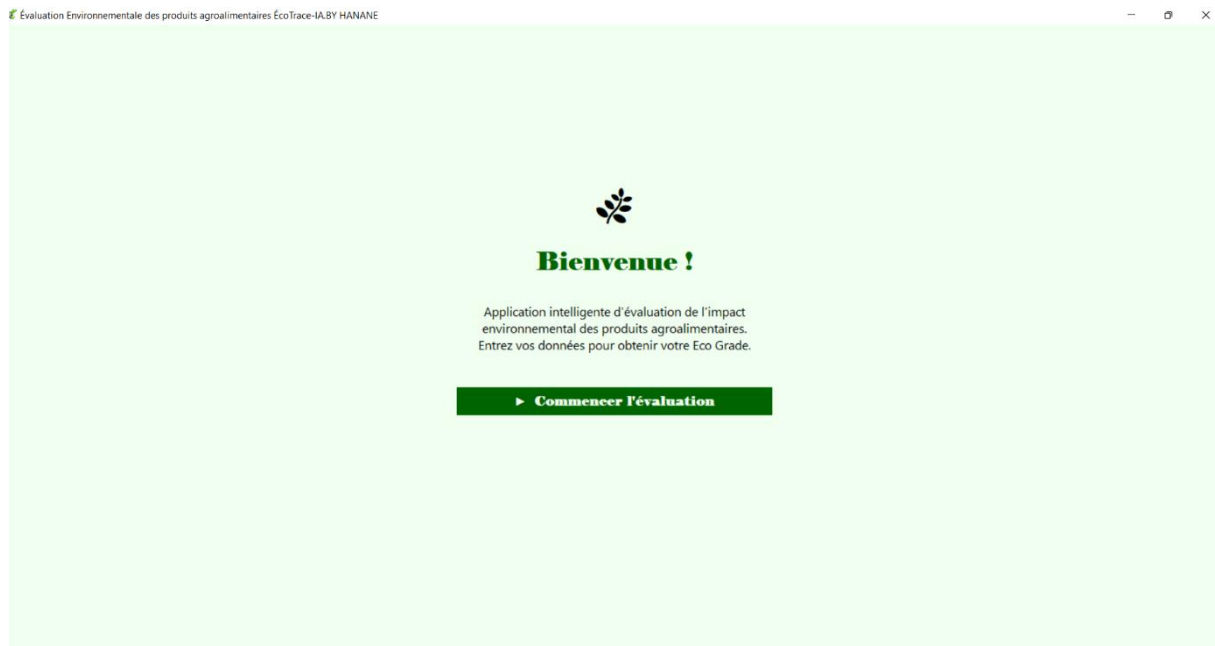


Figure III.11 : Interface d'accueil du système ÉcoTrace-IA

Tableau de bord d'Évaluation : La Figure III.12 présente la plateforme principale, structurée de manière à séparer logiquement l'entrée de données et la lecture des résultats :

Zone de Saisie (Colonne de gauche) : Cette section est dédiée au renseignement des 16 critères environnementaux. Pour assurer la rigueur des données, nous avons opté pour :

Une iconographie dédiée : Chaque critère est illustré par une icône spécifique, facilitant l'identification rapide par l'opérateur industriel.

Des champs normalisés : L'usage de listes déroulantes (*Combobox*) restreint les choix aux modalités prédéfinies, évitant ainsi les erreurs de saisie et garantissant que chaque entrée est directement convertible en score numérique.

Zone de Résultat et Recommandation (Colonne de droite) : Cette zone est le centre décisionnel de l'interface :

Visualisation du grade : Un espace central est réservé à l'affichage dynamique de l'Éco-Grade obtenu (de A à E) et du score calculé.

Bloc de recommandation : Une fenêtre interactive génère des conseils d'optimisation basés sur les faiblesses détectées dans les critères saisis.

Statut système : Un indicateur de statut en bas de page assure une rétroaction en temps réel, informant l'utilisateur si l'application est "En attente" ou si le calcul est "Terminé".

Ce design modulaire respecte les principes d'ergonomie cognitive : il réduit la charge mentale de l'utilisateur tout en garantissant que le moteur de classification dispose de données structurées pour effectuer son inférence vectorielle.

Évaluation Environnementale des produits agroalimentaires ÉcoTrace-IA BY HANANE

ÉcoTrace-IA

- Type d'emballage: Plastique vierge
- Type d'énergie: Charbon
- Consommation d'énergie: < 0.5 kWh/kg
- Type de transport amont: Avion
- Distance de transport amont: Local
- Type de transport aval: Avion
- Distance de transport aval: Local
- Poids de l'emballage (%): < 5%
- Type de déchets: Organique
- Déchets en décharge (%): < 5%
- Source d'eau: Eau de pluie/puits
- Consommation d'eau (l/kg): < 100
- Agriculture: Bio
- Classe alimentaire: Neutre (céréales, légumineuses)
- Type de sol: Tourbeux
- Fin de vie de l'emballage: Recycle

Évaluer le produit

Résultat

Grade :

Score :

Recommandations d'optimisation :

Aucune recommandation générée. Veuillez évaluer un produit.

Statut :
En attente de données...

Figure III.12 : Interface principale d'évaluation : saisie des critères et zone de résultats.

