



Université Mohammed-V de Rabat
École Nationale Supérieure d'Informatique
et d'Analyse des Systèmes



Prévision de la Demande Internationale des Produits du Terroir Marocain

Data Driven Decision Making

Réalisé par :

AMELLOUK Nour

BAOU Zakaria

Encadré par :

M. TABII Youness

Année Académique 2023/2024

Introduction

Résumé

Ce projet vise à prévoir la demande des produits du terroir marocain principalement sur le marché international, ainsi qu'à définir des stratégies efficaces pour leur développement et commercialisation. Les données seront collectées à partir de sites officiels et subiront une phase de nettoyage approfondie pour identifier les produits pertinents et les caractéristiques clés pour la formulation des stratégies.

Nous effectuerons une analyse approfondie des données et explorerons diverses techniques d'apprentissage automatique pour déterminer le meilleur modèle possible. Le choix des algorithmes sera justifié, et la visualisation des données sera obligatoire pour présenter les mesures de performance et les insights basés sur différents axes. L'étude fournira des stratégies exploitables basées sur une analyse des données pour améliorer la présence sur le marché et la compétitivité des produits du terroir marocain.

Abstract

This project aims to forecast the demand for Moroccan local products mainly in international markets and to develop effective strategies for their development and commercialization. Data will be collected from official websites and undergo a thorough cleaning phase to identify relevant products and key characteristics for strategy formulation.

We will conduct deep data analysis and explore various machine learning techniques to determine the best possible model. The choice of algorithms will be justified, and data visualization will be mandatory to present performance metrics and insights based on various axes. The study will provide actionable strategies grounded in data-driven analysis to enhance the market presence and competitiveness of Moroccan local products.

Introduction

Les produits du terroir marocain, tels que l'huile d'argan, les dattes, le safran et les épices traditionnelles, représentent la richesse culturelle et gastronomique du Maroc. Souvent cultivés de manière artisanale et respectueuse de l'environnement, ces produits jouent un rôle clé dans la préservation des traditions locales et le soutien économique des communautés rurales.

L'économie marocaine valorise ces produits pour leur potentiel à stimuler la croissance économique, diversifier les exportations et promouvoir la culture marocaine à l'international. La demande mondiale croissante pour des produits naturels et authentiques renforce l'attrait des produits du terroir marocain, offrant ainsi des opportunités significatives pour les producteurs locaux de pénétrer de nouveaux marchés et d'améliorer leurs revenus.

Étudier la demande pour ces produits, tant sur le marché national qu'international, est donc essentiel pour comprendre les dynamiques actuelles et futures de ce secteur. En prévoyant cette demande, nous pouvons développer des stratégies de commercialisation efficaces, adaptées aux besoins et aux préférences des consommateurs. Ces stratégies doivent être fondées sur des données précises et pertinentes, collectées à partir de sources officielles.

Notre projet débutera par une phase de collecte et de nettoyage des données pour identifier les produits concernés et leurs caractéristiques spécifiques. Nous justifierons le choix des algorithmes de prévision en fonction des caractéristiques des produits et des exigences du marché. Une visualisation des données sera réalisée pour illustrer les tendances et les résultats obtenus, facilitant ainsi la compréhension des insights dérivés.

Les résultats de cette analyse nous permettront de formuler des recommandations stratégiques pour le développement et la commercialisation des produits du terroir marocain. En optimisant les approches de vente et en ciblant efficacement les marchés potentiels, ce projet vise à renforcer la compétitivité des produits du terroir marocain et à stimuler leur croissance sur la scène internationale.

Table des matières

| | | |
|----------|--|-----------|
| 1 | Analyse exploratoire des données | 6 |
| 1.1 | Sources de données | 6 |
| 1.2 | Description des données et Identification des produits | 7 |
| 1.3 | Traitement de données | 7 |
| 1.4 | Analyse des données et Visualisation | 8 |
| 1.4.1 | Produits agricoles | 9 |
| 1.4.2 | Coton | 10 |
| 1.4.3 | Cuir | 11 |
| 1.4.4 | Produits alimentaires transformés et produits agroalimentaires | 12 |
| 1.4.5 | Total | 13 |
| 2 | Méthodologie de Prévision et Résultats | 15 |
| 2.1 | Algorithmes de prévision sélectionnés | 15 |
| 2.1.1 | LinearRegression | 15 |
| 2.1.2 | XGBRegressor | 15 |
| 2.1.3 | LSTM | 16 |
| 2.2 | Métrique de performance | 17 |
| 2.2.1 | Explained Variance Score (EVS) | 17 |
| 2.2.2 | R-Squared (R2) | 18 |
| 2.2.3 | Mean Gamma Deviance (MGD) | 19 |
| 2.2.4 | Mean Poisson Deviance (MPD) | 20 |
| 2.3 | Résultats et analyse | 20 |
| 2.3.1 | LinearRegression | 20 |
| 2.3.2 | XGBRegressor | 20 |
| 2.3.3 | LSTM | 21 |
| 2.3.4 | Comparaison des modèles | 22 |
| 2.4 | Prédiction du modèle | 22 |
| 3 | Stratégies de Développement et de Commercialisation | 24 |
| 3.1 | Synthèse des résultats et perspectives futures | 24 |

Table des figures

| | | |
|-----|--|----|
| 1.1 | Diagrammes en boîte a moustache des valeurs d'exportation pour différentes catégories de produits | 9 |
| 1.2 | Tendances D'exportation Des Produits Agricoles | 10 |
| 1.3 | Tendances D'exportation Du Coton | 11 |
| 1.4 | Tendances D'exportation Du Cuir | 11 |
| 1.5 | Tendances D'exportation Des Produits alimentaires transformés et produits agroalimentaires | 12 |
| 1.6 | Tendances D'exportation Des Produits de terroir marocain | 13 |
| 1.7 | Pourcentage de contribution de chaque catégorie de produit dans le total des produits de terroir en 2022 | 14 |
| 1.8 | Contribution de chaque catégorie de produit dans le total des produits de terroir en 2022 | 14 |
| 2.1 | Comparaison entre les valeurs originales et les valeurs prédites par LinearRegression | 21 |
| 2.2 | Comparaison entre les valeurs originales et les valeurs prédites par XGBRegressor | 21 |
| 2.3 | Comparaison entre les valeurs originales et les valeurs prédites par LSTM . . . | 22 |
| 2.4 | 15 dernières séries chronologiques trimestrielles VS 30 prochaines séries chronologiques trimestrielles | 23 |
| 2.5 | Tracer de la valeur entière avec prédiction | 23 |

Liste des tableaux

| | | |
|-----|---|----|
| 1.1 | Aperçu des exportations de produits du terroir marocain (2000-2023) | 8 |
| 1.2 | Statistiques descriptives des donnees | 8 |
| 2.1 | Model Performance Comparison | 22 |

Chapitre 1

Analyse exploratoire des données

1.1 Sources de données

Pour la source des données, après plusieurs tentatives de recherche sur les sites gouvernementaux, aucune piste concluante n'a été trouvée, surtout dans le cas des produits du terroir marocains. Au niveau national, les informations disponibles se limitaient principalement à des articles de presse et des documents similaires, qui ne contenaient pas de données exploitables pour notre projet. Cette absence de données fiables et détaillées a constitué un obstacle majeur, limitant notre capacité à mener une analyse approfondie du marché national des produits du terroir. La difficulté de trouver des données structurées et pertinentes au niveau national a mis en évidence un manque de centralisation et de transparence des informations commerciales sur ces produits spécifiques.

Face à ces défis, nous avons élargi notre recherche de données à l'échelle internationale. Grâce à **L'Office des Changes**, qui est un établissement marocain public à caractère administratif, et à travers le site trademap.org, un site spécialisé dans les statistiques du commerce pour le développement international des entreprises, nous avons pu obtenir des données précieuses. L'Office des Changes, placé sous la tutelle du Ministère de l'économie et des finances, fournit des statistiques commerciales mensuelles, trimestrielles et annuelles, couvrant les importations et exportations en termes de valeurs et volumes, ainsi que les taux de croissance et parts de marché.

