

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITE BADJI MOKHTAR - ANNABA
BADJI MOKHTAR – ANNABA UNIVERSITY



جامعة بادجي مختار – عنابة

Faculté: TECHNOLOGIE

Département: ELECTRONIQUE

Domaine : SCIENCE ET TECHNOLOGIE

Filière : TELECOMMUNICATIONS

Spécialisation: SYSTÈME DE TÉLÉCOMMUNICATIONS

Mémoire

Présenté en vue de l'obtention du diplôme de Master

Thème

Prédiction des rendements agricoles basée sur le Deep Learning

Présenté Par : *HEMICI ABDELMOIZ Et MERABET ABDERRAOUF*

Encadrant : *BOULMAIZ AMIRA*

M.C.A

Université Badji Mokhtar Annaba

Jury de Soutenance :

S.LAFIFI	Professeur	Université Badji Mokhtar Annaba	Président
N.KOUADRIA	Professeur	Université Badji Mokhtar Annaba	Examineur
A. BOULMAIZ	MCA	Université Badji Mokhtar Annaba	Encadrant

Année universitaire : 2023-2024

Remerciement

Nous exprimons notre profonde gratitude à Dieu pour son précieux guide tout au long de ce voyage. Sa bénédiction a été la raison de la rédaction de ce mémoire dans ces circonstances.

Nous tenons également à exprimer notre sincère gratitude à notre superviseur, Mlle Amira Boulmaiz, pour son précieux soutien tout au long de cette étude. Ses conseils et son soutien ont été inestimables, et nous sommes reconnaissants pour son engagement et sa contribution à la réussite de ce travail.

Aux membres du jury, Pr. N.Kouadria et Pr. S.Lafifi, nous exprimons notre gratitude pour votre dévouement à évaluer notre mémoire. Votre expérience a été un élément crucial pour vérifier la validité de notre travail, et nous apprécions votre participation.

Enfin, nous adressons à nos familles, nos amis et tous ceux qui nous ont soutenus, notre sincère gratitude pour votre soutien et vos précieux conseils. Votre présence a été essentielle pour la réalisation de ce projet.

Merci à vous tous.

Dédicaces

Je dédie cette thèse à ma famille, en particulier à mes parents qui ont été mon pilier en me fournissant l'environnement idéal et le soutien matériel et moral pour réaliser cette thèse.

De plus, je tiens à remercier notre superviseur, Mlle Boulmaiz Amira, pour sa supervision et ses conseils tout au long de cette période.

Je remercie également tous mes amis qui sont considérés comme des frères pour moi, pour leurs encouragements et leur aide.

Un merci spécial à tous ceux qui ont été à mes côtés pendant cette période, je vous suis vraiment redevable.

Hemici Abdelmoiz

Dédicaces

je dédie ce projet

À mes chers parents,

Votre amour et votre soutien indéfectible m'ont permis de réaliser ce rêve. Vous avez toujours cru en moi, même dans les moments difficiles. Ce mémoire vous est dédié, en témoignage de ma profonde gratitude.

À mes chers frères,

Merci pour votre présence constante et votre soutien moral tout au long de ce parcours. Vos encouragements ont été précieux pour moi.

À mes chers camarades de promotion,

Ces années passées ensemble ont été enrichissantes et inoubliables. Merci pour les moments de partage, de rires et de soutien mutuel. Ce mémoire est un symbole de notre collaboration et de notre amitié.

À l'Université Badji Mokhtar d'Annaba,

Merci pour la qualité de l'enseignement et les opportunités qui m'ont été offertes. Ce mémoire est un témoignage de mon engagement et de mon respect pour cette institution.

Merabet Abderraouf

Abstract

Agriculture constitutes a fundamental and vital sector in all countries, due to its direct impact on the social and economic status of the country. Climatic factors play a decisive role in agricultural production. The purpose of this study is to present an effective and robust model for predicting agricultural yield in the city of Annaba using Recurrent Neural Networks (RNN) and its types such as LSTM, BiLSTM, GRU, with the aim of reducing losses resulting from climate fluctuations.

The results of our experiments demonstrated the optimal performance of the BiLSTM model with 3 layers using the RMSprop optimizer with an RMSE equal to 2.10.

Keywords: agricultural, Annaba, RNN, LSTM, GRU, BiLSTM, RMSE, RMSprop.

Résumé

L'agriculture constitue un secteur fondamental et vital dans tous les pays, en raison de son impact direct sur la situation sociale et économique du pays. Les facteurs climatiques jouent un rôle décisif dans la production agricole. Le but de cette étude est de présenter un modèle efficace et robuste pour prédire le rendement agricole dans la ville d'Annaba en utilisant les réseaux neuronaux récurrents (RNN) et ses types tels que LSTM, BiLSTM, GRU, dans le but de réduire les pertes résultant des fluctuations climatiques. Les résultats de nos expérimentations ont démontré la performance optimale du modèle BiLSTM avec 3 couches en utilisant l'optimiseur RMSprop avec un RMSE égal 2.10.

Mots clés : Agriculture, Annaba, RNN, LSTM, GRU, BiLSTM, RMSE, RMSprop.

ملخص

الزراعة تشكل قطاعا أساسيا وحيويا في كل الدول، نظرا لأثرها المباشر على الحالة الاجتماعية والاقتصادية للبلاد. العوامل المناخية تلعب دورا حاسما في الإنتاج الزراعي. الغرض من هذه الدراسة هو تقديم نموذج فعال وقوي لتوقع الغلة الزراعية في مدينة عنابة باستخدام الشبكات العصبية المتكررة (RNN) وأنواعها مثل LSTM, BiLSTM, GRU، وذلك بهدف التقليل من الخسائر المترتبة على التقلبات المناخية. أظهرت نتائج تجاربنا الأداء الأمثل لنموذج BiLSTM مع 3 طبقات باستخدام محسن RMSprop مع RMSE يساوي 2.10.

الكلمات المفتاحية: الزراعة، عنابة، RNN، BiLSTM، GRU، LSTM، RMSE، RMSprop.

Liste des Figures

Figure I.1. L'emplacement stratégique d'Annaba en Algérie	6
Figure I.2. Superficies allouées à la culture du Blé en Hectare au cours des 12 dernières années	8
Figure I.3. Production du blé en quintal au cours des 12 dernières années	9
Figure I.4. Rendements agricoles du blé au cours des 12 dernières années	10
Figure I.5. Superficies allouées à la culture des olives en Hectare au cours des 10 dernières années	10
Figure I.6. Production des olives en quintal au cours des 10 dernières années	11
Figure I.7. Rendements agricoles des olives au cours des 10 dernières années	11
Figure II.1. Les branches de l'intelligence Artificielle.....	21
Figure II.2. Réseau neuronal récurrent	22
Figure II.3. Mémoire à long terme (LSTM)	23
Figure II.4. Gated Recurrent Unit (GRU).....	25
Figure II.5. Couche LSTM bidirectionnelle	26
Figure II.6. Comparaison entre le réseau neuronal normal (à gauche) et réseau avec dropout (à droite).....	27
Figure II.7. La couche Batch Normalization	28
Figure II.8. La couche Dense.....	29
Figure III.1. Schéma d'ensemble des étapes du système de prédiction	39
Figure III.2. Carte de la région d'Annaba	41
Figure III.3. Ensemble de données climatiques de Annaba.....	43
Figure III.4. Ensemble de données agricoles de Annaba.....	43
Figure III.5. Modèle de perte pendant l'apprentissage de modèle RNN (Adam)	46
Figure III.6. Résultat de prédiction des rendements du blé basé sur le modèle RNN utilisant l'optimiseur Adam	46
Figure III.7. Modèle de perte pendant l'apprentissage de modèle RNN (RMSprop)	48

Figure III.8. Résultat de prédiction des rendements du blé basé sur le modèle RNN utilisant l'optimiseur RMSprop	48
Figure III.9. Modèle de perte pendant l'apprentissage de modèle LSTM (Adam).....	50
Figure III.10. Résultat de prédiction des rendements du blé basé sur le modèle LSTM utilisant l'optimiseur Adam	50
Figure III.11. Modèle de perte pendant l'apprentissage de modèle LSTM (RMSprop)	52
Figure III.12. Résultat de prédiction des rendements du blé basé sur le modèle LSTM utilisant l'optimiseur RMSprop	52
Figure III.13. Modèle de perte pendant l'apprentissage de modèle GRU (Adam)	54
Figure III.14. Résultat de prédiction des rendements du blé basé sur le modèle GRU utilisant l'optimiseur Adam	54
Figure III.15. Modèle de perte pendant l'apprentissage de modèle GRU (RMSprop)	56
Figure III.16. Résultat de prédiction des rendements du blé basé sur le modèle GRU utilisant l'optimiseur RMSprop	56
Figure III.17. Modèle de perte pendant l'apprentissage de modèle BiLSTM (Adam)	58
Figure III.18. Résultat de prédiction des rendements du blé basé sur le modèle BiLSTM utilisant l'optimiseur Adam	58
Figure III.19. Modèle de perte pendant l'apprentissage de modèle BiLSTM (RMSprop)	60
Figure III.20. Résultat de prédiction des rendements du blé basé sur le modèle BiLSTM utilisant l'optimiseur RMSprop	60

Liste des Tableaux

Tableau III.1. Détails de localisation.....	41
Tableau III.2. Détails sur la base de données	42
Tableau III.3. Résultats du modèle RNN utilisant l’optimiseur Adam en terme de RMSE	45
Tableau III.4. Résultats du modèle RNN utilisant l’optimiseur RMSprop en terme de RMSE..	47
Tableau III.5. Résultats du modèle LSTM utilisant l’optimiseur Adam en terme de RMSE.....	49
Tableau III.6. Résultats du modèle LSTM utilisant l’optimiseur RMSprop en terme de RMSE	51
Tableau III.7. Résultats du modèle GRU utilisant l’optimiseur Adam en terme de RMSE	53
Tableau III.8. Résultats du modèle GRU utilisant l’optimiseur RMSprop en terme de RMSE..	55
Tableau III.9. Résultats du modèle BiLSTM utilisant l’optimiseur Adam en terme de RMSE..	57
Tableau III.10. Résultats du modèle BiLSTM utilisant l’optimiseur RMSprop en terme de RMSE	59
Tableau III.11. La valeur RMSE pour les meilleurs quatre résultats obtenus à partir des quatre modèles	61
Tableau A.1. Architecture de BiLSTM utilisée pour prédire le rendement Blé.....	65

Liste des Abréviations

Abréviations	Description
DGAA	Direction Générale de l'Agriculture d'Annaba
UAV	Unmanned Aerial Vehicle
AI	Intelligence Artificielle
NDVI	l'indice de végétation différentielle normalisé
EVI	l'indice de végétation amélioré
ML	Machine Learning
ANN	Artificial Neural Network
DL	Deep Learning
CNN	Convolutional Neural Networks
GAN	Generative Adversarial Networks
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
BiLSTM	Bidirectionnel LSTM
GRU	Gated Recurrent Unit
Adam	Adaptive Moment Estimation
RMSprop	Root Mean Squared Propagation
ReLU	Rectified Linear Unit
MSE	Mean Squared Error
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
T	Température moyenne annuelle
TM	Température maximale moyenne annuelle
Tm	Température minimale moyenne annuelle
PP	Précipitation totale annuelle de pluie
RA	Nombre total de jours de pluie durant l'année
V	La vitesse moyenne annuelle du vent
SN	Nombre total de jours de neige durant l'année

Table des matières

<i>Remerciement</i>	II
<i>Dédicaces</i>	III
<i>Dédicaces</i>	IV
Résumé	VI
Abstract	V
ملخص	VII
Liste des Figures	VIII
Liste des Tableaux	X
Liste des Abréviations	XI
Table des matières	XII
Introduction Générale	2
CHAPITRE I : Etat de l'art sur la Prédiction des Rendements agricoles à Annaba	4
I.1 Introduction	5
I.2 Présentation Générale de la Région d'Annaba	5
I.3 Importance Économique et Sociale de l'Agriculture	7
I.4 Certaines des cultures fournies par la région	8
I.5 Problématique	12
<i>I.5.1 Pourquoi la Prédiction des Rendements Agricoles ?</i>	12
<i>I.5.2 Le problème de prédiction des rendements agricoles</i>	13
I.6 État de l'art	13
<i>I.6.1 Revue des Travaux Antérieurs</i>	13
<i>I.6.2 Méthodes Traditionnelles et Techniques Modernes</i>	14
I.7 Limitations d'étude	15