III.5.3.2. Moteur de Calcul : Logique Algorithmique et Pondération

La puissance décisionnelle d'ÉcoTrace-IA réside dans son moteur de calcul, conçu pour transformer des données qualitatives en une évaluation environnementale objective. Ce processus s'articule autour de trois étapes clés :

L'étape 1 : Vectorisation et Normalisation

Les données saisies sont transformées en un **vecteur d'entrée** $\vec{X} \in \mathbb{R}^{16}$. Parallèlement, nous définissons un **vecteur de poids** $\vec{W}$ (les coefficients d'importance), qui est normalisé tel que $\sum w_i = 1$. Cette normalisation garantit que l'influence totale des critères reste cohérente quelle que soit la pondération initiale :

$$\vec{W} = \frac{\vec{W}_{raw}}{\sum \vec{W}_{raw}}$$

L'étape 2 : Calcul du Score Moyen Pondéré

Pour modéliser l'impact réel du produit, nous appliquons un produit de Hadamard (multiplication élément par élément) entre le vecteur d'entrée $\vec{X}$ et le vecteur de poids $\vec{W}$, générant ainsi un **vecteur pondéré** $\vec{Y}$:

$$\vec{Y} = \vec{X} \odot \vec{W}$$

Cette étape permet d'amplifier ou d'atténuer l'influence de chaque critère selon son importance écologique sectorielle (par exemple, le poids élevé accordé à l'énergie par rapport au type de sol).

L'étape 3 : Classification par Centroïdes pondérés :

Les vecteurs de référence (centroïdes) $\vec{C}_j$ (pour chaque grade de A à E) sont également projetés dans le même espace vectoriel que $\vec{Y}$ en leur appliquant les mêmes pondérations. La classe d'appartenance est déterminée par la distance euclidienne minimale [36] :

$$Grade = \arg \min_{j \in \{A..E\}} \sqrt{\sum_{k=1}^{16} (y_k - c_{j,k})^2}$$

Enfin, l'indice de confiance est calculé par une mesure de proximité relative : plus la distance $\|\vec{Y} - \vec{C}_j\|$ est faible, plus le score de similarité tend vers 100%.

Traitement des critères non applicables :

La pertinence de certains indicateurs varie selon la nature du produit. À titre d'illustration, les propriétés du sol ou les méthodes culturales ne s'appliquent pas aux produits halieutiques ou strictement industriels. Pour remédier à ces disparités, le système permet l'exclusion des critères non pertinents en leur affectant une valeur nulle. Le score global est alors calculé par une moyenne pondérée uniquement sur les critères actifs, selon la formule suivante :

$$\bar{x} = \frac{\sum_{i=1}^n s_i}{n}$$

Dans cette équation, s_i représente le score de l'indicateur i et n correspond au nombre total de critères effectivement renseignés. Cette méthode de normalisation adaptative permet d'ajuster l'évaluation à la spécificité de chaque produit, évitant ainsi un biais de notation induit par des paramètres non applicables.

III.5.3.3. Résultats et Validation par cas pratiques

Afin de valider la fiabilité de notre moteur d'inférence vectorielle, nous avons testé l'application ÉcoTrace-IA sur trois produits agroalimentaires présentant des profils environnementaux radicalement différents. Cette simulation confirme la cohérence entre la pondération des critères et les grades attribués.

Résultat de l'évaluation du Produit-Type 3 : Produit industriel à haute chaîne logistique (Grade C).

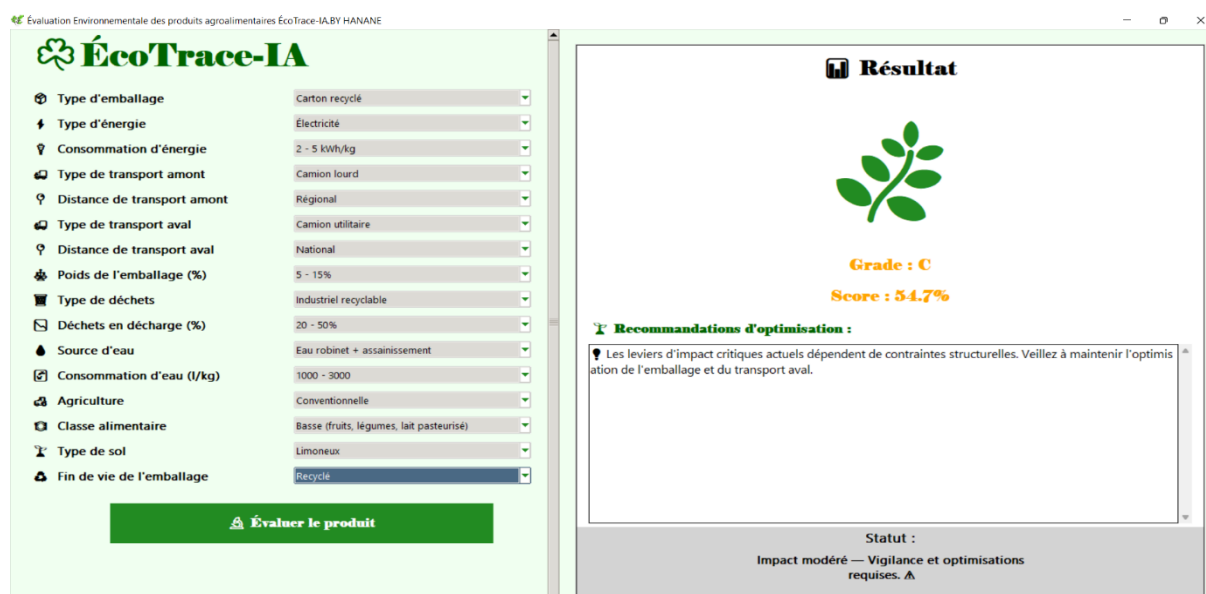


Figure III.13 : Évaluation du Produit-Type 3 (Grade C)



Figure III.14 : Résultat de l'évaluation du Produit-Type 3 : Produit industriel à chaîne logistique modérée (Éco-Grade C).