Ces données nous ont permis de mener une étude au niveau international, de mieux comprendre les dynamiques de demande et d'élaborer des stratégies de développement et de commercialisation basées sur des faits concrets. Les informations détaillées surtout les exportations de ces produits ont fourni une base solide pour notre analyse. Ainsi, malgré les défis rencontrés au niveau national, l'intégration des données internationales a renforcé la pertinence et l'exhaustivité de notre étude, nous permettant de proposer des stratégies éclairées pour accroître la compétitivité des produits du terroir marocains sur les marchés mondiaux.

1.2 Description des données et Identification des produits

Dans le cadre de notre analyse, nous avons utilisé plusieurs tables de données disponibles sur les sites de TradeMap et de l'Office des Changes, couvrant la période de 2000 à 2023. Nous avons choisi de nous concentrer sur les tables des valeurs, représentant la valeur monétaire des produits exportés par le Maroc chaque année de 2000 à 2023. Ces valeurs, mesurées en dollars américains (USD), indiquent les revenus totaux générés par l'exportation de ces produits pour une année donnée. Cette mesure fournit des informations directes sur le montant des exportations et la dynamique de leur évolution.

Le dataset comprend 97 groupes de produits couvrant une vaste gamme de catégories. Ces groupes incluent des biens essentiels tels que les fertilisants, les produits alimentaires (fruits, légumes, viande, poissons, céréales), et les boissons, ainsi que des articles manufacturés comme les véhicules, les machines électriques, et les équipements optiques et médicaux. Les matériaux de construction tels que le ciment, le plâtre, et les produits en métal (fer, acier, aluminium, cuivre) y figurent également. Les produits chimiques, incluant les produits pharmaceutiques et les préparations diverses, occupent une place importante, tout comme les textiles (vêtements, tissus spéciaux, fibres synthétiques). D'autres catégories notables incluent les articles de consommation courante comme les meubles, les articles en plastique, et les produits en caoutchouc.

Nous avons opté pour des séries temporelles trimestrielles, segmentant chaque année en quatre périodes. Cette granularité permet une analyse plus fine des tendances saisonnières et des variations intra-annuelles.

Pour se concentrer sur notre projet, qui porte principalement sur les produits du terroir marocain, nous avons éliminé toutes les catégories de produits n'ayant aucune relation avec notre sujet. Ainsi, nous avons restreint notre analyse à quatre catégories de produits où les produits du terroir marocain sont prédominants : les **produits agricoles**, le **coton**, le **cuir**, et les **aliments transformés et produits agro-alimentaires**.

1.3 Traitement de données

En téléchargeant les fichiers, nous avons obtenu un total de 20 documents .txt, répartis sur quatre catégories : les produits agricoles, le coton, le cuir, et les aliments transformés et produits agro-alimentaires. Chaque catégorie comprenait cinq documents couvrant des périodes de temps différentes. Pour structurer les données de manière cohérente, nous avons d'abord fusionné les documents pour chaque catégorie, créant ainsi des datasets complets pour chaque catégorie couvrant la période de 2000 à 2023.

Après avoir consolidé les données par catégorie, nous avons fusionné ces quatre catégories en un seul dataset global. Ce dataset intégral représente l'ensemble des exportations de produits du terroir marocain sur la période de 2000 à 2023. Pour enrichir notre analyse, nous avons ajouté une colonne supplémentaire indiquant le total des valeurs exportées, permettant ainsi une vue d'ensemble des exportations des produits du terroir marocain sur cette période. Cette approche méthodique nous a permis de centraliser et de synthétiser les données, facilitant ainsi une analyse approfondie et précise des tendances d'exportation.

| Period | Agricultural products | Cotton | Leather | Processed food... | Total |
|---------|-----------------------|--------|---------|-------------------|----------|
| 2000-Q1 | 332441.0 | 169.0 | 58295.0 | 304941.0 | 695846.0 |
| 2000-Q2 | 224801.0 | 244.0 | 55548.0 | 200588.0 | 481181.0 |
| 2000-Q3 | 127046.0 | 23.0 | 63902.0 | 216606.0 | 407577.0 |
| 2000-Q4 | 232257.0 | 0.0 | 58893.0 | 277342.0 | 568492.0 |
| 2001-Q1 | 301359.0 | 72.0 | 65345.0 | 240431.0 | 607207.0 |

TABLE 1.1 – Aperçu des exportations de produits du terroir marocain (2000-2023)

1.4 Analyse des données et Visualisation

Mener une étude d'analyse de données sur les exportations des produits de terroir marocain est essentiel pour prédire la demande du marché international. Cette analyse permet de décrypter les tendances et les préférences des consommateurs, fournissant aux producteurs marocains des insights cruciaux pour adapter leur offre. Identifier les produits les plus prisés et les marchés les plus prometteurs aide à élaborer des stratégies d'exportation optimisées, augmentant ainsi les revenus. Prédire la demande grâce à des données fiables et des modèles statistiques permet d'anticiper les fluctuations du marché, réduisant les risques de volatilité des ventes. Une telle étude améliore également la gestion des ressources et la planification, renforçant la compétitivité des produits de terroir marocains à l'international.

| | Agricultural Products | Cotton | Leather | Processed Food... | Total |
|-------|-----------------------|--------|----------|-------------------|----------|
| Count | 94 | 94 | 94 | 94 | 94 |
| Mean | 814227.1 | 29.0 | 90058.4 | 531264.9 | 1435579 |
| Std | 484119.2 | 66.1 | 18846.3 | 225590 | 701155 |
| Min | 127046 | 0 | 35914 | 152765 | 407577 |
| 25% | 442440.8 | 0 | 79998.3 | 340247 | 873366.8 |
| 50% | 700444 | 1 | 92831 | 508790.5 | 1371174 |
| 75% | 1069214 | 25 | 102630.8 | 702016.8 | 1845925 |
| Max | 2120093 | 467 | 125270 | 1060766 | 3288152 |

TABLE 1.2 – Statistiques descriptives des données

Ce tableau présente les statistiques descriptives des valeurs d'exportation pour différentes catégories de produits de terroir marocains. Les produits agricoles montrent une moyenne élevée de 814227,1 avec une grande variabilité (écart-type de 484119,2). Les valeurs vont de 127046 à 2120093, avec une médiane de 700444, indiquant une distribution largement étalée. Le coton, en revanche, a des valeurs d'exportation très faibles avec une moyenne de 29,0 et une médiane de seulement 1, et une variabilité modérée (écart-type de 66,1). Le cuir présente une moyenne de 90058,4 et un écart-type de 18846,3, avec des valeurs allant de 35914 à 125270, et une médiane de 92831, montrant une certaine constance dans les exportations. Les produits alimentaires transformés et agro-industriels ont une moyenne de 531264,9 et un écart-type de 225590, avec des valeurs allant de 152765 à 1060766, et une médiane de 508790,5, indiquant une distribution relativement large mais moins variable que les produits agricoles. Le total des valeurs d'exportations des produits de terroir marocains présente une moyenne de 1435579 avec un écart-type de 701155, allant de 407577 à 3288152, et une médiane de 1371174, suggérant une large gamme de valeurs d'exportation avec des fluctuations importantes.