I.7.1 Limites potentielles du projet.....	15
I.8 Conclusion	16
CHAPITRE II : Introduction aux Réseaux neuronaux récurrents.....	17
II.1 Introduction	18
II.2 Aperçu sur les différentes branches de l'intelligence artificielle	18
II.2.1 Intelligence artificielle (AI)	18
II.2.2 Machine Learning (ML)	19
II.2.3 Réseaux de neurones artificiels (ANN).....	19
II.2.4 L'apprentissage profond (DL)	20
II.3 Les réseaux de neurones récurrents (RNN).....	21
II.3.1 Structure du réseau neuronal récurrent LSTM	23
II.3.2 Structure du réseau neuronal récurrent (GRU)	24
II.3.3 Structure de Bidirectionnel LSTM (BiLSTM)	26
II.4 Amélioration de l'apprentissage.....	27
II.4.1 La couche Dropout	27
II.4.2 La couche Batch Normalization	28
II.4.3 La couche Dense	29
II.5 Paramètres d'apprentissage en profondeur.....	30
II.5.1 Adam (Adaptive Moment Estimation)	30
II.5.2 RMSprop (Root Mean Square Propagation)	30
II.6 Hyperparamètres d'un réseau de neurones	31
II.6.1 Unités par couche (Number of Units in Each Layer).....	31
II.6.2 Nombre de couches cachées (Number of Hidden Layers).....	32
II.6.3 Nombre d'époques (Number of Epochs)	32
II.6.4 Taux d'apprentissage (Learning Rate).....	32

II.6.5 Fonctions d'activation (Activation Fonctions)	32
II.7 Métriques d'évaluation	34
II.7.1 Mean squared Error (MSE)	34
II.7.2 Mean Absolute Error (MAE)	35
II.7.3 Root Mean Square Error (RMSE)	35
II.8 Conclusion	36
CHAPITRE III : Système proposé pour la prédiction des rendements agricoles	37
III.1 Introduction	38
III.2 Les principales étapes du système de prédiction proposé	38
III.3 Outils et matériels utilisés	40
III.4 Collecte de données	40
III.5 Entraînement du modèle	44
III.6 Résultats et discussion	44
a. Résultats du modèle RNN	45
b. Résultats du modèle LSTM	49
c. Résultats du modèle GRU	53
d. Résultats du modèle BiLSTM	57
III.7 Conclusion	62
Conclusion Générale	63
Annexe	64
Bibliographie	66
Webographie	70

Introduction Générale

L'agriculture joue un rôle crucial dans l'économie mondiale, assurant la sécurité alimentaire et générant des revenus pour des millions de personnes. Cependant, la prévision des rendements agricoles reste un défi majeur en raison de la complexité des facteurs influençant la production, tels que les conditions climatiques, les pratiques agricoles et les caractéristiques du sol. Ce mémoire se concentre sur l'application des réseaux neuronaux récurrents (RNN) pour améliorer la précision des prédictions des rendements agricoles, en particulier dans la région d'Annaba.

La région d'Annaba, en Algérie, est une zone agricole importante dont l'économie dépend en grande partie de la production agricole. Malgré son potentiel, l'agriculture à Annaba fait face à de nombreux défis, notamment les variations climatiques, les techniques agricoles traditionnelles et les contraintes socio-économiques. Ces défis compliquent la prévision des rendements agricoles, rendant nécessaire l'utilisation de techniques avancées pour améliorer la précision des prédictions.

La prévision précise des rendements agricoles est essentielle pour les agriculteurs, les décideurs et les planificateurs économiques. Elle permet de mieux gérer les ressources, de planifier les opérations agricoles et de prendre des décisions éclairées pour assurer la sécurité alimentaire. En utilisant des techniques avancées d'apprentissage automatique, ce mémoire vise à contribuer à l'amélioration des prédictions des rendements agricoles, apportant ainsi une valeur ajoutée significative à la région d'Annaba et potentiellement à d'autres régions agricoles similaires.

Ce mémoire a plusieurs objectifs qui peuvent être résumer dans les points suivants:

Fournir un aperçu complet de la région d'Annaba et de son importance agricole, en identifiant les principaux défis auxquels elle est confrontée.

Explorer l'application des réseaux neuronaux récurrents et de leurs variantes, comme les LSTM (Long Short-Term Memory) et les GRU (Gated Recurrent Unit), dans la prévision des rendements agricoles.

Proposer un système de prédiction des rendements agricoles en compilant un ensemble de données robustes pour la région d'Annaba et en analysant l'impact des hyperparamètres sur les performances des modèles.

Ce mémoire est structuré en trois chapitres principaux :

Chapitre 1 : Ce chapitre fournit un aperçu détaillé de la région d'Annaba, en soulignant son importance agricole et les principaux défis rencontrés. Nous discutons également de l'importance de la prévision des rendements agricoles et des différentes méthodologies utilisées par les chercheurs pour aborder ce problème.

Chapitre 2 : Ce chapitre explore en profondeur les réseaux neuronaux récurrents et leurs variantes, telles que les LSTM, les BiLSTM (LSTM bidirectionnelle) et les GRU. Nous examinons comment ces modèles peuvent capturer les dépendances temporelles dans les données agricoles et discutons des hyperparamètres clés pour leur construction.

Chapitre 3 : Ce chapitre propose un système de prédiction des rendements agricoles pour Annaba, détaillant la collecte et le prétraitement des données, ainsi que la comparaison des performances des différents modèles RNN. Nous analysons également l'impact des hyperparamètres sur l'efficacité des modèles.

En conclusion, ce mémoire vise à combler les lacunes dans les méthodes traditionnelles de prévision des rendements agricoles en utilisant des approches modernes basées sur les réseaux neuronaux récurrents, offrant ainsi des solutions potentielles pour améliorer la précision des prévisions agricoles dans la région d'Annaba.

CHAPITRE I : Etat de l'art sur la Prédiction des Rendements agricoles à Annaba

I.1 Introduction

Ce chapitre fournit un aperçu complet de la région d'Annaba, en soulignant son importance agricole. Nous nous penchons sur l'importance économique et sociale de l'agriculture, en soulignant les principales cultures produites dans la région. Cependant, l'agriculture à Annaba fait face à plusieurs défis, dont nous discuterons dans la section « Problématique ».

L'essentiel de ce chapitre réside dans l'exploration de la prévision du rendement agricole. Nous discuterons des raisons pour lesquelles la prévision des rendements agricoles est cruciale et des complexités inhérentes à ce processus. Nos objectifs de recherche, tant généraux que spécifiques, visent à aborder ces complexités et à contribuer à l'ensemble des connaissances existantes.

Un examen des travaux antérieurs dans ce domaine nous aidera à comprendre les méthodologies utilisées et les résultats obtenus par d'autres chercheurs. Nous comparerons et comparerons également les méthodes traditionnelles de prévision du rendement avec les techniques modernes, en fournissant un compte rendu détaillé des deux.

Cependant, notre étude n'est pas sans limites. Nous reconnaitrons ces limites et définirons les limites de notre projet. Enfin, nous identifierons les domaines que notre étude ne couvrira pas ou seulement brièvement.

Ce chapitre sert de guide à travers le paysage complexe de la prévision des rendements agricoles dans la région d'Annaba, préparant le terrain pour les discussions détaillées dans les chapitres suivants.

I.2 Présentation Générale de la Région d'Annaba

Annaba est une ville située dans le nord-est de l'Algérie, sur la rive sud du bassin algéro-provençal [1]. Elle est distante de 600 km de la capitale Alger et 106 km la séparent de la frontière tunisienne [1]. La ville s'élève au fond d'une baie ouverte à l'est sur le golfe d'Annaba [1]. La wilaya d'Annaba est reliée aux grandes villes d'Algérie et à la Tunisie distante de 100 km par un réseau routier et ferroviaire dense [2]. Skikda, Constantine, Alger, Tunis et Guelma sont ainsi directement desservies à partir de la wilaya d'Annaba [2].

Annaba est connue pour être une ville accueillante, hospitalière et touristique très appréciée par les visiteurs qui lui ont donnée le nom affectif de « la coquette » [3]. Elle possède un centre-ville très animé, un front de mer et une corniche qui a gardé toute son authenticité.[3]

En ce qui concerne l'agriculture, Annaba dispose de grandes potentialités agricoles [4]. Elle possède une plaine aux terres fertiles et une pluviométrie favorable, ce qui la rend propice à diverses activités agricoles [4]. Cependant, malgré ces atouts, la production agricole d'Annaba est relativement faible par rapport à son potentiel [4]. En moyenne, la wilaya ne produit annuellement que 26 milliards de dinars, alors que d'autres wilayas moins nanties en potentialités enregistrent chacune 250 milliards de DA/an [4].

Pour remédier à cette situation, des efforts sont déployés pour développer les cultures et toutes les filières agricoles [4]. Par exemple, le Salon International de l'Investissement Agricole et de la Technologie (Agritech expo 2023) a été organisé à Annaba pour promouvoir l'investissement intelligent pour un développement durable [4]. Ce salon a permis de présenter des technologies de pointe et des acquis de la recherche dans les domaines agricole, agro-alimentaire, de pêche et des services qui leur sont liés [4].



Figure I.1. L'emplacement stratégique d'Annaba en Algérie

I.3 Importance Économique et Sociale de l'Agriculture

L'agriculture joue un rôle central dans les sociétés du monde entier, servant de pierre angulaire pour la sécurité alimentaire, la création d'emplois et la résilience économique. En Algérie, l'industrie agricole a été un élément clé de la structure économique et sociale du pays.

I.3.1. Importance économique

Le secteur agricole en Algérie renforce l'expansion économique du pays en offrant des opportunités de revenus pour les agriculteurs, en soutenant les économies rurales et en générant des profits grâce au commerce des produits agricoles. L'évolution des coopératives agricoles en Algérie, dont Tabacoop et Cotocoop, a été essentielle pour moderniser les méthodes agricoles, augmenter la productivité et faciliter l'accès aux marchés. L'application de techniques agricoles spécialisées, en particulier dans la culture du tabac et du coton, a amélioré de manière significative la qualité des produits agricoles destinés à l'exportation.

I.3.2. Importance sociale

L'agriculture est un secteur d'emploi majeur en Algérie, en particulier dans les régions rurales où une grande partie de la population dépend de l'agriculture pour ses moyens de subsistance. Des initiatives telles que les écoles agricoles et les centres de formation ont été cruciales pour éduquer la population locale sur les pratiques agricoles, favorisant ainsi le développement des compétences et l'autonomisation. La formation de coopératives agricoles et de syndicats a non seulement amélioré les méthodes agricoles, mais a également cultivé un sentiment de camaraderie et de coopération entre les agriculteurs.

I.3.3. Défis et opportunités

Malgré son importance, le secteur agricole algérien est confronté à plusieurs défis, dont la pénurie d'eau, la dégradation des sols et l'accès limité aux technologies avancées. Cependant, le secteur offre également des opportunités de croissance économique supplémentaire grâce à des pratiques durables, à l'innovation technologique et à la production de produits agricoles à valeur ajoutée. Enfin, l'industrie agricole en Algérie a une importance économique et sociale considérable. Les efforts pour soutenir et améliorer le secteur agricole peuvent contribuer de manière significative à la prospérité et au bien-être général de la nation. [1].

I.4 Certaines des cultures fournies par la région

Annaba, une cité algérienne, est célèbre pour son exceptionnelle productivité agricole. Elle couvre une superficie totale de 133 972 hectares consacrée à l'agriculture, qui comprend 43 850 hectares de terres arables, 75 424 hectares de forêts et 14 698 hectares réservés aux pâturages et aux terres non productives. De ce fait, la wilaya d'Annaba est considérée comme un centre significatif pour le développement agricole et une culture remarquable.