L'évaluation révèle un impact modéré, caractéristique d'une chaîne logistique industrielle standard. Le système souligne ici une marge d'optimisation sur le transport aval, illustrant la capacité du modèle à identifier les leviers de progression environnementale.

Résultat de l'évaluation du Produit-Type 2 : Produit de consommation courante (Grade B).

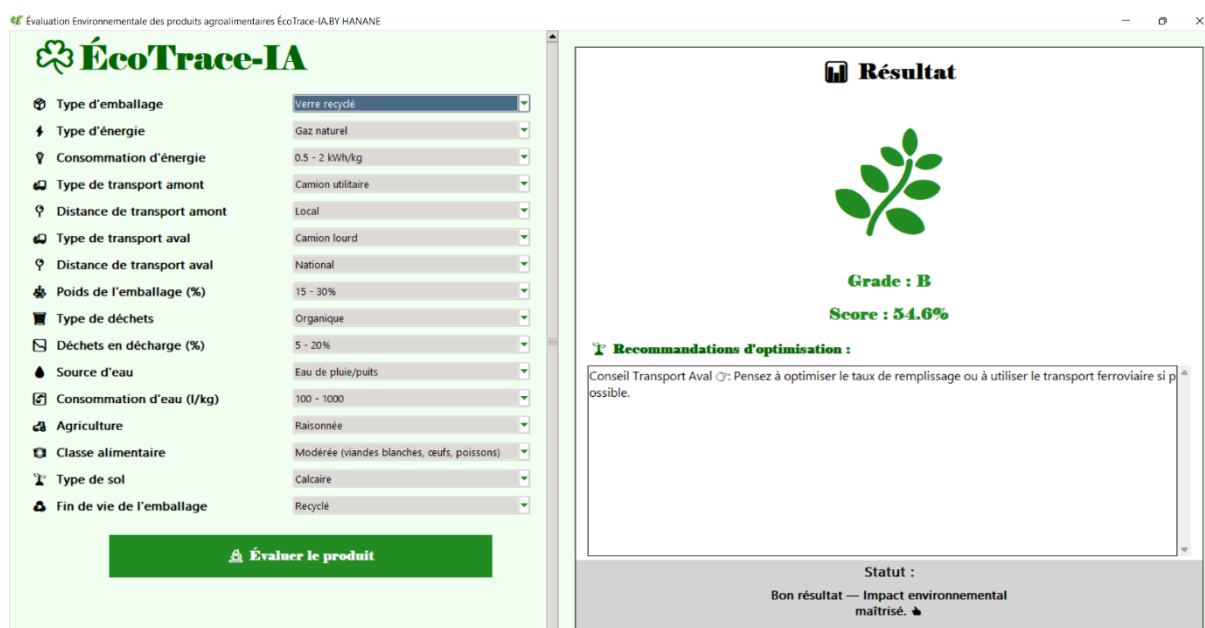


Figure III.15 : Évaluation du Produit-Type 2 (Grade B)

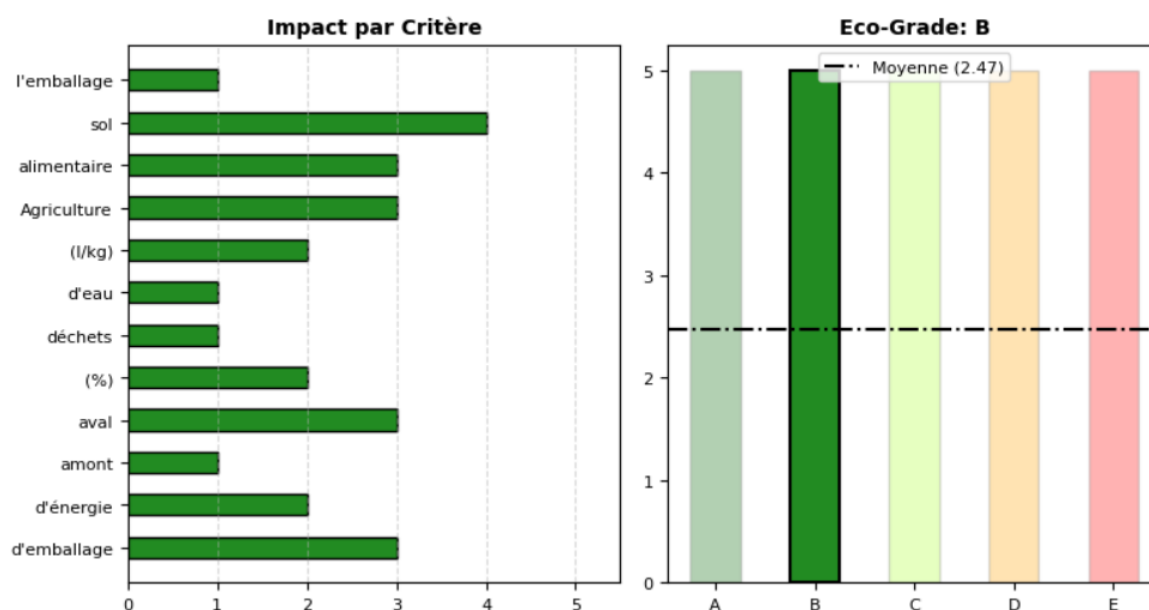


Figure III.16 : Résultat de l'évaluation du Produit-Type 2 : Produit de consommation courante (Éco-Grade B).

Ce résultat reflète une maîtrise optimisée des impacts. L'algorithme détecte une gestion efficiente des ressources énergétiques et agricoles, positionnant naturellement le produit dans la catégorie B, conforme aux attentes de durabilité sectorielle.