Diagrammes en boîte des valeurs d'exportation pour différentes catégories de produits

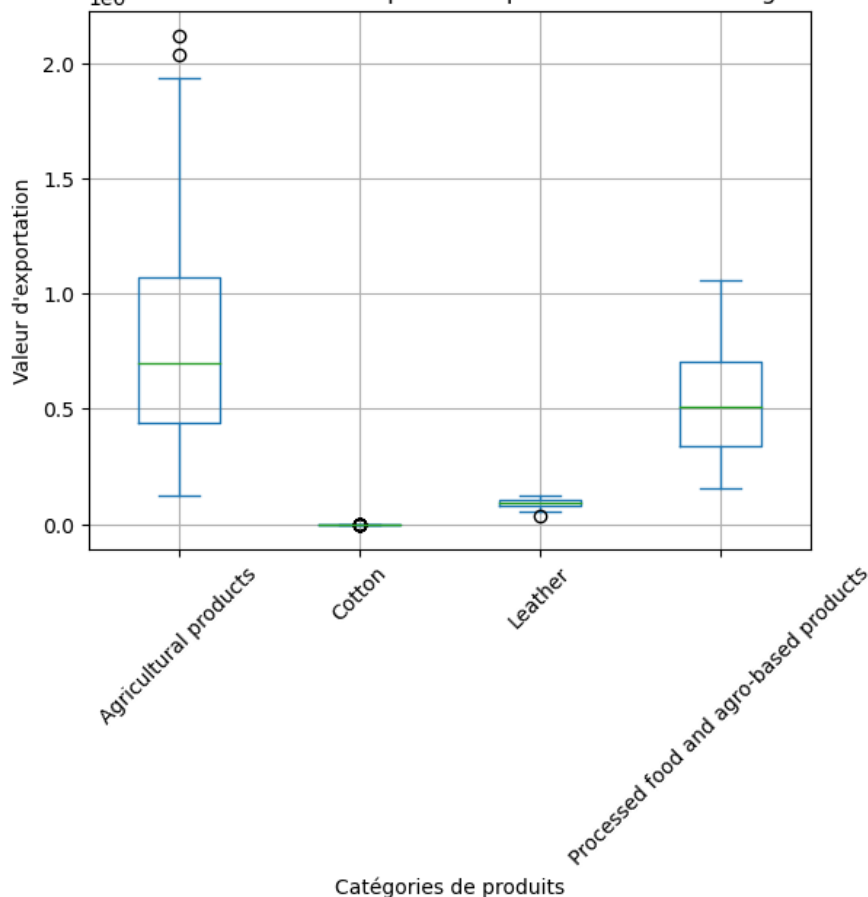


FIGURE 1.1 – Diagrammes en boîte a moustache des valeurs d’exportation pour différentes catégories de produits

Le diagramme à moustaches montre une large gamme de valeurs d’exportation pour chaque catégorie de produits. Les produits agricoles ont une médiane élevée et une large distribution, indiquant des exportations significatives mais aussi une grande variabilité dans les valeurs exportées. Les valeurs interquartiles pour cette catégorie montrent une dispersion notable, avec plusieurs valeurs aberrantes élevées, soulignant des périodes de forte exportation. Le coton, par contre, affiche des valeurs d’exportation très faibles avec une médiane proche de zéro et une distribution très serrée, suggérant une exportation très limitée et stable. Le cuir a une médiane modérée avec une distribution légèrement plus large que le coton, mais reste beaucoup moins variable que les produits agricoles. Les produits alimentaires transformés et agro-industriels montrent une médiane relativement haute et une dispersion comparable aux produits agricoles, avec plusieurs valeurs aberrantes vers le haut, ce qui indique des fluctuations importantes dans les volumes d’exportation.

1.4.1 Produits agricoles

Les tendances d’exportation des produits agricoles ont montré une trajectoire ascendante remarquable. À partir d’un niveau de base plus bas en 2000, on observe une montée régulière avec des pics et des creux périodiques, indiquant des variations saisonnières typiques des cycles agricoles.

Malgré ces fluctuations, la tendance générale est à la croissance significative, suggérant des améliorations dans la productivité agricole, la capacité d'exportation et possiblement la demande du marché. Cette tendance à la hausse pourrait également refléter les avancées technologiques dans le secteur agricole, de meilleurs accords commerciaux internationaux et une augmentation de la consommation mondiale de produits agricoles.

En 2024, la valeur des produits agricoles exportée a plus que triplé par rapport aux niveaux observés en 2000, démontrant une croissance robuste et soutenue du secteur des exportations agricoles sur une période de 24 ans.

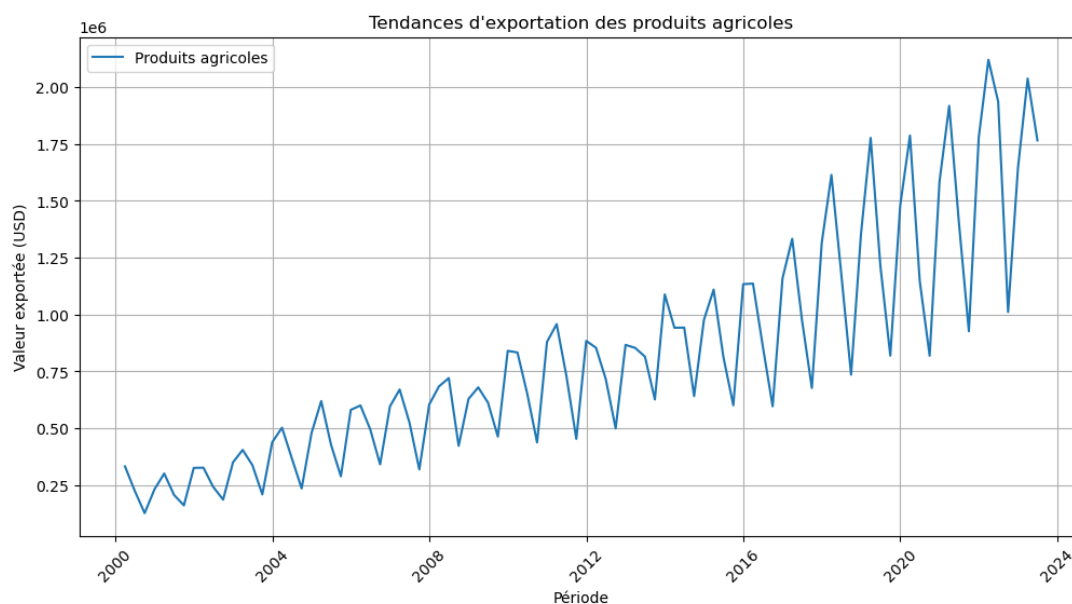


FIGURE 1.2 – Tendances D'exportation Des Produits Agricoles

1.4.2 Coton

Dans le cas du coton, les données montrent une grande variabilité avec un écart type (66,14) qui dépasse la moyenne (29,02). La valeur maximale (467) est nettement supérieure à la médiane (1), indiquant quelques valeurs exceptionnellement élevées.

Nous avons des variations importantes avec des périodes de forte activité suivies de longues périodes de faible valeur d'exportation. Au début des années 2000, les valeurs d'exportations de coton étaient relativement élevées, atteignant des pics autour de 200 USD exportées. Cependant, cette tendance a rapidement décliné, avec des fluctuations sporadiques et des périodes prolongées de faibles exportations. Entre 2008 et 2018, les valeurs d'exportations de coton sont restées presque négligeables.

À partir de 2019, une reprise graduelle est observée, culminant en 2024 avec une augmentation spectaculaire des valeurs d'exportations de coton, atteignant plus de 400 USD, avant de redescendre brusquement.

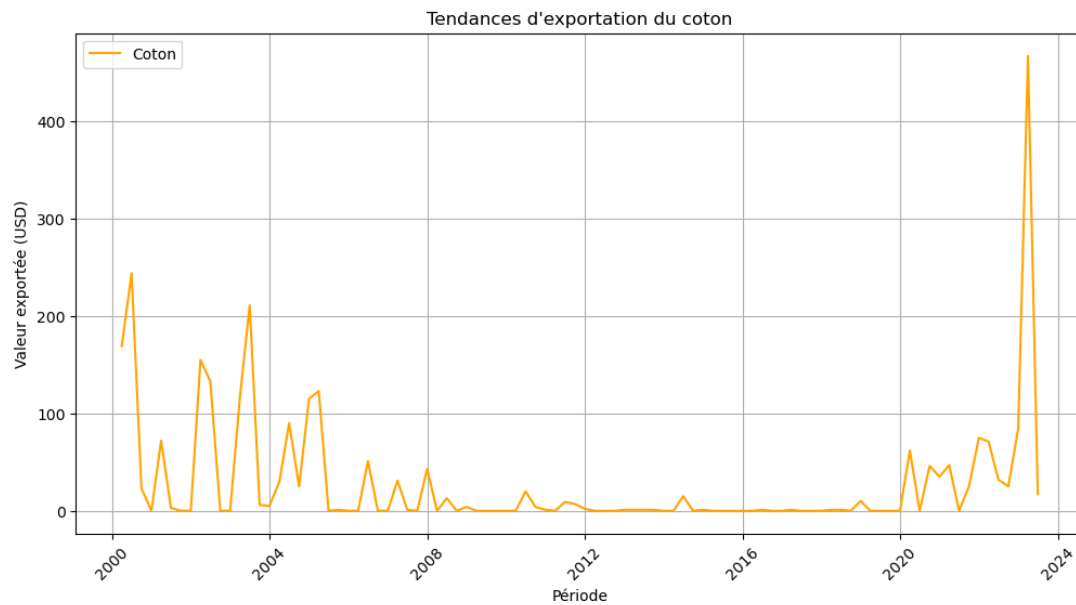


FIGURE 1.3 – Tendances D'exportation Du Coton

1.4.3 Cuir

La valeur exporté du cuir a connu une croissance générale avec des fluctuations notables. Au début des années 2000, les valeurs d'exportations de cuir se situaient autour de 60 000 USD.