Le blé a été introduit dans les pratiques agricoles locales dans la région de Bône à Annaba. Où il a été cultivé sur les sols fertiles de la plaine d'Annaba. La culture industrielle du blé a été promue par l'Union Agricole de l'Est qui a converti des terres fertiles en terres arables. Dans le processus de production et de vente du blé, les coopératives agricoles ont joué un rôle vital en normalisant les méthodes agricoles et en facilitant les exportations vers les marchés internationaux.[1-2]

Selon la direction générale de l'agriculture de la ville d'Annaba, les superficies allouées à la culture du blé ont été couvertes de 2012 à 2023 et se répartissaient comme suit :

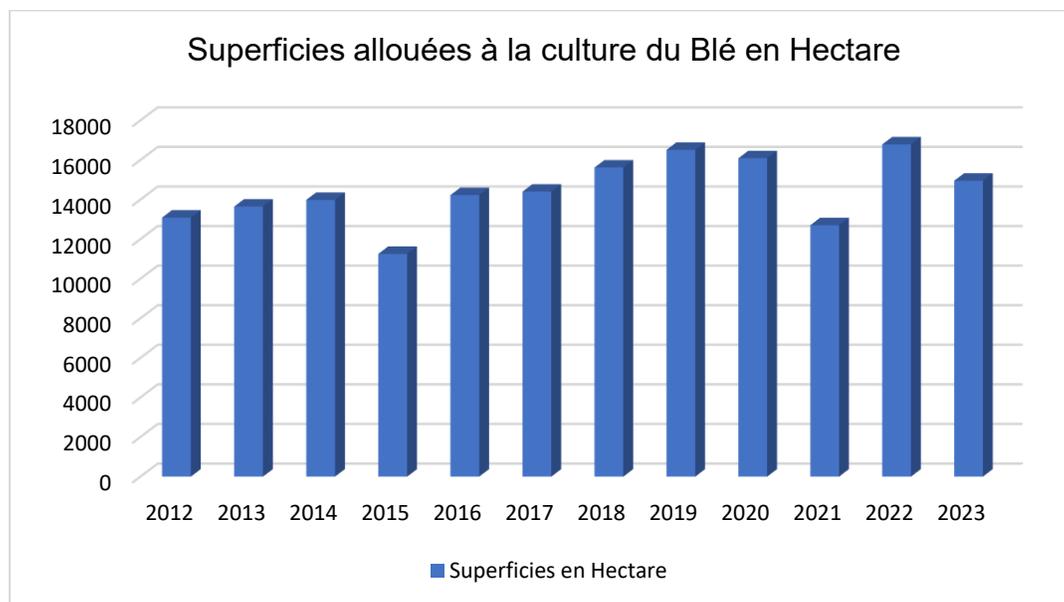


Figure I.2. Superficies allouées à la culture du Blé en Hectare au cours des 12 dernières années

L'excellente préparation de ces zones agricoles a donné une production de blé élevée. La **Figure I.3** montre la production de blé au cours de cette période.

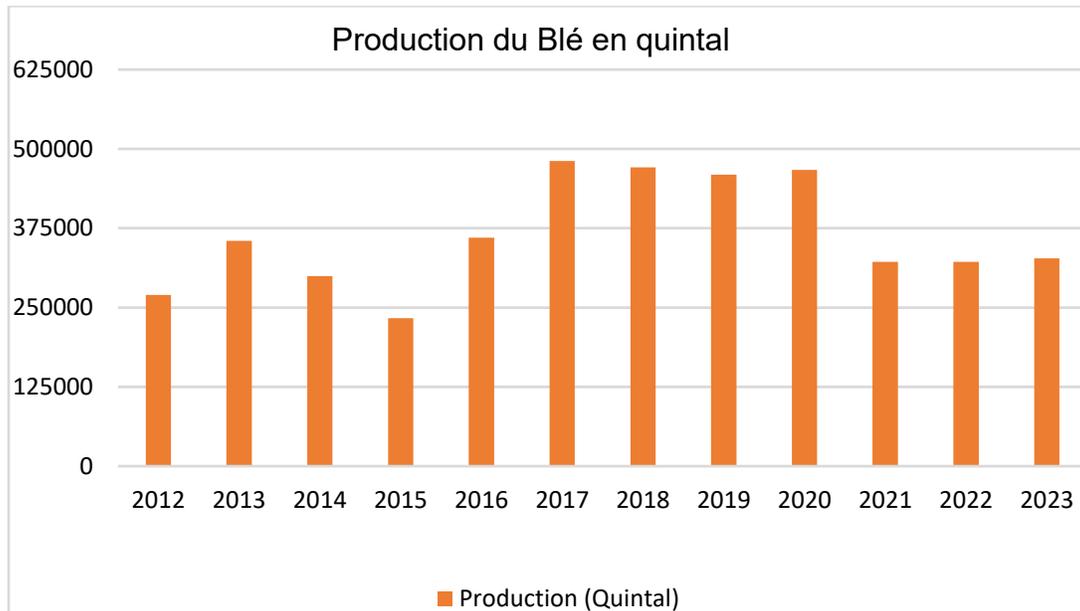


Figure I.3. *Production du blé en quintal au cours des 12 dernières années*

Les rendements des cultures, définis comme la quantité de produits agricoles par unité de surface, sont essentiels en agriculture. Ils sont influencés par divers facteurs tels que la météo, la qualité du sol et les techniques de gestion des cultures. L'optimisation des rendements est cruciale pour la rentabilité des agriculteurs et la sécurité alimentaire mondiale face à une population croissante.

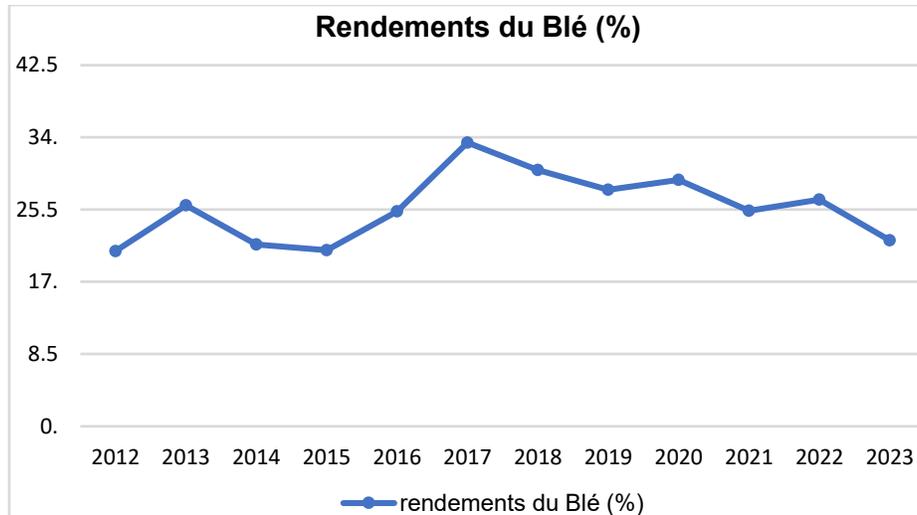


Figure I.4. Rendements agricoles du blé au cours des 12 dernières années

La formule mathématique pour calculer le rendement se présente comme suit :

$$\text{Rendement (\%)} = \frac{\text{Production en quintal}}{\text{Superficie en hectare}} \quad (\text{I.1})$$

De même, l'olive est considérée comme un aliment de base dans la région d'Annaba et dans les zones côtières de la mer Méditerranée. De plus, elle est destinée à l'exportation mondiale en raison du fait qu'elle est la matière première à partir de laquelle l'huile des olives est extraite. D'après la DGAA Les superficies allouées à la culture d'olive ont été couvertes de 2013 à 2023 et se répartissaient comme suit :



Figure I.5. Superficies allouées à la culture des olives en Hectare au cours des 10 dernières années

La production des olives au cours de ces années est représentée dans la **Figure I.6**.

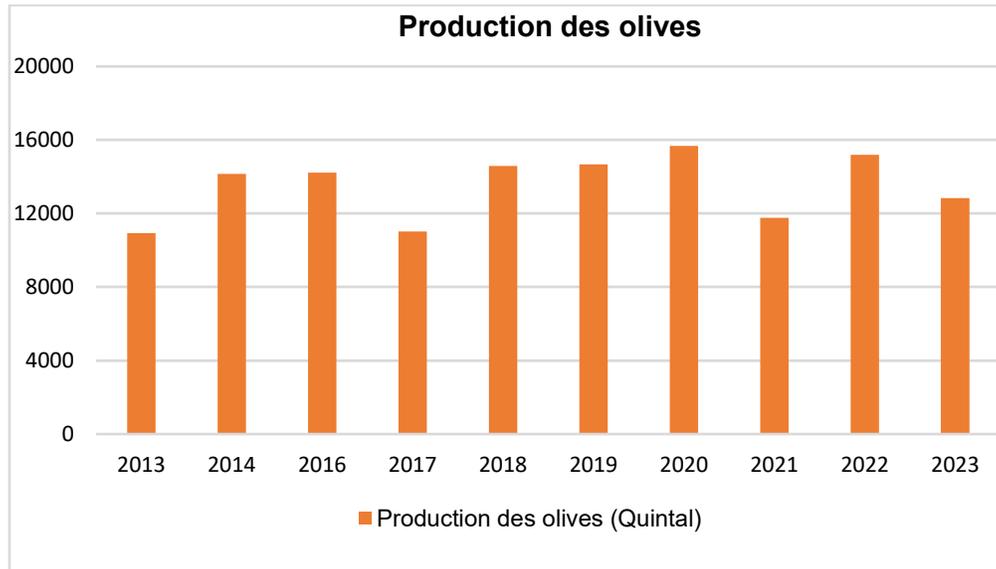


Figure I.6. Production des olives en quintal au cours des 10 dernières années

La **Figure I.7** représente le rendement des olives au cours des dix dernières années.

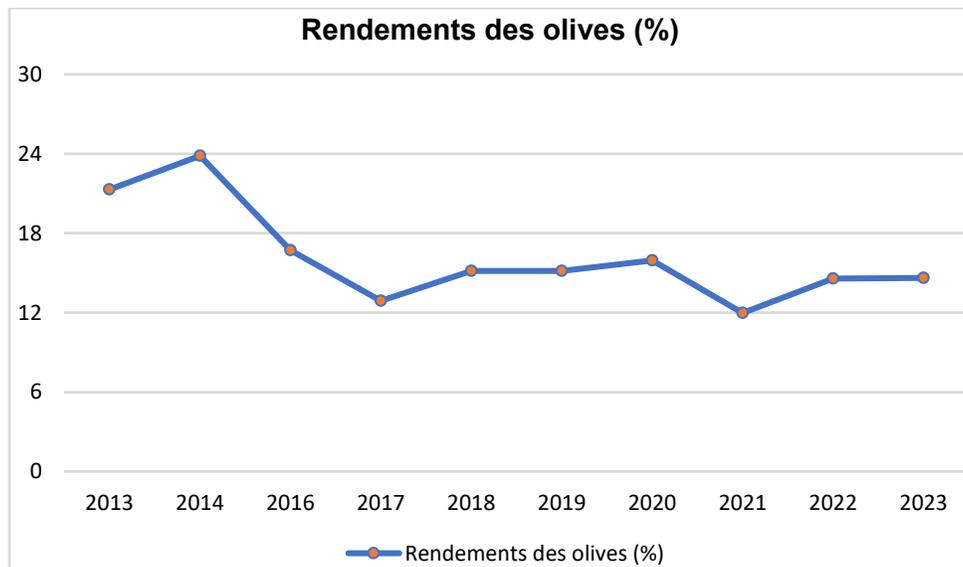


Figure I.7. Rendements agricoles des olives au cours des 10 dernières années

I.5 Problématique

I.5.1 Pourquoi la Prédiction des Rendements Agricoles ?

Les prévisions de rendement des cultures jouent un rôle essentiel dans l'agriculture contemporaine pour diverses raisons :

- La prévision précise de la disponibilité alimentaire permet d'évaluer la disponibilité alimentaire, ce qui est crucial pour faire face aux défis mondiaux en matière de sécurité alimentaire . [4]
- La prévention de la famine peut être facilitée par des prévisions précoces qui permettent de fournir des informations sur les éventuelles pénuries alimentaires et de permettre des interventions en temps voulu [5].
- L'agriculture durable est soutenue par les prévisions en optimisant l'utilisation des ressources et en augmentant la productivité.
- Anticiper la demande future : Étant donné l'anticipation d'une augmentation de la population mondiale, il est crucial de prévoir avec précision les rendements afin de faire face à la demande croissante de nourriture.
- Les agriculteurs, les agronomes et les décideurs ont la possibilité de prendre des décisions éclairées en se basant sur les prévisions de rendement afin d'améliorer les pratiques de gestion des cultures . [5]
- Amélioration des ressources : Les prévisions contribuent à améliorer l'efficacité de ressources telles que l'eau, les engrais et les pesticides, ce qui permet d'économiser des dépenses et de bénéficier de l'environnement. [5]
- Afin de mieux gérer les risques liés à la production agricole, les agriculteurs peuvent prévoir les variations éventuelles des rendements en anticipant les fluctuations potentielles .[5]

I.5.2 Le problème de prédiction des rendements agricoles

Comme nous l'avons déjà rappelé, la prévision des rendements agricoles revêt une grande importance à plusieurs égards. Bien entendu, de nombreux problèmes empêchent d'atteindre les résultats escomptés, comme en témoignent deux points critiques :

- Facteurs environnementaux : les conditions climatiques comprennent la température, l'humidité et le taux de précipitations, car ils affectent directement les cultures agricoles. Aussi la qualité du sol et la préparation des zones pour l'agriculture avec l'équipement nécessaire pour s'adapter à des situations spéciales. Ces facteurs ont une incidence directe sur les cultures. Si les cultures sont faibles, elles auront naturellement une incidence sur les prévisions de rendement.
- Manque de données : les données sont cruciales pour construire un modèle agricole efficace et précis. Le manque de données peut résulter d'une négligence administrative et du non-respect de la nécessité de surveiller en permanence les résultats des cultures lors de la construction du modèle.