Résultat de l'évaluation du Produit-Type 1 : Produit à faible transformation (Grade A).



Figure III.17 : Évaluation du Produit-Type 1 (Grade A)

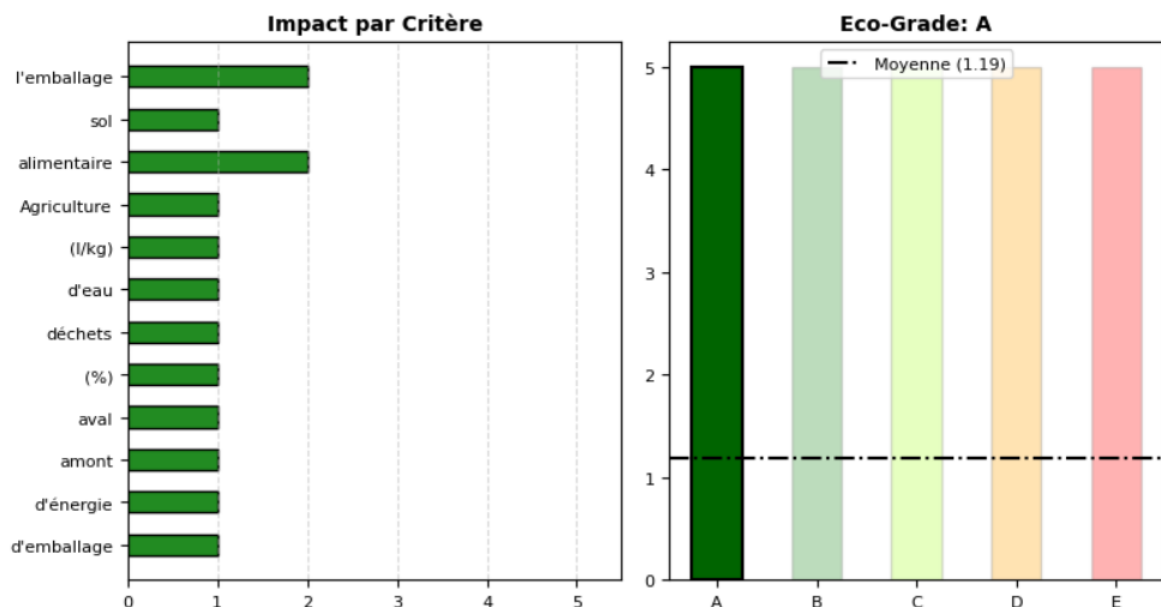


Figure III.18 : Résultat de l'évaluation du Produit-Type 1 : Produit à faible transformation (Éco-Grade A).

Le résultat illustre un modèle d'excellence environnementale. La faible pondération des critères de transformation et la proximité locale des circuits de distribution permettent au système d'attribuer un Éco-Grade A, validant ainsi la pertinence des seuils de performance définis.

Résultat de l'évaluation du Produit-Type 5 : Produit importé à empreinte carbone critique (Grade E).

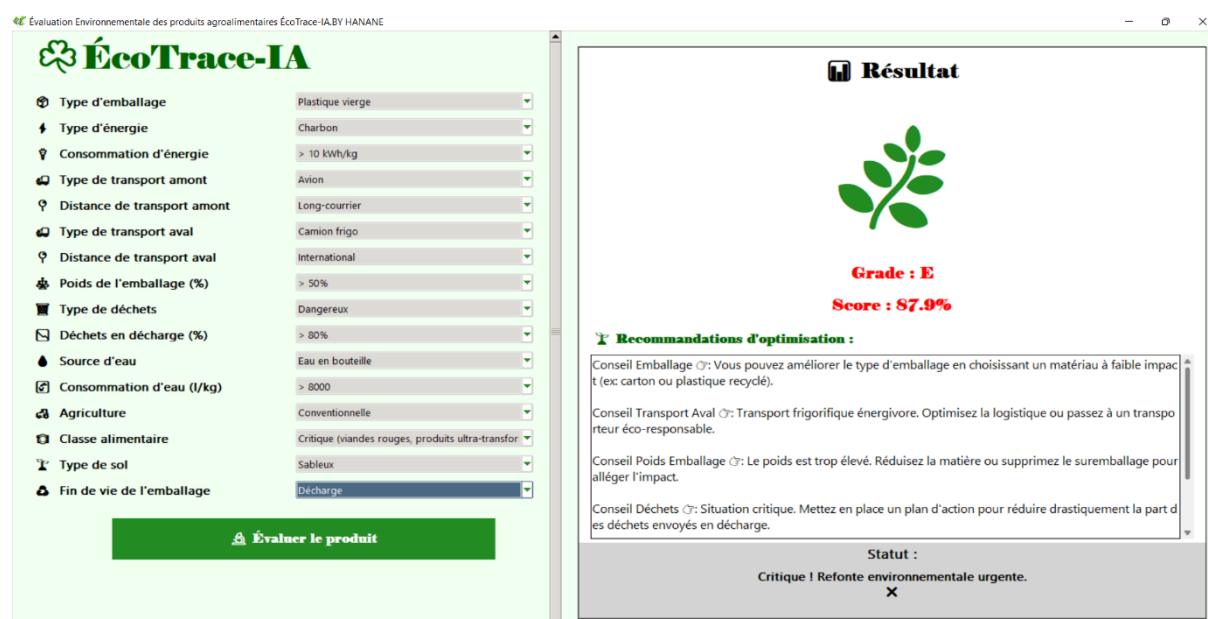


Figure III.19 : Évaluation du Produit-Type 5 (Grade E)



Figure III.20 : Résultat de l'évaluation du Produit-Type 5 : Produit importé à empreinte environnementale critique (Éco-Grade E).