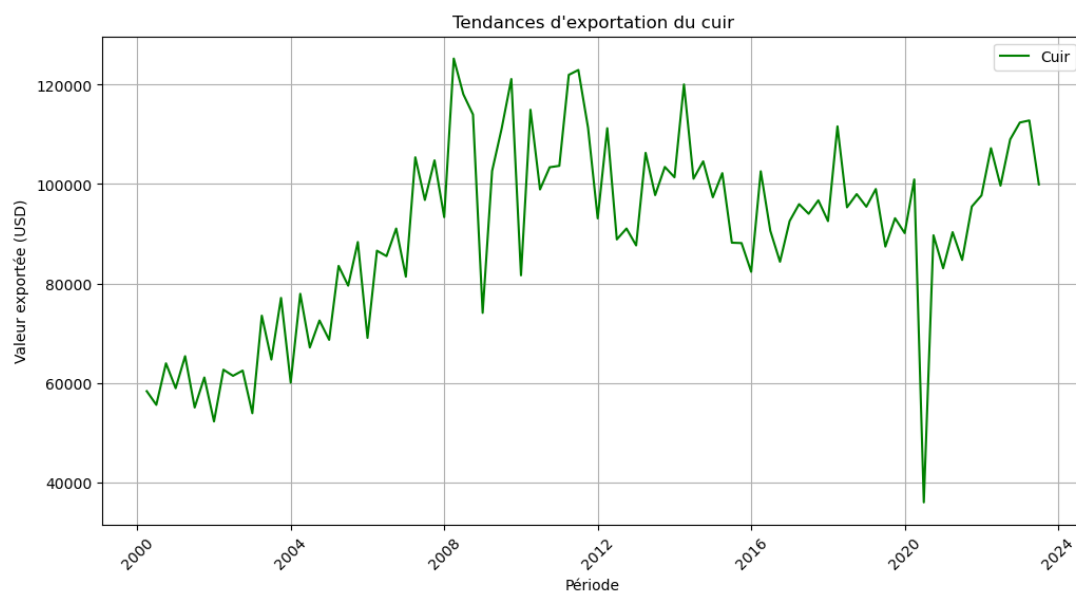


FIGURE 1.4 – Tendances D'exportation Du Cuir

On observe une augmentation progressive jusqu'en 2008, où les valeurs d'exportations atteignent leur premier sommet à plus de 120 000 USD. Entre 2008 et 2012, les valeurs d'exportations restent élevées mais montrent des variations importantes, indiquant une instabilité dans le marché ou la production.

Après 2012, bien que les valeurs d'exportations continuent de fluctuer, elles maintiennent généralement un niveau élevé autour de 100 000 USD. Un déclin marqué est visible autour de 2019-2020, probablement dû à des perturbations économiques ou de production, mais les valeurs d'exportations se redressent rapidement par la suite.

En 2024, les niveaux de valeurs d'exportation de cuir restent robustes, proches des pics précédents.

Les valeurs de cuir sont plus concentrées autour de la moyenne (90 058,40) avec un écart type modéré (18 846,34). La médiane (92 831) est proche de la moyenne, ce qui indique une distribution relativement symétrique des valeurs.

1.4.4 Produits alimentaires transformés et produits agroalimentaires

les tendances de valeurs d'exportation des produits alimentaires transformés et agroalimentaires montrent une augmentation constante avec des fluctuations saisonnières.

Au début des années 2000, les valeurs d'exportations se situaient autour de 200 000 USD. Une augmentation graduelle est observée jusqu'en 2008, où les valeurs d'exportations atteignent un premier pic à environ 600 000 USD.

Après une légère baisse, la tendance reprend une trajectoire ascendante régulière à partir de 2012. Les valeurs d'exportations continuent de croître, atteignant régulièrement des niveaux supérieurs à 800 000 USD à partir de 2018 et culminant à plus de 1 000 000 USD en 2024, malgré quelques fluctuations.

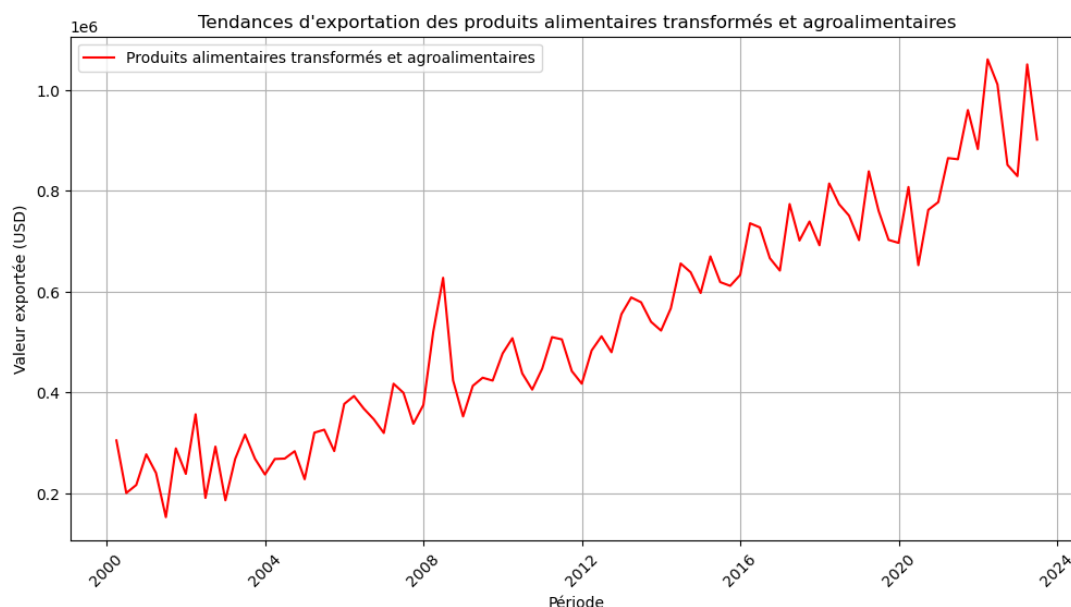


FIGURE 1.5 – Tendances D'exportation Des Produits alimentaires transformés et produits agroalimentaires

Ces produits présentent une moyenne (531 264,90) et un écart type (225 590) relativement élevés, suggérant une variabilité importante. La médiane (508 790,50) est légèrement inférieure

à la moyenne.

1.4.5 Total

On observe une nette augmentation des valeurs exportées au fil du temps, passant d'environ 0,5 million USD au début des années 2000 à plus de 3 millions USD en 2024. Cette croissance soutenue reflète une augmentation de la capacité d'exportation et de la demande internationale pour les produits marocains. Les fluctuations saisonnières sont également visibles, avec des pics et des creux réguliers chaque année, ce qui peut être attribué aux cycles de production et aux variations de la demande mondiale.