I.6 État de l'art

I.6.1 Revue des Travaux Antérieurs

La prédiction des rendements agricoles est un domaine de recherche bien établi, avec un accent particulier sur l'utilisation de l'apprentissage automatique et de l'apprentissage profond. Voici quelques points clés :

- Dans [6], une analyse approfondie a été réalisée sur l'utilisation de l'apprentissage profond pour prévoir les rendements agricoles. Il a été constaté que les architectures d'apprentissage profond les plus couramment utilisées pour cette tâche sont les réseaux neuronaux convolutionnels et les LSTM (Long Short-Term Memory), en particulier pour les céréales comme le blé et le maïs. Bien que l'imagerie satellite soit couramment utilisée dans ces études, l'utilisation des véhicules aériens sans pilote (UAVs ou Unmanned aerial vehicle) pour la collecte de données est en augmentation.
- Une autre recherche a mis en évidence l'importance de comprendre les liens entre le rendement des cultures, les caractéristiques du sol et les caractéristiques topographiques

pour la gestion agricole [7]. Cette recherche a été réalisée à partir d'un ensemble de données provenant de plusieurs champs situés dans le sud-ouest de l'Ontario, au Canada. Les forêts aléatoires ont été identifiées comme étant les plus efficaces pour prédire le rendement, avec une valeur R^2 de 0,85 pour le maïs et 0,94 pour le soja.

- Une autre étude [8] a souligné que l'estimation et la prédiction du rendement des cultures fournissent des informations précieuses sur la dynamique de croissance des cultures, le développement génétique et la physiologie. En agronomie, les facteurs tels que les graines, les sols, les conditions météorologiques et les stratégies de gestion des champs sont étroitement liés au rendement des cultures.

Il est essentiel de noter que les résultats peuvent varier en fonction des régions et des cultures spécifiques étudiées. De plus, l'efficacité des techniques d'apprentissage automatique et d'apprentissage profond dépend de la qualité des données. Il est donc crucial de disposer de données précises et à jour pour la prédiction des rendements agricoles.

I.6.2 Méthodes Traditionnelles et Techniques Modernes

Au fil des ans, les techniques de prédiction des rendements agricoles ont progressé, passant de méthodes traditionnelles à des approches plus avancées et modernes. Voici un résumé de ces techniques :

a. Méthodes traditionnelles

- Identification des cultures : Les techniques traditionnelles, comme l'identification des cultures, ont été utilisées par les agriculteurs et les agronomes pour estimer les rendements. [6]
- Jugement d'expert : Les prévisions de rendement reposaient souvent sur le savoir-faire des experts et des agriculteurs.[9]
- Modèles statistiques : Des modèles statistiques, tels que l'analyse de régression et la prévision de séries temporelles, ont été utilisés pour prédire les rendements futurs en combinant des facteurs de gestion et environnementaux avec des données de rendement historiques.[6]

b. Techniques modernes

- Modèles d'apprentissage automatique : Des algorithmes d'apprentissage automatique, tels que les forêts aléatoires, les réseaux de neurones artificiels, les machines à vecteurs de support, entre autres, ont été employés pour prédire les rendements agricoles. [10-7]
- Apprentissage profond : Les architectures d'apprentissage profond, comme les réseaux neuronaux convolutionnels et les LSTM (Long Short-Term Memory), sont de plus en plus utilisées pour la prédiction des rendements agricoles.[6]
- Télédétection : La télédétection est une méthode permettant d'obtenir des informations sur des cibles sans contact direct, et les informations ou les données sont recueillies par des capteurs à distance[8]. Des indices de végétation normalisés, tels que l'indice de végétation différentielle normalisé (NDVI) et l'indice de végétation amélioré (EVI), sont couramment utilisés pour surveiller la croissance des cultures et prédire les rendements.[3]

Il est essentiel de souligner que l'efficacité de ces méthodes peut varier en fonction de divers facteurs, y compris la qualité des données disponibles, la spécificité de la culture et de la région, et les conditions environnementales. De plus, l'utilisation de techniques modernes nécessite souvent une expertise technique et une capacité à interpréter et à appliquer correctement les résultats des modèles.

I.7 Limitations d'étude

I.7.1 Limites potentielles du projet

Lorsqu'on envisage de mettre en place un projet qui vise à prédire les rendements agricoles en utilisant des modèles d'apprentissage profond, il est crucial de prendre en compte plusieurs obstacles potentiels :

- Disponibilité des données : La disponibilité des données nécessaires pour entraîner le modèle peut être un défi. Ces données peuvent inclure des informations sur les conditions météorologiques, la qualité du sol, l'utilisation d'engrais et de pesticides, etc. Il peut être difficile d'obtenir ces données pour toutes les régions et pour une période de temps suffisamment longue.

- Qualité des données : Même si les données sont disponibles, leur qualité peut varier. Par exemple, les mesures de température ou de précipitations peuvent être inexactes, ou les données sur les rendements agricoles peuvent être incomplètes ou biaisées.
- Déterminer le type de culture pour notre modèle : Nous devons déterminer la culture qui contient une grande base de données et de bonne qualité, et dans ce cas, le blé représente la meilleure base de données, tandis que les olives contiennent presque la même taille que la base de données, mais avec une qualité inférieure, comme nous l'avons mentionné précédemment.
- Ressources : L'apprentissage profond nécessite généralement des ressources informatiques importantes, notamment en termes de puissance de calcul et de stockage. Cela peut être une contrainte, surtout si vous travaillez avec de grands ensembles de données.

I.8 Conclusion

Ce premier chapitre a posé les bases d'une étude détaillée sur la prédiction des rendements agricoles dans la région fertile d'Annaba. En soulignant le rôle crucial de l'agriculture pour l'économie locale et la société, et en mettant en avant l'importance du blé et des olives comme éléments fondamentaux de l'agriculture régionale, nous avons créé un contexte riche et détaillé pour notre enquête. La clarification précise de notre question de recherche et de nos objectifs, associée à une analyse approfondie de l'état actuel des connaissances, constitue le point de départ pour la création de modèles prédictifs novateurs.

Dans le chapitre suivant, nous nous concentrons sur l'étude de l'apprentissage profond afin de trouver des solutions efficaces pour prévoir les rendements du blé. Notre ambition est grande : contribuer de manière significative à l'amélioration des pratiques agricoles et influencer positivement la politique agricole dans la région d'Annaba.

CHAPITRE II : Introduction aux Réseaux neuronaux récurrents

II.1 Introduction

Ce chapitre explore l'application des réseaux de neurones récurrents (RNNs) et leurs variantes, telles que la mémoire à long terme (LSTM), LSTM bidirectionnelle (BiLSTM), et les unités récurrentes synchronisées (GRU), dans la prévision des rendements agricoles.

Les modèles d'apprentissage automatique traditionnels ont souvent du mal à capturer les dépendances temporelles, ce qui conduit au développement de l'ARN et de ses variantes. Les ARN ont montré des résultats prometteurs dans des tâches telles que la prévision du rendement agricole. Le chapitre explore le modèle LSTM, une amélioration des réseaux de neurones récurrents (RNN), conçus pour surmonter les limitations des RNN., le modèle BiLSTM, qui étend les capacités du LSTM en traitant les données dans les deux sens, et le modèle GRU, qui offre une alternative efficace en calcul avec des performances comparables à LSTM.

Le chapitre traite également de divers hyperparamètres impliqués dans la construction de ces modèles, tels que le nombre d'unités par couche, le nombre de couches, le nombre de fois et le taux d'apprentissage. Le chapitre vise à approfondir notre compréhension de ces modèles, de leurs forces et faiblesses, et de leur application efficace dans la prévision des rendements agricoles.

II.2 Aperçu sur les différentes branches de l'intelligence artificielle

L'intelligence artificielle (IA) est un domaine vaste et en constante évolution.

II.2.1 Intelligence artificielle (AI)

L'intelligence artificielle (IA) permet aux machines d'imiter ou de surpasser les capacités cognitives humaines. Elle se base sur diverses théories et technologies, et est interprétée principalement à travers le test de Turing et la perspective de Marvin Minsky. L'IA se présente sous plusieurs formes [11]:

- Les **machines réactives** qui réagissent à leur environnement sans utiliser de souvenirs ou d'expériences passées.
- La **mémoire limitée** qui utilise les expériences passées pour prendre des décisions actuelles avec une capacité de mémoire limitée.
- La **théorie de l'esprit** qui comprend et raisonne sur les états mentaux.

- La **conscience de soi** qui montre une conscience de soi et peut réfléchir à ses propres états mentaux et émotions.

L'IA a le potentiel de transformer divers secteurs, mais soulève également des préoccupations éthiques et sociétales, notamment le déplacement d'emplois et les questions de confidentialité.[12]

II.2.2 Machine Learning (ML)

L'apprentissage automatique (ML), une sous-branche de l'IA, se concentre sur le développement d'algorithmes qui permettent aux machines d'apprendre à partir de données et de prendre des décisions sans programmation explicite. Les algorithmes de ML sont classés en trois catégories :

- **Apprentissage supervisé** : Il apprend la relation entre les entrées et les sorties connues pour faire des prédictions précises sur de nouvelles données.[13]
- **Apprentissage non supervisé** : Il découvre des modèles cachés dans les données sans valeurs cibles connues.[13]
- **Apprentissage par renforcement** : Il apprend en interagissant avec un environnement et en recevant des rétroactions sous forme de récompenses ou de punitions.[13]

Ces algorithmes utilisent des modèles mathématiques et des algorithmes d'optimisation pour apprendre et faire des prédictions. Ils trouvent des applications dans divers domaines, y compris le traitement du langage naturel, la vision par ordinateur, la prévision des rendements des cultures et les systèmes de recommandation.

II.2.3 Réseaux de neurones artificiels (ANN)

Les réseaux de neurones artificiels (ANN ou Artificiel Neural Network), inspirés par le cerveau humain, sont des algorithmes d'apprentissage automatique composés de neurones artificiels interconnectés. Ils sont formés sur des données étiquetées par apprentissage supervisé, ajustant les poids des connexions pour minimiser l'erreur de prédiction. Une fois formés, ils peuvent faire des prédictions sur de nouvelles données.

Les ANN sont utilisés dans diverses tâches, y compris la classification d'images, la prévision des rendements des cultures, le traitement du langage naturel et les systèmes de recommandation. Il existe différents types d'ANN, chacun adapté à différents types de problèmes.

Cependant, les ANN ont des limites [14]. Ils peuvent être gourmands en calcul, sujets au surajustement, et ont souvent du mal à expliquer leurs prédictions. Ces défis font l'objet de recherches en cours dans le domaine des ANN et de l'apprentissage automatique.

II.2.4 L'apprentissage profond (DL)

Les algorithmes de Deep Learning (DL), une catégorie spécialisée d'algorithmes Machine Learning, sont conçus pour entraîner des réseaux neuronaux profonds multicouches. Ces algorithmes facilitent l'apprentissage de modèles complexes et de représentations à partir de vastes ensembles de données en utilisant la structure hiérarchique des réseaux neuronaux profonds.