Ce cas souligne la sensibilité du moteur d'inférence face aux critères critiques. La forte empreinte logistique et le mode de production ultra-transformé entraînent une classification immédiate en Grade E, confirmant la robustesse du système dans l'identification des produits à fort risque écologique.

III.6. Déploiement et mise en place de l'application (Format .exe)

La mise en place opérationnelle des trois modules du système nécessite une structuration rigoureuse des fichiers sources. Pour garantir le bon fonctionnement de chaque application, le répertoire de travail doit impérativement regrouper le script principal, l'icône associée, ainsi que le fichier de persistance des centroïdes R.pkl pour les modules d'apprentissage et de classification.

La procédure de compilation pour obtenir un fichier exécutable autonome s'effectue via l'outil PyInstaller, installé via la commande :

```
pip install pyinstaller
```

III.6.1. Unité d'apprentissage — LogiClass

```
pyinstaller --noconsole --onefile --icon="logicclass.ico" LogiClass.py --clean
```

Cette application génère, à l'issue de la phase d'entraînement, le fichier de configuration R.pkl contenant les centroïdes stabilisés. Il est impératif que l'exécutable généré et ce fichier soient conservés ensemble au sein du même répertoire, afin que le module puisse sauvegarder correctement les résultats d'apprentissage lors de son exécution.

III.6.2. Unité d'apprentissage (RT-KMC)

```
pyinstaller --noconsole --onefile --icon="shutterstock_1580728858.ico" RT-KMC.py --clean
```

Cette opération génère un fichier exécutable unique facilitant sa distribution. Afin d'assurer l'exécution correcte du moteur de classification, l'application générée et le fichier de configuration **R.pkl** — produit par LogiClass — doivent impérativement être conservés

ensemble au sein du même répertoire. Cette coexistence physique garantit que le module puisse charger correctement les centroïdes lors de son lancement.

III.6.3. Unité d'évaluation (ÉcoTrace-IA)

```
pyinstaller --noconsole --onefile --icon="herb-emoji-clipart-xl.ico" ÉcoTrace-IA.py --clean
```

Contrairement aux deux modules précédents, l'application ÉcoTrace-IA est conçue pour être totalement autonome. Le processus de compilation intègre l'intégralité des ressources et des paramètres nécessaires au fonctionnement du moteur d'inférence directement au sein du fichier exécutable, sans dépendance à un fichier externe.

III.7. Conclusion

Ce chapitre a été consacré à la présentation technique des trois composantes logicielles de notre système. Nous avons commencé par **LogiClass**, l'unité d'apprentissage qui implémente l'algorithme K-Means pour la génération et la persistance des centroïdes. Ensuite, nous avons présenté **RT-KMC**, qui exploite ces données apprises pour une classification en temps réel dans un espace bidimensionnel.

Enfin, l'application **ÉcoTrace-IA** représente l'aboutissement fonctionnel de notre démarche. En s'inspirant du principe de classification par distance euclidienne de K-Means, avec des centroïdes définis de manière experte et déterministe dans un espace vectoriel pondéré à 16 dimensions, elle démontre la robustesse et la scalabilité de notre modèle initial face à des problématiques métier complexes, telles que l'évaluation environnementale des produits agroalimentaire.

Chapitre IV : Étude de faisabilité et modèle économique

Introduction

Toute solution technologique, aussi performante soit-elle sur le plan algorithmique, ne peut prétendre à un impact réel sans une assise économique solide. Ce chapitre aborde donc la dimension entrepreneuriale d'ÉcoTrace-IA à travers deux outils complémentaires : le Business Model Canvas, qui structure la logique de création de valeur, et le Business Plan, qui en traduit la faisabilité en chiffres concrets sur cinq années d'exploitation.

Business Model Canvas

Business Model Canvas (BMC)

Le Business Model Canvas, développé par Alexander Osterwalder, est un outil de planification stratégique structuré en neuf blocs interconnectés. Il permet de représenter de façon synthétique la manière dont une solution crée, délivre et capture de la valeur. Le tableau suivant présente le modèle économique d'ÉcoTrace-IA selon ces neuf axes [42] :

Partenaires Clés	Activités Clés	Proposition de Valeur	Relations Clients	Segments Clients
Ministère des Startup et des PME : soutien juridique et institutionnel. Fonds d'Amorçage (ALGEX / ANADE / ANSEJ) : soutien et financement des startups.	Développement et mise à jour de l'algorithme d'évaluation environnementale.	Évaluation environnementale immédiate avec notation Éco-Grade A à E. Diagnostic préventif des points faibles avant exportation. Conformité aux exigences européennes (MACF). Outil accessible et à faible coût pour les PME.	Support technique dédié par entreprise.	PME agroalimentaires algériennes petites et moyennes. Entreprises exportatrices exposées aux taxes carbonees. Organismes gouvernementaux et universités. Startups émergentes en recherche de certifications environnementales.
	Gestion de la base de données des normes environnementales.		Ateliers de formation des utilisateurs.	
	Garantie de qualité et sécurité des données / Développement de l'interface graphique.		Contrats de maintenance annuels.	
			Résultats fiables, transparents et traçables.	
	Ressources Clés		Canaux	
	Base de données des 16 critères environnementaux pondérés.		Visites directes B2B vers les entreprises et usines.	
	Compétences techniques en développement Python.		Participation aux foires et événements universitaires.	
	Infrastructure technique : serveurs locaux sécurisés.		Communication directe par e-mail et téléphone.	
Structure des Coûts		Flux de Revenus		
Serveur haute performance : 14 000 000 DA. Aménagement un bureau : 3 000 000 DA. Flash disks clients : 120 000 DA par client par Année. Frais marketing, licences et conformité légale.		Abonnement annuel : 4 800 000 DA par entreprise (400 000 DA/mois). Formation initiale : 500 000 DA par nouveau client. Contrats de maintenance et support technique annuel.		