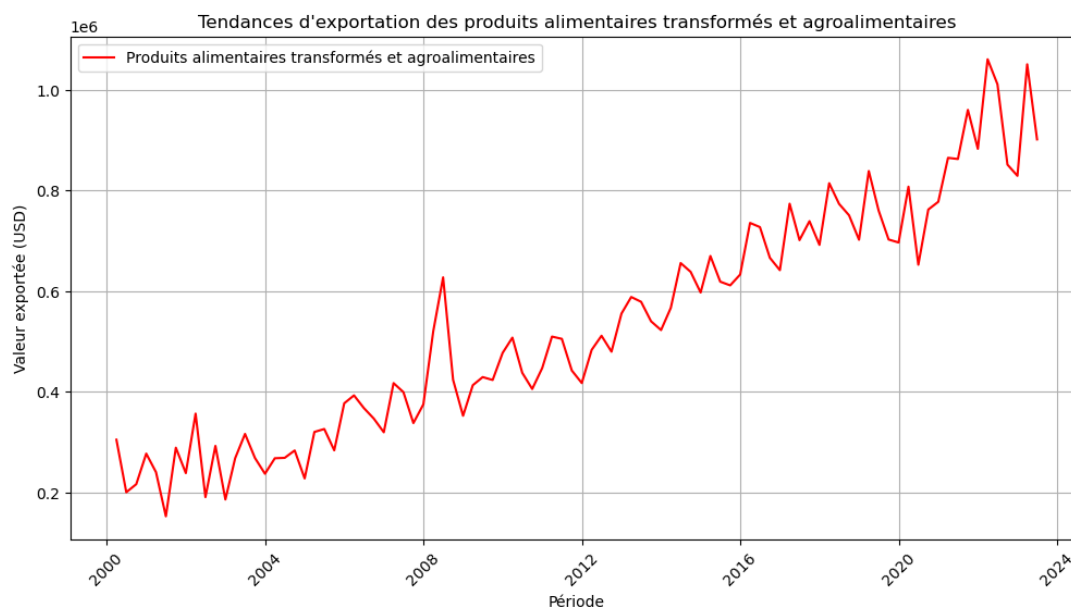


FIGURE 1.6 – Tendances D'exportation Des Produits de terroir marocain

Les données statistiques précédentes corroborent cette tendance à la hausse, avec une moyenne des exportations totales de 1,44 million USD et un écart-type de 701155 USD, indiquant une variabilité considérable. La médiane des exportations totales, à 1,37 million USD, est proche de la moyenne, ce qui suggère une distribution relativement symétrique autour de la moyenne. Les produits agricoles et les produits alimentaires transformés dominent les exportations, contribuant de manière significative à cette croissance globale.

En 2022, les produits agricoles représentent la majeure partie des exportations de produits de terroir marocains, avec une contribution de 61,6%. Cela souligne l'importance des produits agricoles dans l'économie d'exportation du Maroc.

Les produits alimentaires transformés et agro-industriels constituent le deuxième groupe le plus important, contribuant à hauteur de 34,5% du total des exportations. Cette part significative reflète la valeur ajoutée par la transformation et le traitement des produits agricoles.

En revanche, le cuir ne représente que 3,9% des exportations totales, ce qui indique une contribution modeste mais néanmoins notable.

Pourcentage de contribution de chaque catégorie de produit dans le total des produits de terroir en 2022

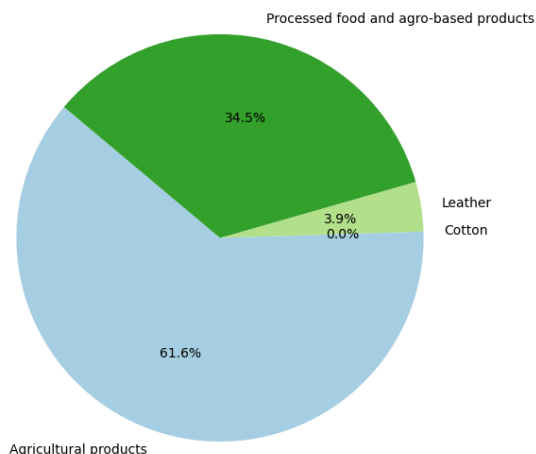


FIGURE 1.7 – Pourcentage de contribution de chaque catégorie de produit dans le total des produits de terroir en 2022

Le coton, avec une contribution proche de 0%, joue un rôle négligeable dans le total des exportations. Ces données montrent clairement que les exportations marocaines sont dominées par les produits agricoles et alimentaires transformés, qui constituent l'essentiel de l'activité exportatrice du pays.

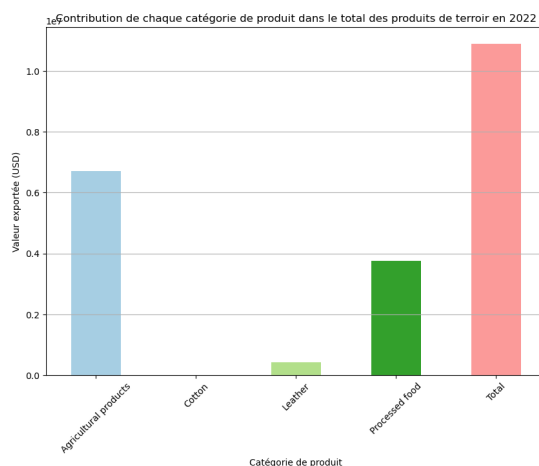


FIGURE 1.8 – Contribution de chaque catégorie de produit dans le total des produits de terroir en 2022

Chapitre 2

Méthodologie de Prévision et Résultats

2.1 Algorithmes de prévision sélectionnés

2.1.1 LinearRegression

La régression linéaire est une méthode statistique utilisée pour établir une relation linéaire entre une variable dépendante y et une variable indépendante x . Mathématiquement, cette relation est représentée par l'équation de la droite :

$$y = \beta_0 + \beta_1 x + \epsilon$$

où β_0 est l'ordonnée à l'origine (intercept), β_1 est le coefficient de régression (pente), et ϵ est le terme d'erreur. Pour estimer les valeurs de β_0 et β_1 , on utilise la méthode des moindres carrés, qui minimise la somme des carrés des erreurs (SSE) :

$$SSE = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2$$

Les formules pour les coefficients sont :

$$\beta_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

et

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

où \bar{x} et \bar{y} sont les moyennes de x et y . La régression linéaire permet ainsi de prédire y en fonction de x et de quantifier la force de cette relation.

2.1.2 XGBRegressor

XGBRegressor est un algorithme de régression basé sur l'ensemble des méthodes de boosting de gradient, implémenté dans la bibliothèque XGBoost. XGBoost est réputé pour sa performance et son efficacité, et il est couramment utilisé dans les compétitions de machine learning.

Le modèle XGBRegressor peut être défini mathématiquement comme une somme pondérée d'arbres de décision :

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i)$$

où \hat{y}_i est la prédiction pour l'observation i , K est le nombre total d'arbres, et f_k représente l'arbre de décision k . Chaque f_k appartient à l'ensemble des arbres de décision \mathcal{F} .

L'objectif de XGBRegressor est de minimiser une fonction de coût composée d'une fonction de perte et d'un terme de régularisation :

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

où l est la fonction de perte (par exemple, l'erreur quadratique), Ω est le terme de régularisation pour éviter le surapprentissage, et ϕ représente les paramètres du modèle.

Le terme de régularisation Ω pour un arbre de décision est défini comme suit :

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

où T est le nombre de feuilles dans l'arbre, γ est un paramètre de régularisation qui contrôle la complexité du modèle, λ est un autre paramètre de régularisation, et w_j est le poids de la feuille j .

XGBRegressor utilise une approche itérative pour ajouter des arbres de décision afin de corriger les erreurs résiduelles des prédictions précédentes. À chaque itération t , un nouvel arbre f_t est ajouté pour minimiser la fonction de coût :

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t)$$

où $\hat{y}_i^{(t-1)}$ est la prédiction à l'itération précédente.

XGBRegressor combine des arbres de décision de manière additive en utilisant le boosting de gradient, tout en incorporant des termes de régularisation pour éviter le surapprentissage et améliorer la généralisation.

2.1.3 LSTM

Les réseaux de neurones à mémoire à long terme (LSTM, pour Long Short-Term Memory) sont une variante des réseaux de neurones récurrents (RNN) conçue pour mieux capturer les dépendances à long terme dans les données séquentielles. Les LSTM sont particulièrement efficaces pour traiter et prédire des séries temporelles, du texte et d'autres types de séquences.

Un bloc LSTM se compose de plusieurs cellules de mémoire et de trois portes (gate) principales : la porte d'entrée, la porte de sortie et la porte d'oubli. Ces portes régulent le flux d'information dans le réseau.

1. **Porte d'oubli** : La porte d'oubli décide quelles informations de l'état précédent h_{t-1} doivent être oubliées. Elle est définie par :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

où : - f_t est le vecteur de la porte d'oubli. - σ est la fonction sigmoïde. - W_f et b_f sont les poids et biais associés.