Voici quelques algorithmes d'apprentissage profond courants [15] :

- a. Réseaux neuronaux convolutionnels (CNN) (Convolutional Neural Networks) : Principalement utilisés dans les tâches de vision par ordinateur, les CNN excellent dans la reconnaissance d'images, la détection d'objets et la segmentation d'images. Ils utilisent des couches convolutionnelles pour extraire des caractéristiques des images d'entrée et sont reconnus pour leur capacité à capturer des hiérarchies spatiales dans les données.
- b. Réseaux neuronaux récurrents (RNN) (Recurrent Neural Networks) : Les RNN sont adaptés pour les tâches de traitement de données séquentielles, telles que le traitement du langage naturel et l'analyse de séries temporelles. Ils possèdent des boucles de rétroaction qui leur permettent de conserver des informations sur les entrées précédentes, les rendant aptes pour les tâches impliquant des séquences de données.
- c. Réseaux antagonistes génératifs (GAN) (Generative Adversarial Networks) : Les GAN comprennent deux réseaux neuronaux, un générateur et un discriminateur, qui sont entraînés simultanément dans un cadre compétitif. Les GAN sont utilisés pour générer de nouveaux échantillons de données, tels que des images, de la musique ou du texte, et trouvent des applications dans la génération d'images, le transfert de style et l'augmentation des données.

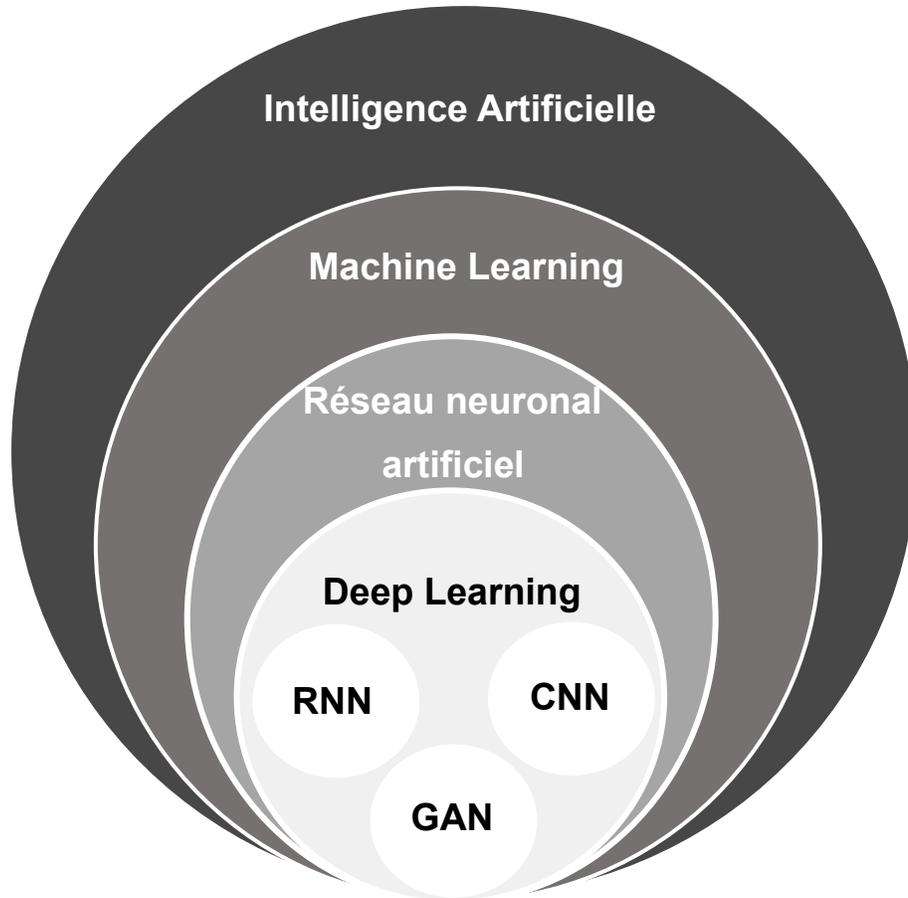


Figure II.1. *Les branches de l'intelligence Artificielle*

Selon la nature de nos données actuelles, il est judicieux d'utiliser des modèles de type RNN pour deux raisons. Premièrement, nos données sont numériques et non pas des images ou autres. Deuxièmement, nous avons besoin de modèles qui exploitent les observations précédentes pour effectuer des prédictions futures. Cela est possible grâce aux réseaux neuronaux récurrents où l'information circule à travers une boucle, permettant de se souvenir des résultats précédents. Cette caractéristique nous a orientés vers les modèles RNN plutôt que les modèles CNN.

II.3 Les réseaux de neurones récurrents (RNN)

Les réseaux de neurones récurrents (RNN ou Recurrent Neural Network) sont une classe spécifique de réseaux de neurones artificiels conçus pour traiter des données séquentielles comme des informations de séries temporelles, des chaînes d'événements et des tâches liées au traitement du langage naturel. L'architecture unique des RNN, qui comprend des boucles internes pour le

retour d'information, leur permet d'apprendre et de reconnaître des motifs dans les séquences et de modéliser les dépendances dans le temps en formant une mémoire des entrées précédentes.[15]

Cependant, les conceptions conventionnelles de RNN rencontrent des problèmes tels que le problème du gradient qui s'évanouit, où l'impact des souvenirs antérieurs s'estompe avec le temps, entraînant des défis dans l'apprentissage des dépendances sur une longue durée. Pour atténuer cela, des conceptions de RNN plus sophistiquées ont été introduites, y compris les réseaux de mémoire à long terme (LSTM). Les LSTM intègrent des mécanismes complexes, y compris des mécanismes d'attention, pour améliorer la rétention et l'utilisation des informations des étapes de temps précédentes, facilitant ainsi un apprentissage plus efficace des dépendances à long terme dans les données séquentielles.[15]

Les RNN sont fréquemment utilisés dans une variété de tâches, y compris la prévision de séries temporelles, la prédiction du comportement des processus et des tâches de traitement du langage naturel telles que la transduction de séquences et la traduction de machines neuronales. Grâce à leur capacité à reconnaître des motifs séquentiels et des dépendances temporelles, les RNN ont démontré leur valeur dans un large éventail d'applications où l'interprétation et le traitement des données séquentielles sont cruciaux.

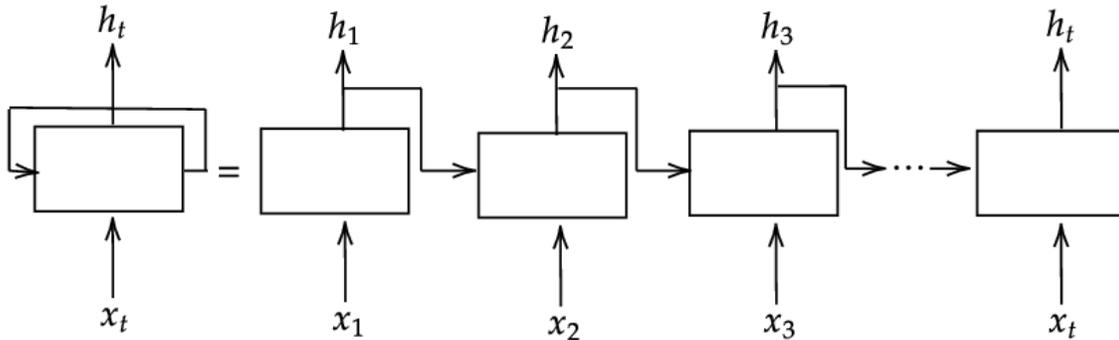


Figure II.2. Réseau neuronal récurrent

La sortie d'une unité RNN est calculée par équation (II.1) :

$$h_t = \tanh(W_x x_t + W_h h_{t-1} + b_t) \quad (\text{II.1})$$

Où h_{t-1} est la sortie récurrente de l'étape précédente, x_t est l'entrée à l'heure actuelle, et W_x, W_h, b_t sont les poids et biais du réseau à former lors de la formation. Le problème avec RNN est

que si l'entrée séquentielle est grande, le gradient de la fonction de perte peut se terminer à zéro, empêchant effectivement les poids du modèle d'être mis à jour.[16]

II.3.1 Structure du réseau neuronal récurrent LSTM

Les réseaux de neurones à mémoire à long terme (LSTM ou Long Short Term Memory) sont une amélioration des réseaux de neurones récurrents (RNN), conçus pour surmonter les limitations des RNN. Les RNN ont du mal à retenir des informations pendant de longues périodes et à gérer les "dépendances à long terme". Ils rencontrent également des problèmes tels que les gradients explosifs et disparus lors de la rétropropagation. En revanche, les LSTM éliminent presque complètement le problème du gradient disparu, gèrent le bruit, les représentations distribuées et les valeurs continues. Ils offrent une large gamme de paramètres, y compris les taux d'apprentissage, et les biais d'entrée et de sortie.

L'unité LSTM est équipée d'une porte d'oubli qui peut être utilisée pour instruire les neurones individuels sur l'importance de l'information et sa durée d'importance. Une unité LSTM standard est composée d'une entrée de bloc z_t , d'une porte d'entrée i_t , d'une porte d'oubli f_t , d'une porte de sortie o_t et d'une cellule de mémoire c_t (voir **Figure II.3**).

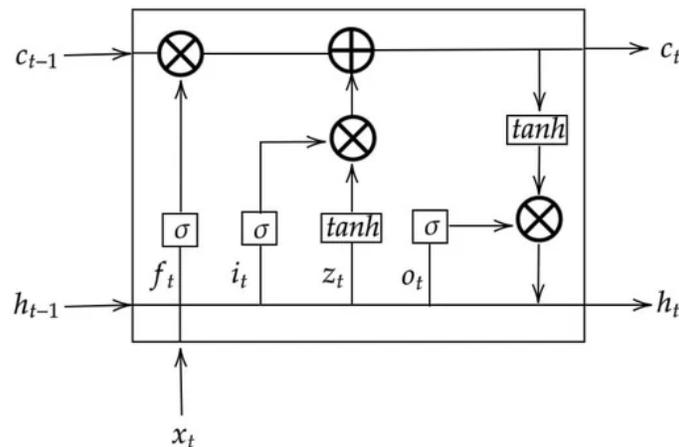


Figure II.3. Mémoire à long terme (LSTM)

La porte d'oubli f_t est utilisée pour éliminer les informations de l'état de la cellule qui ne sont plus nécessaires, selon l'équation (II.2). L'entrée à un moment spécifique x_t et la sortie de la cellule précédente h_{t-1} sont fournies à la porte, multipliées par des matrices de poids, puis ajoutées au biais. Le résultat est traité par une fonction sigmoïde qui donne un nombre entre 0 et 1. Si la sortie

est 0, l'information est oubliée pour un état de cellule donné ; si la sortie est 1, l'information est conservée pour une utilisation future.

La porte d'entrée i_t est utilisée pour ajouter des informations précieuses à l'état de la cellule, selon l'équation (II.3). Initialement, l'information est régulée par la fonction sigmoïde, qui filtre les valeurs à stocker, de manière similaire à la porte d'oubli. Par la suite, un vecteur de nouvelles valeurs candidates de h_{t-1} et x_t est produit avec la porte de bloc z_t en utilisant l'équation (II.4), qui donne des sorties de -1 à $+1$. Les valeurs du vecteur et les valeurs régulées sont multipliées pour obtenir des informations utiles selon l'équation (II.5).[17]

La porte de sortie o_t détermine quelle information dans la cellule est utilisée pour calculer la sortie de l'unité LSTM, selon l'équation (II.6).

$$f_t = \sigma(W_{fx} x_t + W_{fh} h_{t-1} + b_f) \quad (\text{II.2})$$

$$i_t = \sigma(W_{ix} x_t + W_{ih} h_{t-1} + b_i) \quad (\text{II.3})$$

$$z_t = \tanh(W_{zx} x_t + W_{zh} h_{t-1} + b_z) \quad (\text{II.4})$$

$$c_t = f_t * c_{t-1} + i_t * z_t \quad (\text{II.5})$$

$$o_t = \sigma(W_{ox} x_t + W_{oh} h_{t-1} + b_o) \quad (\text{II.6})$$

Au départ, un vecteur est formé en appliquant la fonction \tanh à la cellule. Par la suite, l'information est normalisée à l'aide de la fonction sigmoïde, qui filtre les valeurs à stocker en tenant compte des entrées h_{t-1} et x_t . Les valeurs du vecteur et les valeurs normalisées sont multipliées selon l'équation (II.7) pour être envoyées en sortie et en entrée à la cellule suivante :

$$h_t = o_t * \tanh(c_t) \quad (\text{II.7})$$

II.3.2 Structure du réseau neuronal récurrent (GRU)

Le réseau neuronal récurrent GRU ou Gated Recurrent Unit, une forme de réseau neuronal récurrent, a été proposée en 2014 comme une alternative moins complexe aux réseaux de mémoire à long terme. Semblable à LSTM, GRU peut gérer des données séquentielles. Il utilise des

mécanismes de portage, à savoir la porte de réinitialisation et la porte de mise à jour, pour modifier sélectivement l'état caché du réseau à chaque étape de temps, gérant ainsi le flux d'informations.