Tableau IV.1 : Business Model Canvas d'ÉcoTrace-IA

Business Plan

Business Plan

Le Business Plan traduit la vision stratégique d'ÉcoTrace-IA en projections financières concrètes sur cinq exercices. Le modèle de revenus repose sur un abonnement annuel de 4 800 000 DA par entreprise cliente et une formation initiale de 500 000 DA par nouveau client, pour une base progressant de 8 entreprises en première année à 38 en cinquième année. L'investissement initial s'élève à 17 000 000 DA, comprenant un serveur haute performance et un espace de travail, intégralement amorti sur la durée du plan. Les tableaux suivants présentent les principales composantes financières du projet :

Amortissements

Désignation du matériel	Fonctionnalité du matériel	Prix unitaire	Année 01	Année 02	Année 03	Année 04	Année 05
PC / Serveur puissant (IA + stockage)	Traitement algorithme K-Means & données	14 000 000	14 000 000	0	0	0	0
Aménagement Bureau 1	Espace de travail	3 000 000	3 000 000	0	0	0	0

[Retour au sommaire](#)

Achats directs

			N-1 (historique)	Année 01	Année 02	Année 03	Année 04	Année 05
Produit/Service N°01	Kit livraison client (docs, câbles, packaging)	Flash disk livraison client ÉcoTrace-IA (renouvelé/an)						
	Quantité :	108		8	14	20	28	38
	Prix unitaire :	120000		120000	120000	120000	120000	120000
	Sous-Total :		0	960000	1680000	2400000	3360000	4560000
			N-1 (historique)	Année 01	Année 02	Année 03	Année 04	Année 05
Produit/Service N°02	Désignation :	Kit livraison (docs, packaging, câbles)						
	Quantité :	0		8	14	20	28	38
	Prix unitaire :	5000		5000	5000	5000	5000	5000
	Sous-Total :		0	40000	70000	100000	140000	190000
TOTAL :			0	1000000	1750000	2500000	3500000	4750000
Progression :			-	75%	43%	40%	36%	

Chiffre d'affaires

			N-1 (historique)	Année 01	Année 02	Année 03	Année 04	Année 05
Produit/Service N°01	Désignation :	Abonnement annuel ÉcoTrace-IA (400 000 DA/mois × 12)						
	Quantité :	108		8	14	20	28	38
	Prix unitaire :	4800000		4800000	4800000	4800000	4800000	4800000
	Sous-Total :		0	38400000	67200000	96000000	134400000	182400000
			N-1 (historique)	Année 01	Année 02	Année 03	Année 04	Année 05
Produit/Service N°02	Désignation :	Formation utilisation ÉcoTrace-IA (one-time / nouveau client)						
	Quantité :	38		8	6	6	8	10
	Prix unitaire :	500000		500000	500000	500000	500000	500000
	Sous-Total :		0	4000000	3000000	3000000	4000000	5000000
TOTAL :			0	42400000	70200000	99000000	138400000	187400000
Progression :			-	66%	41%	40%	35%	

Cash-flow réel

	Année 1	Année 2	Année 3	Année 4	Année 5
CA	42 400 000	70 200 000	99 000 000	138 400 000	187 400 000
– Achats	–1 000 000	–1 750 000	–2 500 000	–3 500 000	–4 750 000
– Invest. (Cash réel)	–17 000 000	0	0	0	0
= Cash-flow réel	24 400 000	68 450 000	96 500 000	134 900 000	182 650 000

Tableau IV.2 : Projections financières d'ÉcoTrace-IA (2027–2032)

Conclusion

Ce chapitre démontre qu'ÉcoTrace-IA répond à un besoin économique identifié et quantifiable. Le BMC révèle un positionnement différencié sur un marché encore peu couvert par des solutions locales, tandis que les projections financières confirment la viabilité du modèle avec un cash flow réel positif dès la première année, atteignant 24 400 000 DA. ÉcoTrace-IA s'affirme ainsi comme un projet entrepreneurial concret, ancré dans les réalités du tissu économique agroalimentaire algérien.

Conclusion générale

Conclusion générale

Au terme de ce travail, trois applications ont été conçues et développées, formant ensemble un système cohérent et progressif allant de la maîtrise algorithmique jusqu'à l'outil industriel opérationnel.

LogiClass constitue le point de départ de cette architecture. En implémentant l'algorithme K-Means dans sa forme itérative complète, cette application a permis d'observer et de valider expérimentalement la dynamique de convergence des centroïdes, posant ainsi les fondements théoriques et pratiques du système.

RT-KMC prolonge cette démarche en exploitant les centroïdes persistés par **LogiClass** pour assurer une classification en temps réel. Ce module a démontré la robustesse du mécanisme de décision par distance euclidienne, tout en offrant une interface de test interactive permettant de valider le comportement du modèle face à des observations nouvelles.

ÉcoTrace-IA représente l'aboutissement naturel de cette progression. En transposant les principes acquis dans un espace vectoriel pondéré à 16 dimensions, cette application répond directement aux besoins des PME agroalimentaires algériennes confrontées aux exigences environnementales internationales, notamment le mécanisme MACF. Elle attribue automatiquement un Éco-Grade de A à E, offrant un outil de pré-évaluation simple, autonome et immédiatement déployable.