2. Porte d'entrée : La porte d'entrée décide quelles nouvelles informations seront stockées dans l'état de la cellule. Elle comprend deux étapes :

- La fonction sigmoïde décide quelles valeurs seront mises à jour :

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

- La fonction tangente hyperbolique crée un vecteur de nouvelles valeurs candidates \tilde{C}_t qui pourraient être ajoutées à l'état de la cellule :

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

3. Mise à jour de l'état de la cellule : L'état de la cellule C_t est mis à jour en combinant l'état précédent, modulé par la porte d'oubli, et les nouvelles valeurs candidates, modulées par la porte d'entrée :

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

où \odot représente le produit élément par élément (Hadamard).

4. Porte de sortie : La porte de sortie décide quelles parties de l'état de la cellule seront sorties. Elle est définie par :

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

Le vecteur de sortie h_t est ensuite calculé en appliquant la fonction tangente hyperbolique à l'état de la cellule mis à jour, puis en le modulant par la sortie de la porte de sortie :

$$h_t = o_t \odot \tanh(C_t)$$

les LSTM utilisent des portes pour contrôler le flux d'information et capturer les dépendances à long terme dans les données séquentielles. Les formules mathématiques des LSTM permettent de réguler l'oubli, l'ajout et la sortie d'informations, ce qui les rend particulièrement efficaces pour des tâches telles que la modélisation des séries temporelles et le traitement du langage naturel.

2.2 Métrique de performance

2.2.1 Explained Variance Score (EVS)

Le Explained Variance Score est une métrique utilisée pour évaluer la performance d'un modèle de régression. Elle mesure la proportion de la variance de la variable dépendante qui est expliquée par le modèle. Cette métrique est particulièrement utile pour comparer la performance de différents modèles de régression.

La formule de Explained Variance Score est donnée par :

$$\text{Explained Variance} = 1 - \frac{\text{Var}(y - \hat{y})}{\text{Var}(y)}$$

où : - y représente les valeurs réelles.

- \hat{y} représente les valeurs prédites par le modèle.

- Var désigne la variance.

Plus précisément :

- $\text{Var}(y - \hat{y})$ est la variance des erreurs de prédiction (ou des résidus).

- $\text{Var}(y)$ est la variance des valeurs réelles.

- Si Explained Variance Score est égal à 1, cela signifie que le modèle explique parfaitement la variance des données.

- Si Explained Variance Score est égal à 0, cela signifie que le modèle n'explique aucune variance des données, ce qui est équivalent à utiliser la moyenne des valeurs réelles comme prédiction.

- Un score négatif indique que le modèle est pire que la moyenne des valeurs réelles.

le Explained Variance Score est une mesure clé pour évaluer la performance des modèles de régression, en indiquant la proportion de la variance des données qui est expliquée par le modèle. Un score élevé indique un bon modèle de régression.

2.2.2 R-Squared (R²)

Le coefficient de détermination, communément appelé R^2 , est une métrique statistique utilisée pour évaluer la qualité d'ajustement d'un modèle de régression. Il indique la proportion de la variance de la variable dépendante qui est expliquée par la variable indépendante dans le modèle.

La formule de R^2 est donnée par :

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

où :

- SS_{res} est la somme des carrés des résidus (ou des erreurs) :

$$SS_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

avec y_i les valeurs réelles et \hat{y}_i les valeurs prédites par le modèle.

- SS_{tot} est la somme totale des carrés, qui mesure la variance totale des valeurs réelles :

$$SS_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2$$

où \bar{y} est la moyenne des valeurs réelles y_i .

- Si $R^2 = 1$, cela signifie que le modèle explique parfaitement toute la variance des données.
- Si $R^2 = 0$, cela signifie que le modèle n'explique aucune variance des données, ce qui est équivalent à utiliser la moyenne des valeurs réelles comme prédiction.
- Un R^2 négatif indique que le modèle est pire que la moyenne des valeurs réelles.

Le coefficient de détermination R^2 est une mesure clé pour évaluer la performance des modèles de régression, en indiquant la proportion de la variance des données qui est expliquée par le modèle. Un R^2 élevé indique un bon modèle de régression.

2.2.3 Mean Gamma Deviance (MGD)

Le Mean Gamma Deviance est une métrique utilisée pour évaluer la performance des modèles de régression lorsque les erreurs suivent une distribution Gamma. Cette métrique est particulièrement utile pour les données asymétriques ou hétéroscédastiques, courantes dans les modèles de régression Gamma.

La déviance Gamma moyenne est définie par :

$$\text{mean_gamma_deviance}(y, \hat{y}) = \frac{2}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{\hat{y}_i} - \log \left(\frac{y_i}{\hat{y}_i} \right) \right)$$

où :

- y_i sont les valeurs réelles.
- \hat{y}_i sont les valeurs prédites par le modèle.
- n est le nombre total d'observations.
- Un Mean Gamma Deviance de 0 indique une prédiction parfaite.
- Des valeurs positives de Mean Gamma Deviance indiquent des écarts entre les valeurs réelles et prédites, la magnitude de la valeur reflétant la gravité de ces écarts.

La déviance Gamma moyenne est une mesure clé pour évaluer les modèles de régression adaptés aux données avec des distributions Gamma, offrant une évaluation de la qualité de la prédiction en tenant compte de l'asymétrie et de l'hétéroscédasticité des données.

2.2.4 Mean Poisson Deviance (MPD)

Le Mean Poisson Deviance est une métrique utilisée pour évaluer la performance des modèles de régression lorsque les erreurs suivent une distribution de Poisson. Cette métrique est particulièrement adaptée aux données de comptage ou aux événements rares.

La déviance de Poisson moyenne est définie par :

$$\text{mean_poisson_deviance}(y, \hat{y}) = \frac{2}{n} \sum_{i=1}^n \left(y_i \log \left(\frac{y_i}{\hat{y}_i} \right) - (y_i - \hat{y}_i) \right)$$

où :

- y_i sont les valeurs réelles.
- \hat{y}_i sont les valeurs prédites par le modèle.
- n est le nombre total d'observations.
- Un Mean Poisson Deviance de 0 indique une prédiction parfaite.
- Des valeurs positives de Mean Poisson Deviance indiquent des écarts entre les valeurs réelles et prédites, la magnitude de la valeur reflétant la gravité de ces écarts.

La déviance de Poisson moyenne est une mesure clé pour évaluer les modèles de régression adaptés aux données de comptage, offrant une évaluation de la qualité de la prédiction en tenant compte de la distribution de Poisson des erreurs.

2.3 Résultats et analyse

2.3.1 LinearRegression

LinearRegression montre des résultats plus équilibrés entre l'entraînement et le test. Pendant l'entraînement, l'EVS et le R2 sont à 0.92, tandis que les valeurs de MGD et MPD sont respectivement 0.00 et 5296.01, soulignant des erreurs très faibles.

En test, ce modèle conserve une performance supérieure par rapport aux autres modèles, avec un EVS de 0.78 et un R2 de 0.43. Les valeurs de MGD et MPD sont aussi les plus faibles en test, à savoir 0.02 et 58919.44, ce qui indique une bonne capacité de généralisation.

2.3.2 XGBRegressor

XGBRegressor se distingue par des performances remarquables pendant l'entraînement avec un EVS (Explained Variance Score) et un R2 tous deux à 0.99, ce qui indique une excellente capacité à expliquer la variance des données.