Les GRU éliminent l'état de la cellule et utilisent l'état caché pour la transmission d'informations [18]. Cette structure comprend seulement deux portes : la porte de mise à jour z_t et la porte de réinitialisation r_t . Analogues aux portes LSTM, les portes GRU sont formées pour exclure sélectivement toutes les informations non essentielles tout en conservant les informations précieuses, et peuvent être ajustées en utilisant les équations (II.8) – (II.11) :

$$z_t = \sigma(W_{ux} x_t + W_{uh} h_{t-1} + b_u) \quad (\text{II.8})$$

$$r_t = \sigma(W_{rx} x_t + W_{rh} h_{t-1} + b_r) \quad (\text{II.9})$$

$$o_t = \tanh(W_{ox} x_t + W_{oh} h_{t-1} + b_o) \quad (\text{II.10})$$

$$h_t = (1 - z_t) * o_t + z_t * h_{t-1} \quad (\text{II.11})$$

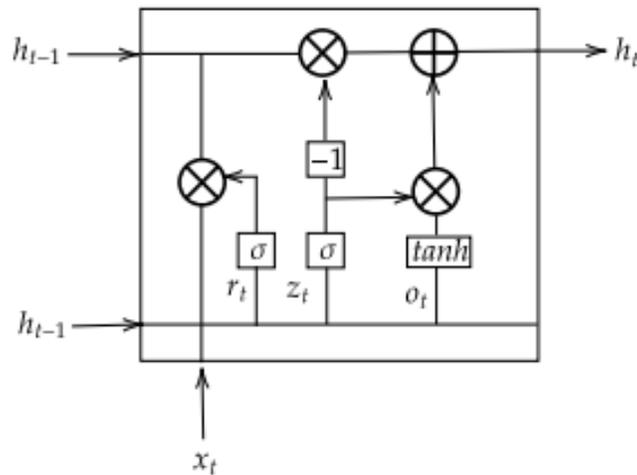


Figure II.4. Gated Recurrent Unit (GRU)

Les fonctions \tanh et σ ajoutent une non-linéarité au réseau. Ces fonctions permettent au modèle de saisir les relations non linéaires entre les entrées et les sorties du modèle. Au début de l'entraînement, les poids et les biais dans les équations (II.2)– (II.4), (II.6) et (II.8)– (II.10) sont définis de manière aléatoire. Pendant l'entraînement, le modèle essaie de régler les poids et les

biais de manière à minimiser la fonction de perte. Par conséquent, l'algorithme d'un modèle RNN est un problème d'optimisation. Les modèles LSTM et GRU ont des unités similaires, mais ils diffèrent également. Par exemple, dans l'unité LSTM, la quantité de contenu de mémoire vue ou utilisée par d'autres unités du réseau est contrôlée par la porte de sortie. En revanche, le GRU libère tout son contenu sans aucun contrôle [17]. À partir de ces seules similitudes et différences, il est difficile de conclure quel modèle fonctionne mieux sur un problème qu'un autre. Dans cet article, les deux modèles ont été mis en œuvre pour voir quel modèle fonctionne mieux sur le problème d'estimation du rendement.

II.3.3 Structure de Bidirectionnel LSTM (BiLSTM)

BiLSTM est capable de bien mémoriser une séquence, de résoudre le problème de dépendance sur des durées plus longues et présente un avantage considérable pour améliorer la corrélation des séries temporelles. Basé sur le réseau LSTM, les structures de réseau avant et arrière forment une boucle fermée d'information, ce qui permet de mieux vérifier et corriger les informations d'erreur du processus tout en conservant les informations de données bidirectionnelles, ce qui confère une robustesse accrue.[19]

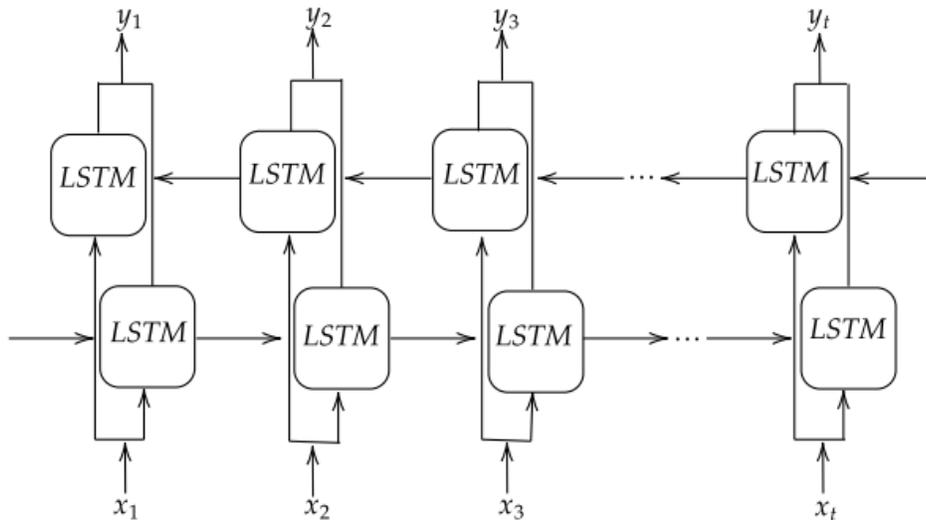


Figure II.5. Couche LSTM bidirectionnelle

II.4 Amélioration de l'apprentissage

II.4.1 La couche Dropout

Le surapprentissage se produit lorsque le modèle apprend le bruit dans les données d'entraînement et la généralisation du modèle aux données invisibles est peu fiable et ne peut pas faire des estimations précises.

Il est proposé d'utiliser des techniques de régularisation pour prévenir le surapprentissage. L'une d'elles est le « dropout », qui consiste à sélectionner aléatoirement certains neurones de la couche cachée et à bloquer leur sortie afin qu'ils ne soient pas évalués dans les algorithmes d'apprentissage, puis, après un certain temps, à libérer les sorties des neurones bloqués et à bloquer d'autres neurones[20]. Cela conduit à ce que le réseau neuronal devienne plus général et ne dépende pas uniquement d'un groupe de neurones pour prendre certaines décisions.

Dans les modèles LSTM, le dropout peut être ajouté aux entrées des couches, aux sorties des couches, et aux sorties récurrentes[20]. La taille du dropout est un hyperparamètre qui doit être défini pendant l'entraînement (voir **Figure II.6**).

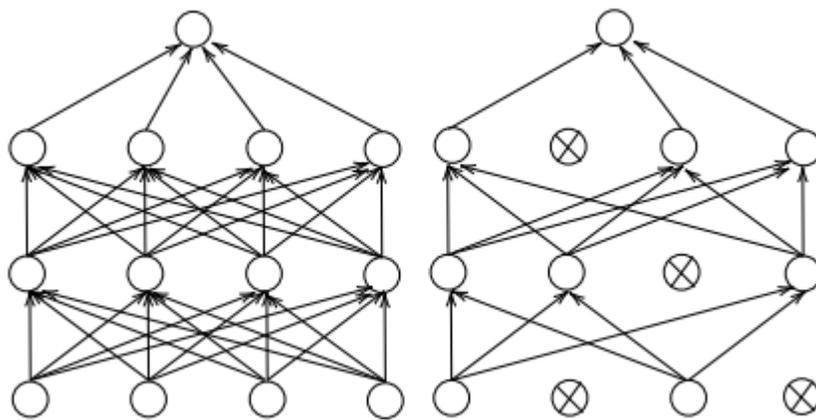


Figure II.6. Comparaison entre le réseau neuronal normal (à gauche) et réseau avec dropout (à droite)

II.4.2 La couche Batch Normalization

La normalisation par lots, est une technique conçue pour accélérer et stabiliser le processus d'entraînement des réseaux de neurones artificiels.

Le principe de base de cette technique est d'ajuster les entrées de chaque couche afin que leur activation de sortie moyenne soit nulle et que leur écart-type soit de un. Ceci est accompli en normalisant, re-entrant, et re-caling les entrées de chaque couche.

À l'origine, la normalisation par lots a été développée pour résoudre le problème du « décalage des covariables internes ». Ce phénomène se réfère aux altérations de la distribution d'entrée aux couches internes pendant le processus d'entraînement, qui peuvent être influencées par le caractère aléatoire de l'initialisation des paramètres et la variabilité des données d'entrée.

Pour résumer, la normalisation par lots est une méthode puissante qui peut stabiliser le processus d'apprentissage, réduire considérablement le nombre d'époques de formation nécessaires pour former des réseaux profonds et offrir une certaine régularisation, réduisant ainsi les erreurs de généralisation.[8]

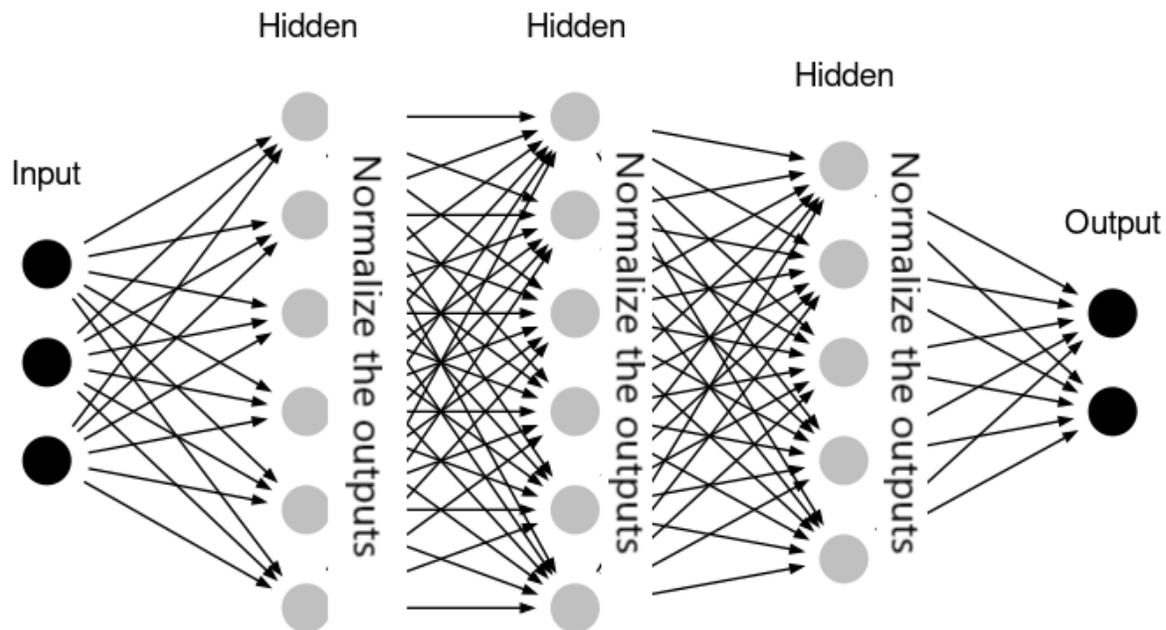


Figure II.7. *La couche Batch Normalization*

II.4.3 La couche Dense

Une « couche dense », également appelée « couche entièrement connectée », est un élément clé des réseaux de neurones. Dans une telle couche, chaque neurone est lié à chaque neurone de la couche précédente, créant une connexion profonde, c'est pourquoi on l'appelle une couche "dense".

L'opération d'une couche dense implique une multiplication de vecteurs matriciels. Sa fonction principale est de modifier les dimensions du vecteur d'entrée et d'implémenter des transformations comme la rotation, l'échelle et la traduction.