Ces trois réalisations, bien que distinctes dans leurs objectifs, s'inscrivent dans une démarche unifiée : démontrer que l'intelligence artificielle peut constituer un levier concret et accessible au service de la transition écologique de l'industrie agroalimentaire algérienne.

Sur le plan économique, l'étude de faisabilité conduite dans le quatrième chapitre confirme la viabilité commerciale du projet. Le Business Model Canvas a mis en évidence une proposition de valeur différenciée, tandis que les projections financières sur la période 2027–2032 attestent d'un potentiel de rentabilité dès la première année, avec un cash flow réel de 24 400 000 DA en année 1, fondé sur un modèle d'abonnement annuel de 4 800 000 DA par entreprise cliente.

En définitive, **ÉcoTrace-IA** s'impose non seulement comme une contribution académique à l'intersection de l'intelligence artificielle et de l'évaluation environnementale, mais également comme une solution concrète, évolutive et à fort potentiel d'impact pour l'industrie agroalimentaire algérienne. Les perspectives d'évolution sont nombreuses : enrichissement du

Conclusion générale

référentiel de critères environnementaux, intégration de données réelles issues de PME partenaires, et développement d'une version web accessible à l'ensemble des acteurs de la filière.

Perspectives

Bien que les résultats obtenus soient hautement satisfaisants, ce projet ouvre la voie à plusieurs axes d'amélioration et d'extension :

- **Amélioration continue des trois applications et de leurs interfaces** : faire évoluer LogiClass, RT-KMC et ÉcoTrace-IA vers des versions plus abouties, avec des interfaces utilisateur plus ergonomiques, multilingues, et davantage orientées vers l'expérience de l'utilisateur final — qu'il soit chercheur, technicien ou industriel.
- **Intégration de données réelles pour ÉcoTrace-IA** : dès que des entreprises partenaires fourniront des bilans environnementaux réels, LogiClass pourra entraîner de véritables centroïdes à partir de ces données, permettant à ÉcoTrace-IA d'évoluer d'une approche déterministe vers un système appris et validé empiriquement.
- **Renforcement de la sécurité des données** : intégrer un système de chiffrement des données stockées, une authentification par mot de passe pour chaque entreprise cliente, ainsi qu'une gestion des droits d'accès (JSSP) garantissant que chaque société ne puisse consulter que ses propres données, avec une architecture de sécurité renforcée avant tout déploiement à grande échelle.

Références bibliographiques

- [1] FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture), "Greenhouse gas emissions from agrifood systems: Global, regional and national trends, 2000-2020", FAO Statistical Development Series, Rome, 2021.
- [2] Ministère de l'Environnement et des Énergies Renouvelables (Algérie), "Rapport National sur l'État et les Perspectives de l'Environnement et de la Transition Écologique en Algérie", MEER, Alger, 2020. [En ligne]. Disponible sur : <https://www.me.gov.dz> [Consulté le : mai 2025].
- [3] Organisation Internationale de Normalisation (ISO), "ISO 14040:2006 — Management environnemental — Analyse du cycle de vie — Principes et cadre", ISO, 2006.
- [4] Organisation Internationale de Normalisation (ISO), "ISO 14044:2006 — Management environnemental — Analyse du cycle de vie — Exigences et lignes directrices", ISO, 2006. [En ligne]. Disponible sur : <https://www.iso.org/fr/standard/38498.html> [Consulté le : mai 2025].
- [5] J. B. Guinée, R. Heijungs, H. A. Udo de Haes et A. W. Sleeswijk, "Life cycle assessment: past, present, and future," *Environmental Science & Technology*, vol. 45, no. 1, pp. 90–96, 2011. doi: 10.1021/es101316v
- [6] B. Bergougui, "Moving toward environmental mitigation in Algeria: Asymmetric impact of fossil fuel energy, renewable energy and technological innovation on CO₂ emissions," *Energy Strategy Reviews*, vol. 51, 2024. doi: 10.1016/j.esr.2023.101281
- [7] J. D. Sachs, *The Age of Sustainable Development*, Columbia University Press, New York, 2015.
- [8] GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat), "Résumé à l'intention des décideurs — Rapport spécial sur les conséquences d'un réchauffement planétaire de 1,5 °C," IPCC/GIEC, 2018.
- [9] Nations Unies — CCNUCC, "Adoption de l'Accord de Paris — FCCC/CP/2015/L.9," UNFCCC, 2015.
- [10] Commission Européenne, "Communication de la Commission — Le Pacte vert pour l'Europe," COM(2019) 640 final, EUR-Lex, 2019.

[11] Ministère de la Transition Écologique (France), "Mécanisme d'ajustement carbone aux frontières (MACF)," 2024.

[12] Banque de France, "La taxonomie verte européenne," 2023.

[13] Direction Générale des Entreprises (DGE), "Guide pratique CSRD," Ministère de l'Économie et des Finances, 2024.

[14] X. He, F. Zhai et J. Ma, "Global Impact of the Carbon Border Adjustment Mechanism — A Quantitative Assessment," *Journal of Globalization and Development*, vol. 16, no. 1, pp. 29–53, 2025. doi: 10.1515/jgd-2023-0098

[15] Éco-conception.fr, "Analyse du Cycle de Vie (ACV) : méthodologie et phases", *Éco-conception.fr*, 2022. [En ligne]. Disponible sur : <https://www.eco-conception.fr/static/analyse-du-cycle-de-vie-acv.html> [Consulté le : mai 2025].