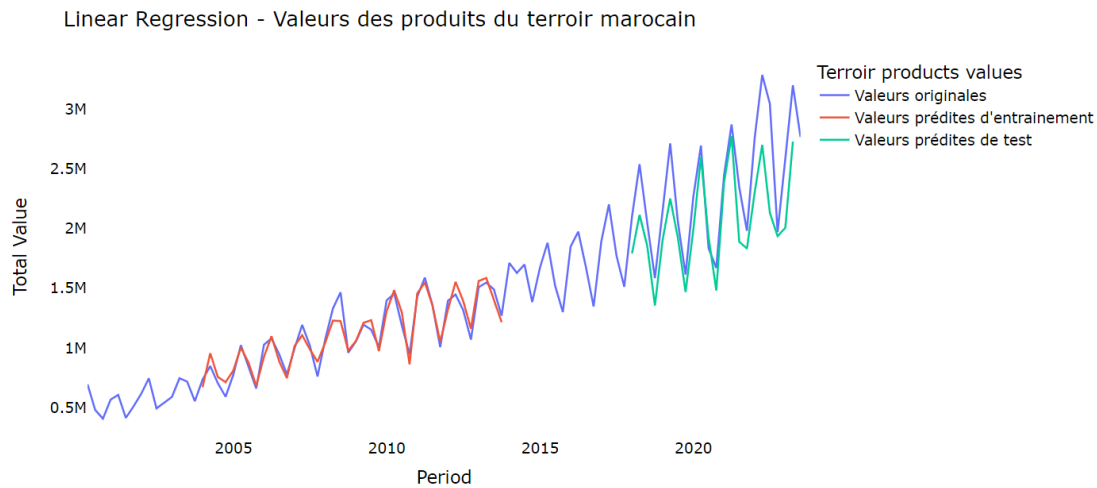


FIGURE 2.1 – Comparaison entre les valeurs originales et les valeurs prédites par LinearRegression

Cependant, ses performances sur le jeu de test chutent drastiquement, avec un EVS à 0.00 et un R^2 à -2.94, indiquant une surapprentissage. Le MGD (Mean Gradient Deviation) et le MPD (Mean Prediction Deviation) sont également très élevés en test, avec respectivement 0.27 et 502634.27, suggérant des erreurs de prédiction significatives.

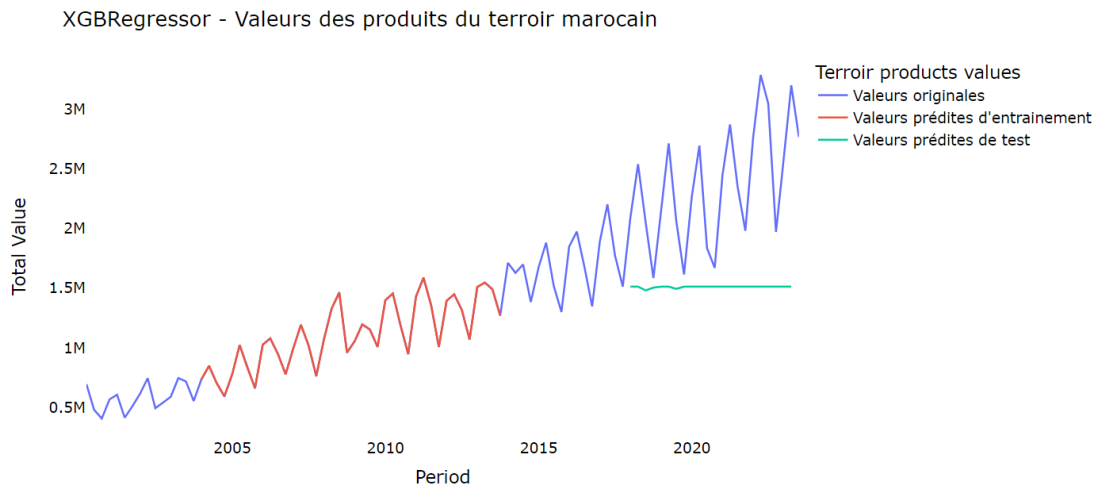


FIGURE 2.2 – Comparaison entre les valeurs originales et les valeurs prédites par XGBRegressor

2.3.3 LSTM

LSTM affiche des performances relativement modestes, avec un EVS et un R^2 de 0.59 lors de l'entraînement. Le modèle excelle néanmoins en termes de MPD pendant l'entraînement avec une valeur de 0.08, mais présente des difficultés importantes en phase de test.

Le EVS tombe à 0.15 et le R^2 à -0.64, signifiant une faible capacité explicative et prédictive. Les valeurs de MGD et MPD en test (26921.81 et 180126.53 respectivement) révèlent de larges erreurs, illustrant une mauvaise généralisation du modèle.

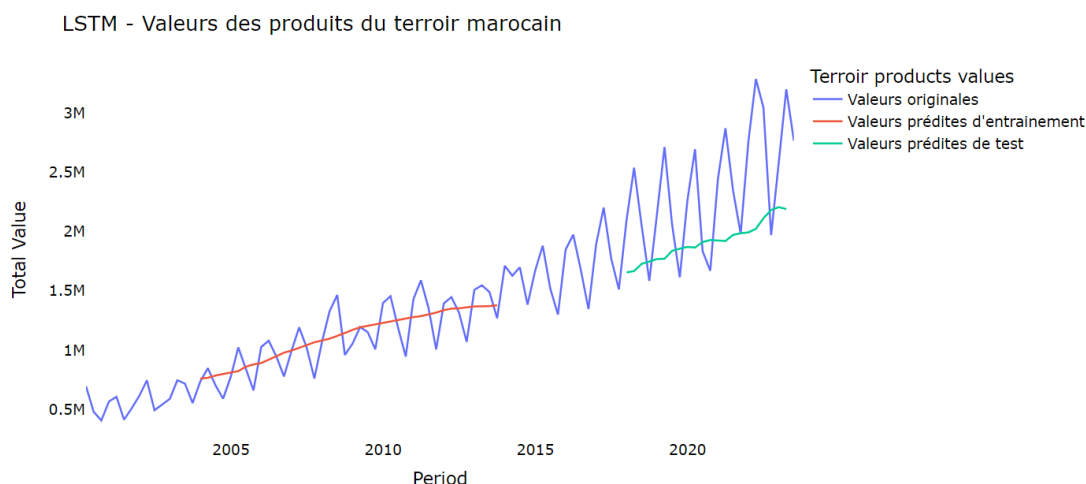


FIGURE 2.3 – Comparaison entre les valeurs originales et les valeurs prédites par LSTM

2.3.4 Comparaison des modèles

Le tableau présente une comparaison des performances de trois modèles différents : XGBRegressor, LinearRegression, et LSTM, pour les phases d'entraînement et de test.

TABLE 2.1 – Model Performance Comparison

| Model | Train | | | | Test | | | |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-----------------|
| | EVS | R2 | MGD | MPD | EVS | R2 | MGD | MPD |
| XGBRegressor | 0.99 | 0.99 | 2.23 | 1.95 | 0.00 | -2.94 | 0.27 | 502634.27 |
| LinearRegression | 0.92 | 0.92 | 0.00 | 5296.01 | 0.78 | 0.43 | 0.02 | 58919.44 |
| LSTM | 0.59 | 0.59 | 0.02 | 0.08 | 0.15 | -0.64 | 26921.81 | 180126.53 |

LinearRegression offre la meilleure performance globale avec des erreurs minimales et une bonne généralisation, tandis que XGBRegressor souffre de surapprentissage et LSTM ne parvient pas à bien généraliser les données de test.

2.4 Prédiction du modèle

Les valeurs prédites pour les prochaines 30 séries chronologiques trimestrielles, allant de 2023-Q3 à 2030-Q4, basées sur le modèle de régression linéaire des exportations des produits de terroir marocains en USD, montrent des tendances intéressantes et significatives. Le premier graphique compare les valeurs réelles des 15 dernières périodes avec les valeurs prédites pour les 30 prochaines périodes. Les valeurs réelles oscillent entre environ 1,6 million USD et 3,2 millions USD, affichant une forte variabilité. Les valeurs prédites pour les 30 prochaines périodes semblent suivre un schéma similaire, avec des oscillations régulières, suggérant une saisonnalité forte et récurrente dans les données des exportations.