Fondamentalement, une couche dense est une composante essentielle de nombreux modèles en raison de sa capacité à apprendre d'un vaste éventail de fonctionnalités. Il est généralement utilisé après un bloc d'extraction d'entités (telles que la convolution, l'encodeur ou le décodeur), comme couche de sortie finale ou pour projeter un vecteur d'une dimension à une autre.[9]

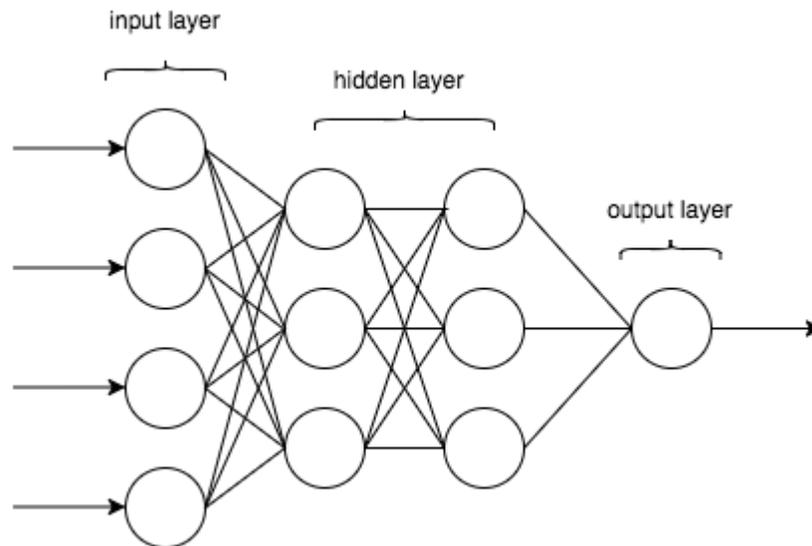


Figure II.8. *La couche Dense*

II.5 Paramètres d'apprentissage en profondeur

II.5.1 Adam (Adaptive Moment Estimation)

Adam est une technique d'optimisation largement reconnue, fréquemment utilisée dans le domaine de l'apprentissage profond pour ajuster les paramètres d'un réseau neuronal tout au long du processus d'entraînement. Elle fusionne ingénieusement les principes de Momentum et de taux d'apprentissage adaptatif pour faciliter des modifications de paramètres efficaces et puissantes.[21]

Le principe fondamental sous-jacent à Adam est l'ajustement du taux d'apprentissage pour chaque paramètre, qui est basé sur les estimations des premières et deuxièmes moments des gradients. Il maintient une moyenne en déclin exponentiel des gradients passés (le premier moment) et de leurs carrés (le deuxième moment). Ces estimations sont ensuite utilisées pour mettre à jour les paramètres en ajustant convenablement les gradients.

L'ajustement du taux d'apprentissage adaptatif d'Adam est essentiel pour gérer des gradients de différentes échelles et des besoins divers en taux d'apprentissage à travers les paramètres. Cette caractéristique s'avère particulièrement avantageuse dans des scénarios où les gradients sont clairsemés ou sujets au bruit. L'algorithme intègre également des étapes pour corriger le biais, ce qui permet de traiter le biais introduit par les estimations initiales des moments.

En raison de ses performances impressionnantes sur un large éventail de tâches d'apprentissage profond, Adam a gagné une popularité considérable. Il atteint souvent la convergence plus rapidement que les algorithmes d'optimisation conventionnels tels que la descente de gradient stochastique (SGD), et est capable de gérer efficacement de grands ensembles de données. Adam est réputé pour sa robustesse, sa convivialité et ses performances généralement fiables à travers différentes architectures de réseau et domaines de problèmes.[22]

II.5.2 RMSprop (Root Mean Square Propagation)

RMSprop, ou Propagation quadratique moyenne, est une technique d'optimisation répandue dans le domaine de l'apprentissage profond. Elle offre des solutions aux lacunes des méthodes

conventionnelles de descente de gradient en ajustant le taux d'apprentissage pour chaque paramètre, en tenant compte des gradients historiques. Une caractéristique significative de RMSprop est sa capacité à gérer des gradients de différentes échelles en préservant une moyenne mobile de gradients carrés, permettant des mises à jour stables et équilibrées sur toutes les dimensions de poids.[23]

De plus, RMSprop intègre un taux de déclin ou un terme de *momentum* qui influence la contribution des gradients précédents. L'inclusion du *momentum* confère une inertie aux mises à jour de poids, facilitant la convergence dans les zones avec des gradients constants. Cette caractéristique aide à traverser les régions à haute courbure et accélère le processus d'optimisation.

Le taux d'apprentissage adaptatif et le *momentum* de RMSprop en font un choix idéal pour l'optimisation des réseaux neuronaux profonds. Il relève les défis posés par les magnitudes de gradient fluctuantes et les mises à jour de poids problématiques qui pourraient entraver la convergence ou induire des oscillations.

La mise en œuvre de RMSprop est simple, étant donné sa disponibilité immédiate dans les bibliothèques d'apprentissage profond [24]. Ces bibliothèques offrent des fonctions intégrées qui permettent une intégration facile de RMSprop en tant qu'algorithme d'optimisation de choix.

II.6 Hyperparamètres d'un réseau de neurones

Les configurations qui définissent la conception et le comportement d'entraînement d'un réseau de neurones sont appelées hyperparamètres. Voici quelques réglages typiques des hyperparamètres :

II.6.1 Unités par couche (Number of Units in Each Layer)

Le nombre d'unités (ou neurones) dans chaque couche est un facteur clé qui influence la capacité et la puissance expressive du réseau. Un nombre plus élevé d'unités pourrait permettre au réseau de capturer des relations plus complexes au sein des données, mais pourrait également augmenter le risque de surajustement. Le nombre optimal d'unités dépend de la complexité du problème et du volume de données disponibles.

II.6.2 Nombre de couches cachées (Number of Hidden Layers)

Le terme 'couches cachées' désigne les couches qui se trouvent entre les couches d'entrée et de sortie d'un réseau neuronal. L'ajout de couches cachées supplémentaires permet au réseau d'apprendre des représentations sophistiquées et de détecter potentiellement des motifs plus complexes dans les données. Cependant, l'augmentation du nombre de couches cachées augmente également la complexité du modèle, ce qui pourrait nécessiter plus de données et de ressources informatiques pour une formation efficace.

II.6.3 Nombre d'époques (Number of Epochs)

Une 'époque' signifie un balayage complet de l'ensemble des données d'entraînement. Le nombre d'époques détermine le nombre de fois que le réseau sera formé sur l'ensemble des données. Un nombre plus élevé d'époques permet au réseau de revoir les données plus fréquemment et de converger potentiellement vers une solution améliorée. Cependant, un nombre excessif d'époques d'entraînement peut entraîner un surajustement, où le modèle devient excessivement adapté aux données d'entraînement et sous-performe sur les données inédites.

II.6.4 Taux d'apprentissage (Learning Rate)

Le taux d'apprentissage régit l'ampleur des étapes lors des mises à jour des paramètres dans le processus d'optimisation. Il dicte la vitesse à laquelle le réseau apprend à partir des gradients. Un taux d'apprentissage plus élevé pourrait accélérer la convergence, mais pourrait également induire une instabilité. À l'inverse, un taux d'apprentissage plus faible pourrait ralentir la convergence mais produire des résultats plus précis.

II.6.5 Fonctions d'activation (Activation Functions)

Les fonctions d'activation jouent un rôle crucial dans les réseaux de neurones car elles introduisent la non-linéarité, permettant au modèle d'apprendre des modèles complexes et de faire des prédictions. Voici trois couramment utilisés fonctions d'activation :

a. La fonction «*sigmoid*» :

La fonction *sigmoid* est une forme spéciale de la fonction logistique utilisée comme fonction d'activation dans les réseaux neuronaux. Il introduit la non-linéarité dans le modèle, permettant au réseau neuronal d'apprendre des limites de décision plus complexes. La fonction est indiquée par $\sigma(x)$ est donnée par la formule mathématique :

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (\text{II.12})$$

Les principales propriétés de la fonction sigmoïde comprennent un graphique de courbe en forme de S, un domaine de $(-, +)$, une plage de $(0,1)$ une augmentation monotone, continue et différentiable partout dans son domaine, et des valeurs proches de zéro ou un pour les valeurs inférieures à -10 ou supérieures à 10 . La fonction sigmoïde est historiquement importante dans le développement des réseaux de neurones et a de nombreuses applications utiles aujourd'hui.[10]

b. La fonction «*tanh*» :

La fonction *tanh*, également connue sous le nom de tangente hyperbolique, est une fonction d'activation largement utilisée dans les réseaux de neurones. Elle introduit une non-linéarité dans le modèle, permettant au réseau de neurones d'apprendre des frontières de décision plus complexes. La fonction tanh est dénotée par $\tanh(x)$ et est donnée par la formule mathématique :

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (\text{II.13})$$

Il possède des propriétés clés telles qu'un graphique courbe en forme de S, augmentant de manière monotone, continu et différentiable partout dans son domaine. Il est symétrique autour de l'origine $(0, 0)$, ce qui le rend utile pour les données centrées autour de zéro. Dans les réseaux de neurones, La fonction tanh est utilisé comme une fonction d'activation, passant une somme pondérée d'entrées à travers elle, qui sert d'entrée à la couche suivante. Il est historiquement important dans le développement de réseaux de neurones et a de nombreuses applications aujourd'hui.[11]

c. La fonction «*ReLU*» :

L'unité linéaire rectifiée (*ReLU*) est une fonction d'activation populaire utilisée dans les réseaux neuronaux. Il introduit la non-linéarité dans le modèle, permettant au réseau neuronal d'apprendre

des limites de décision plus complexes. La fonction $ReLU$ est généralement désignée par $ReLU(x)$ et est donnée par la formule mathématique suivante :

$$ReLU(x) = \max(0, x) \quad (II.14)$$

La fonction $ReLU$ est une fonction de montée en champ qui produit des entrées positives directement, augmente de manière monotone, continue et différentiable, sauf à zéro, et est nulle pour les entrées négatives et linéaire pour les entrées positives. Il est utilisé comme fonction d'activation dans les réseaux de neurones, passant une somme pondérée d'entrées comme entrée à la couche suivante. Ce processus aide les réseaux de neurones à comprendre et à apprendre des modèles de données complexes. La fonction $ReLU$ est historiquement importante dans le développement de réseaux de neurones et a de nombreuses applications utiles aujourd'hui, en particulier dans les réseaux multicouches de Perceptron et de neurones convolutionnels.[12]

II.7 Métriques d'évaluation

II.7.1 Mean squared Error (MSE)

Erreur Quadratique Moyenne (EQM) ou MSE pour Mean squared Error est une fonction de perte fréquemment utilisée dans les tâches de régression. Dans le cadre de la régression, la MSE évalue la moyenne des carrés des erreurs ou résidus, qui sont les différences entre les valeurs prédites et les valeurs réelles. Mathématiquement, la MSE est calculée comme suit :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (II.15)$$

n : est le nombre d'échantillons

y_i : est la valeur réelle de la variable cible pour le i échantillon

\hat{y}_i : est la valeur prévue de la variable cible pour le i échantillon

L'objectif de minimiser la MSE pendant l'entraînement est de trouver les paramètres du modèle qui réduisent la différence quadratique moyenne entre les valeurs prédites et les valeurs réelles

dans les données d'entraînement. Cependant, dans les tâches de régression déséquilibrées, la MSE peut ne pas bien fonctionner lorsque la distribution des étiquettes est asymétrique, ce qui conduit à une performance sous-optimale sur les étiquettes rares.