[16] INRAE / Agribalyse, "Méthodologie ACV — Documentation Agribalyse", *Agribalyse*, 2023. [En ligne]. Disponible sur : <https://doc.agribalyse.fr/documentation/lesdonnees/methodologie-acv> [Consulté le : mai 2025].

[17] K. Hadasik et al., "Sustainable Development in the Agri-Food Sector in Terms of the Carbon Footprint: A Review," *Sustainability*, vol. 12, no. 16, 6463, 2020. doi: 10.3390/su12166463

[18] HelloCarbo, "Analyse du Cycle de Vie (ACV) : définition, méthode et calcul", *HelloCarbo*, 2023. [En ligne]. Disponible sur : <https://www.hellocarbo.com/blog/calculer/acv/> [Consulté le : mai 2025].

[19] ESU-Services, "SimaPro — Logiciel d'Analyse du Cycle de Vie," *ESU-Services*, [En ligne]. Disponible sur : <https://esu-services.ch/fr/simapro/> [Consulté le : mai 2025].

[20] Consultis Environnement, "OpenLCA et Ecoinvent : outils d'analyse environnementale," *Consultis Environnement*, [En ligne]. Disponible sur : <https://consultis-environnement.com/articles/numerique-responsable/openlca-et-ecoinvent/> [Consulté le : mai 2025].

- [21] Rescoll, "GaBi — Logiciel d'Éco-conception et d'ACV," *Eco-conception.fr*, [En ligne]. Disponible sur : <https://www.eco-conception.fr/data/sources/users/471/7-rescoll.pdf> [Consulté le : mai 2025].
- [22] I. Goodfellow, Y. Bengio et A. Courville, *Deep Learning*, MIT Press, Cambridge, MA, 2016. [En ligne]. Disponible sur : <http://www.deeplearningbook.org>
- [23] A. Zhang, Z. C. Lipton, M. Li et A. J. Smola, *Dive into Deep Learning*, Cambridge University Press, 2023. [En ligne]. Disponible sur : <https://d2l.ai>
- [24] Y. LeCun, Y. Bengio et G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539
- [25] VectorMine, "Machine vs Deep Learning as Artificial Intelligence Principle Outline Diagram," *VectorMine*, 2023. [En ligne]. Disponible sur : <https://vectormine.com/item/machine-vs-deep-learning-as-artificial-intelligence-principle-outline-diagram/> [Consulté le : mai 2025].
- [26] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.
- [27] Codefinity, "Les paradigmes de l'apprentissage automatique," *Codefinity*, 2024. [En ligne]. Disponible sur : <https://codefinity.com/fr/courses/v2/5f9425ad-7012-42d1-9fc0-d5998b1f5fd4/820c63cb-b88e-423f-ba1d-7afb5ffba163/587fcd5d-7022-4ee0-9601-59c5a1841eee> [Consulté le : mai 2025].
- [28] W. S. McCulloch et W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133, 1943.
- [29] CoolIT, "Intelligence artificielle : Histoire et réalités", *CoolIT*, 2023. [En ligne]. Disponible sur : <https://coolit.bzh/intelligence-artificielle-histoire-realites/> [Consulté le : mai 2025].
- [30] V. Nair et G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," *Proceedings of the 27th International Conference on Machine Learning (ICML)*, pp. 807–814, 2010.
- [31] Edem Dogbe, "Activation Functions in Machine Learning: The Hidden Switch that Makes Neural Networks Think," *AI — Plain English*, Medium, 2024. [En ligne]. Disponible sur :

<https://ai.plainenglish.io/activation-functions-in-machine-learning-the-hidden-switch-that-makes-neural-networks-think-5e16efb96702> [Consulté le : mai 2025].

[32] CNAM — Vertigo, "Session 1 — Practical 1 : Réseaux de Neurones," *CNAM*, 2024. [En ligne]. Disponible sur : https://cedric.cnam.fr/vertigo/Cours/US3362-Apprentissage/session1_practical1.html [Consulté le : mai 2025].

[33] K. Hornik, M. Stinchcombe et H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989. doi: 10.1016/0893-6080(89)90020-8

[34] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, no. 4, pp. 303–314, 1989. doi: 10.1007/BF02551274

[35] CNAM — Vertigo, "Cours : Réseaux de Neurones Multicouches," *CNAM*, 2024. [En ligne]. Disponible sur : <https://cedric.cnam.fr/vertigo/Cours/ml/coursReseauxNeuronesMulticouches.html> [Consulté le : mai 2025].

[36] D. E. Rumelhart, G. E. Hinton et R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986. doi: 10.1038/323533a0

[37] D. P. Kingma et J. Ba, "Adam: A method for stochastic optimization," *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015. arXiv:1412.6980

[38] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982. doi: 10.1109/TIT.1982.1056489. [En ligne]. Disponible sur : <https://www.stat.cmu.edu/~rnugent/PCMI2016/papers/LloydKMeans.pdf> [Consulté le : mai 2025].

[39] NVIDIA Corporation, "K-Means Clustering," *NVIDIA Glossary*, [En ligne]. Disponible sur : <https://www.nvidia.com/en-eu/glossary/k-means/> [Consulté le : mai 2025].

[40] D. Arthur et S. Vassilvitskii, "K-Means++: The Advantages of Careful Seeding," *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1027–1035, 2007.

[41] T. Peng et al., "Machine learning for life cycle assessment of agri-food products: A systematic review," *Journal of Cleaner Production*, vol. 341, 2022. doi: 10.1016/j.jclepro.2022.130910

[42] A. Osterwalder et Y. Pigneur, *Business Model Generation : A Handbook for Visionaries, Game Changers, and Challengers*, John Wiley & Sons, Hoboken, NJ, 2010, p. 14.