Le second graphique présente la série chronologique entière, incluant les valeurs prédites. On observe une tendance générale à la hausse continue sur toute la période, avec des augmentations progressives chaque trimestre. La saisonnalité est également évidente, avec des pics et

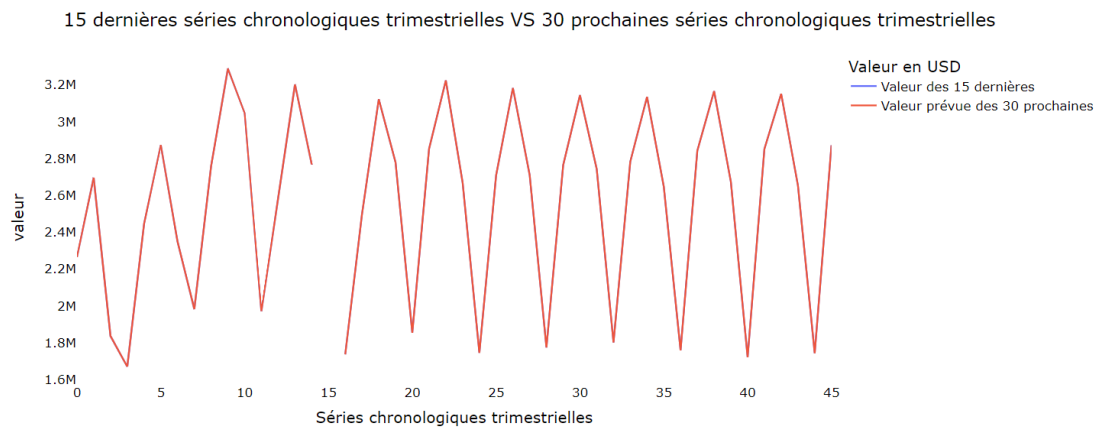


FIGURE 2.4 – 15 dernières séries chronologiques trimestrielles VS 30 prochaines séries chronologiques trimestrielles

des creux réguliers dans les données. À partir de 2023-Q3, les valeurs continuent de montrer une augmentation générale, avec des variations saisonnières jusqu'à atteindre environ 3 millions USD à la fin de la période prédite, soit en 2030-Q4.

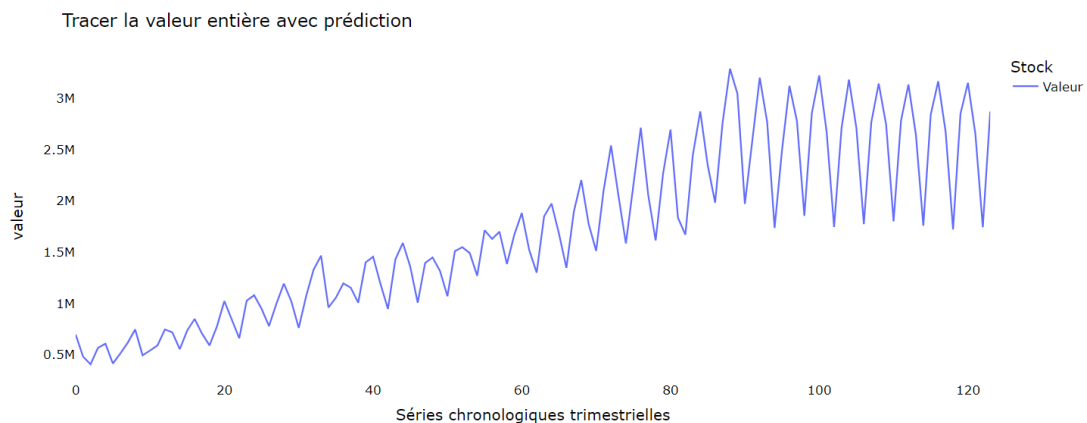


FIGURE 2.5 – Tracer de la valeur entière avec prédiction

Pour les détails spécifiques des valeurs prédites de 2023-Q3 à 2030-Q4, les prévisions commencent autour de 2,5 millions USD pour 2023-Q3 et montrent des oscillations trimestrielles avec une légère tendance à la hausse, atteignant environ 2,8 millions USD à la fin de 2024. De 2025-Q1 à 2026-Q4, la tendance à la hausse se poursuit avec des valeurs oscillant entre 2,6 millions USD et 3 millions USD. Pour la période de 2027-Q1 à 2028-Q4, les valeurs continuent d'augmenter légèrement, oscillant entre 2,7 millions USD et 3,1 millions USD. Enfin, de 2029-Q1 à 2030-Q4, les prédictions montrent des valeurs atteignant jusqu'à 3,2 millions USD, avec des oscillations saisonnières toujours présentes.

Les prédictions indiquent que les exportations des produits de terroir marocains continueront d'augmenter régulièrement jusqu'en 2030, tout en maintenant un schéma saisonnier marqué. Les valeurs prévues pour les trimestres futurs oscillent entre 2,5 millions USD et 3,2 millions USD, reflétant à la fois une croissance globale et des variations trimestrielles typiques.

Chapitre 3

Stratégies de Développement et de Commercialisation

En se basant sur les prévisions de croissance des exportations des produits de terroir marocains jusqu'en 2030, il est important de mettre en œuvre des stratégies de développement et de commercialisation ciblées. Les valeurs prévues oscillant entre 2,5 millions USD et 3,2 millions USD indiquent une tendance positive avec une saisonnalité marquée, offrant ainsi des opportunités de maximiser les revenus.

Pour capitaliser sur cette croissance, il est essentiel d'améliorer les techniques de production afin de garantir une qualité constante et élevée, ce qui renforcera la confiance des consommateurs et stimulera la demande. Diversifier les marchés cibles, en explorant de nouvelles régions et en consolidant les partenariats commerciaux existants, contribuera à stabiliser les revenus et à réduire la dépendance à un marché unique. L'adoption de pratiques durables et le respect des normes internationales valoriseront les produits sur le marché global.

De plus, des campagnes de marketing efficaces, mettant en avant l'authenticité et la qualité des produits de terroir marocains, seront déterminantes pour attirer de nouveaux clients et fidéliser ceux existants. En combinant ces efforts, il sera possible de favoriser une croissance soutenue et de maximiser les exportations des produits de terroir marocains dans les années à venir.

3.1 Synthèse des résultats et perspectives futures

Les prévisions de croissance des exportations des produits de terroir marocains de 2023-Q3 à 2030-Q4 montrent une tendance à la hausse continue avec des valeurs oscillant entre 2,5 millions USD et 3,2 millions USD. Cette analyse est basée sur un modèle de régression linéaire appliqué aux données historiques, révélant une forte saisonnalité et une augmentation progressive des valeurs trimestrielles. Les oscillations régulières suggèrent que les exportations des produits de terroir marocains sont influencées par des facteurs saisonniers récurrents tout en bénéficiant d'une croissance globale positive.

Il est crucial de transformer ces prévisions en stratégies concrètes pour maximiser les avantages et atténuer les risques potentiels. Une première étape essentielle est l'optimisation de la production. Il est impératif d'améliorer les techniques de production pour garantir une qualité

élevée et constante, ce qui peut être réalisé par l'adoption de technologies agricoles avancées et de pratiques durables. Investir dans la formation des agriculteurs et des producteurs pour adopter les meilleures pratiques agricoles et les technologies modernes est également fondamental.

La diversification des marchés est une autre stratégie clé. Explorer de nouveaux marchés régionaux et internationaux peut réduire la dépendance à un nombre limité de marchés et permettre d'identifier les marchés émergents où les produits de terroir marocains peuvent avoir un avantage compétitif. Renforcer les partenariats avec des distributeurs et des détaillants internationaux aidera à élargir la portée et la disponibilité des produits.

Le marketing et le branding jouent un rôle crucial dans cette stratégie. Développer des campagnes de marketing efficaces pour promouvoir l'authenticité, la qualité et l'histoire des produits de terroir marocains est essentiel. Utiliser les médias sociaux et les plateformes en ligne pour atteindre un public mondial permettra de maximiser l'impact. Obtenir des certifications internationales et des labels de qualité renforcera la crédibilité et la confiance des consommateurs.

L'innovation et l'adaptation doivent également être au cœur de la stratégie. Investir dans la recherche et le développement pour créer de nouveaux produits et variantes qui répondent aux tendances du marché et aux préférences des consommateurs est crucial. De plus, être prêt à ajuster les stratégies en fonction des changements du marché et des nouvelles tendances, en utilisant l'analyse des données en temps réel pour prendre des décisions informées, est une pratique à adopter.

La soutenabilité et la responsabilité sociale doivent être intégrées dans la stratégie globale. Adopter des pratiques agricoles durables pour minimiser l'impact environnemental et répondre à la demande croissante des consommateurs pour des produits respectueux de l'environnement est essentiel. Engager les communautés locales et s'assurer que le développement des exportations bénéficie également aux producteurs locaux contribuera au développement économique régional.