La MSE équilibrée, comme proposée dans l'article, répond aux limitations de la MSE dans la régression déséquilibrée en incorporant la distribution des étiquettes d'entraînement avant d'obtenir une prédiction plus équilibrée et d'améliorer les performances sur les étiquettes rares. Elle offre une approche statistiquement fondée pour gérer efficacement les tâches de régression déséquilibrées.[25]

II.7.2 Mean Absolute Error (MAE)

L'Erreur Absolue Moyenne (EAM) ou MAE (Mean Absolute Error) est une mesure fréquemment utilisée dans l'analyse de régression pour évaluer la performance d'un modèle. La MAE mesure la magnitude moyenne des erreurs entre les valeurs prédites et les valeurs réelles. Elle est calculée en prenant la moyenne des différences absolues entre les valeurs prédites et les valeurs réelles pour chaque point de données.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (II.16)$$

La MAE est préférée lorsque les valeurs aberrantes dans les données ne sont pas nécessairement des erreurs et ne devraient pas être fortement pénalisées. Elle fournit une mesure d'erreur plus robuste par rapport à d'autres mesures comme la MSE car elle ne met pas les erreurs au carré, la rendant moins sensible aux grandes erreurs.[26]

II.7.3 Root Mean Square Error (RMSE)

L'Erreur Quadratique Moyenne Racine (EQMR) ou Root Mean Square Error (RMSE) est une autre mesure fréquemment utilisée dans l'analyse de régression pour évaluer la performance d'un modèle. La RMSE mesure la racine carrée de la moyenne des différences au carré entre les valeurs

prédites et les valeurs réelles. Elle fournit une mesure de la dispersion des erreurs dans les prédictions.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (II.17)$$

La RMSE est utile lorsque vous voulez pénaliser plus lourdement les erreurs plus grandes que les petites erreurs. Elle est sensible aux valeurs aberrantes dans les données car elle met les erreurs au carré avant de les moyenner, donnant plus de poids aux grandes erreurs.[26]

II.8 Conclusion

Ce chapitre fournit une analyse approfondie des réseaux de neurones récurrents (RNNs) et de leurs variantes, y compris la mémoire à long terme (LSTM), la LSTM bidirectionnelle (BiLSTM) et les unités récurrentes synchronisées (GRU). Ces modèles ont montré un potentiel significatif pour prédire les rendements agricoles en raison de leur capacité à capturer les dépendances temporelles dans les données séquentielles. Le BiLSTM étend les capacités du LSTM en traitant les données dans les directions avant et arrière. Le modèle GRU simplifie l'architecture LSTM tout en maintenant des performances comparables. Le chapitre explore également les hyperparamètres impliqués dans la construction de ces modèles, tels que le nombre d'unités par couche, les couches cachées, les époques et le taux d'apprentissage, qui jouent tous un rôle crucial dans la performance du modèle. Cependant, l'efficacité de ces modèles dépend de la qualité et de la quantité des données, de la complexité du problème et du réglage minutieux des hyperparamètres du modèle. Par conséquent, une compréhension et une expérimentation approfondies sont cruciales pour une application réussie dans la prévision du rendement agricole.

CHAPITRE III : Système proposé pour la prédiction des rendements agricoles

III.1 Introduction

Dans ce chapitre, nous proposons un système de prédiction des rendements agricoles ainsi que la collecte de données, spécifiquement pour la région d'Annaba. Nous visons à compiler un ensemble de données solide, comprenant les « températures moyennes, maximales et minimales », ainsi que les « taux de précipitations ».

Nous analyserons et comparerons ensuite trois types de modèles de Réseaux Neuronaux Récurrents (RNN) : le « RNN » traditionnel, la « Mémoire à Long Terme (LSTM) », la « LSTM Bidirectionnelle (BiLSTM) », et l' « Unité Récurrente à Portes (GRU) ». Enfin, nous étudierons l' « impact des hyperparamètres » sur ces modèles.

L'objectif est de comprendre comment la collecte de données, le prétraitement et les modèles d'apprentissage automatique sophistiqués peuvent se combiner pour résoudre des tâches de prédiction complexes.

III.2 Les principales étapes du système de prédiction proposé

- Acquisition de données : La première étape consiste à compiler des données historiques qui englobent les variables souhaitées : température minimale, maximale et moyenne, précipitations et rendements des cultures.
- Nettoyage des données : L'étape suivante consiste à prétraiter les données en gérant les valeurs manquantes et les valeurs aberrantes. Il peut également être nécessaire de normaliser ou de standardiser les données. Cette étape est essentielle car la qualité des données peut affecter significativement la performance du modèle.
- Entraînement du modèle : Les données sont divisées en un ensemble d'entraînement et un ensemble de test. L'ensemble d'entraînement est utilisé pour entraîner notre modèle (basé sur RNN, LSTM, BiLSTM ou GRU). Les caractéristiques sont ensuite introduites (données de température et de précipitations) dans le modèle pour prédire la variable cible (rendement des cultures).

- Test et validation : Un ensemble de test (représentant 30% de l'ensemble de la base de données) est utilisé pour évaluer la performance de notre modèle. Cela nous donnera une idée de la manière dont notre modèle est susceptible de se comporter sur des données non vues (n'appartenant pas à l'ensemble d'entraînement).

Affinement du modèle : En fonction des résultats de test obtenues, des ajustements peuvent être apporté sur notre modèle ou ses paramètres pour améliorer les performances. Cela pourrait impliquer de modifier la structure du modèle, d'ajuster le taux d'apprentissage, ou même de collecter plus de données.

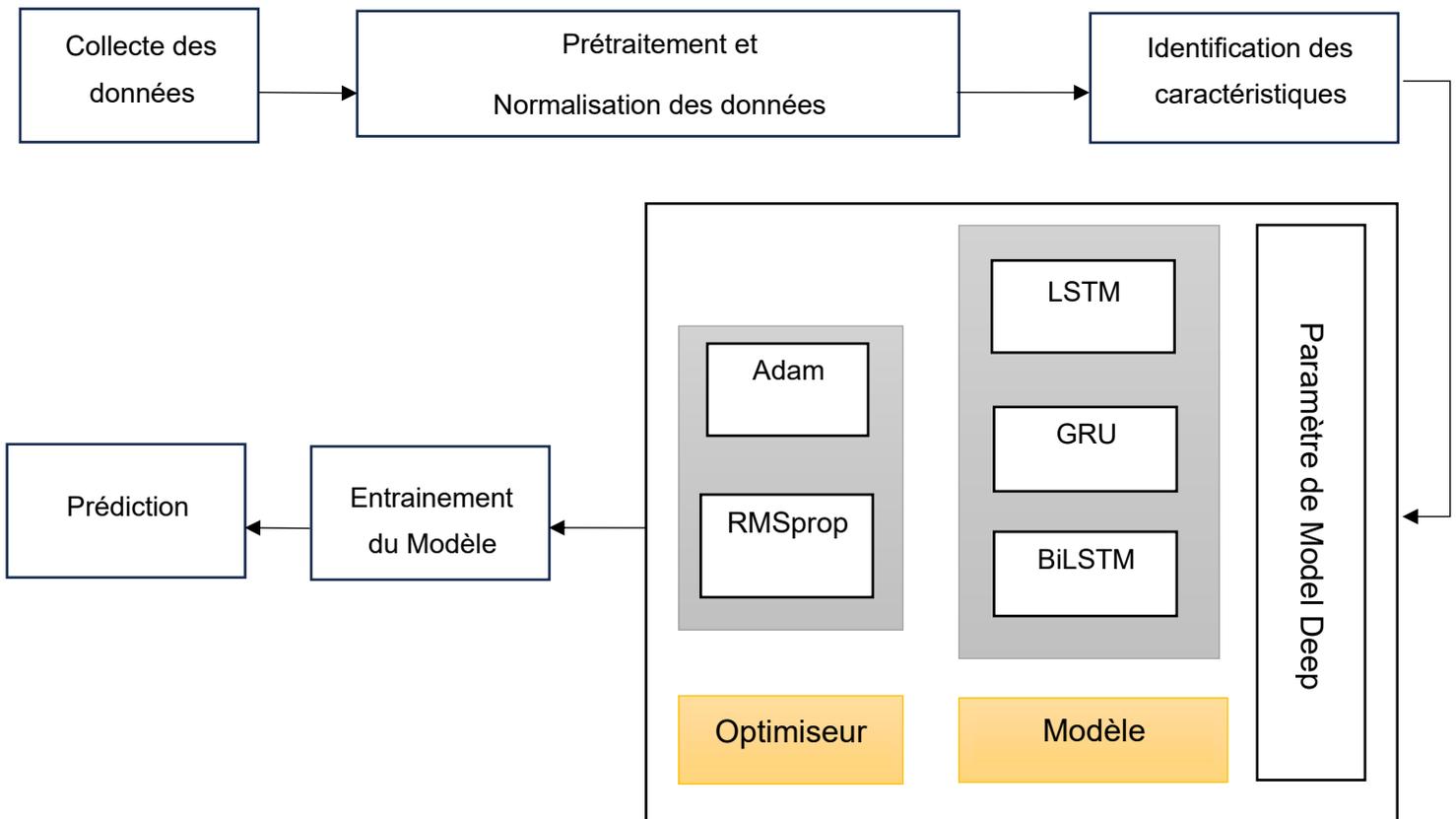


Figure III.1. Schéma d'ensemble des étapes du système de prédiction

III.3 Outils et matériels utilisés

Les prédictions ont été réalisées à l'aide d'un système informatique composé d'un processeur Intel(R) Core (TM) i5-6300U, d'une carte graphique Intel(R) HD Graphics 520 et de 12 Go de RAM.

Anaconda est une distribution des langages de programmation Python pour le calcul scientifique qui vise à simplifier la gestion et le déploiement des paquets.

« TensorFlow » est une bibliothèque logicielle « open source » pour les applications numériques avec des graphes computationnels. Elle a été développée directement par l'équipe « Google Brain » pour l'apprentissage automatique et les réseaux neuronaux profonds.

« Keras » est une bibliothèque de réseaux neuronaux profonds écrite en Python et fonctionnant comme une interface en « TensorFlow » ou « Theano ». Elle a été développée pour permettre une expérimentation rapide. L'environnement Anaconda a été utilisé pour mettre en œuvre le modèle en utilisant la version 3.11.5 de Python et le framework Tensorflow GPU 2.10.0, ainsi que la version 2.10.0 de la bibliothèque « Keras ».

III.4 Collecte de données

Ces informations relatives au climat d'Annaba ont été recueillies d'un site spécialisé dans la surveillance des conditions météorologiques depuis 1942¹, incluant les températures minimales et maximales, les précipitations, etc. En ce qui concerne les données agricoles, elles ont été obtenues auprès de la Direction générale de l'agriculture d'Annaba (DGAA), représentant la superficie, la production et les revenus agricoles pour les céréales sur une période allant de 2012 à 2023.

¹ <https://fr.tutiempo.net/>

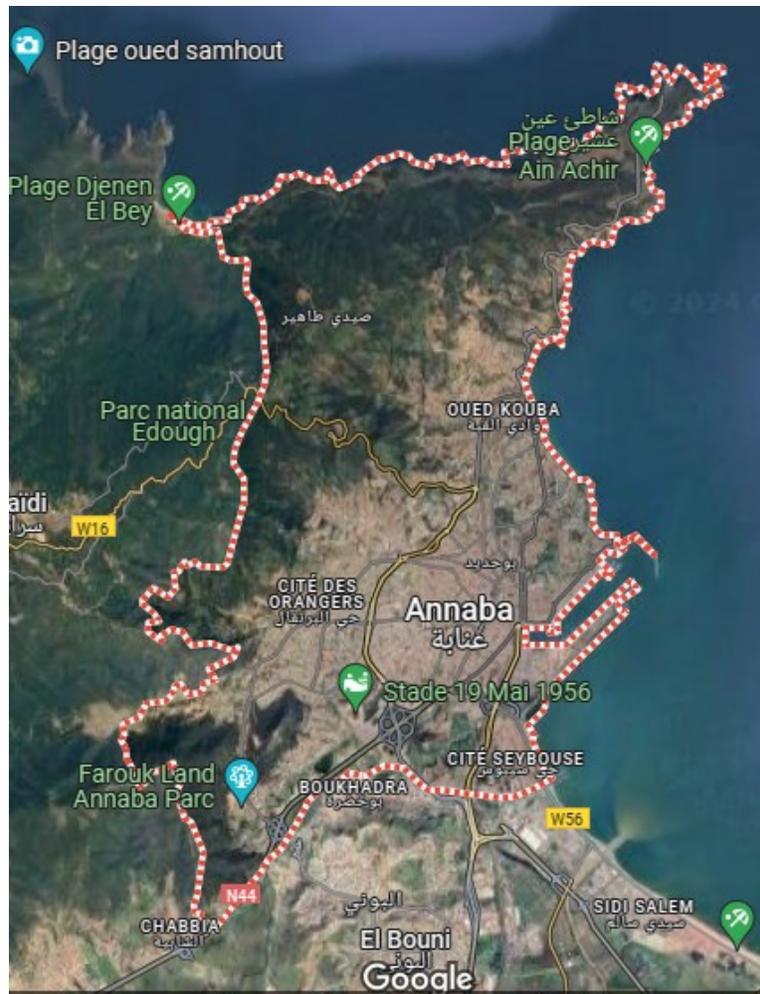


Figure III.2. Carte de la région d'Annaba

Tableau III.1. Details de localisation

Localisation	Longitude	latitude	Année de début	Année de finie	Superficie
ANNABA	36°54'49"N	7°44'40"E	2012	2023	49 Km ²

Notre base de données regroupe des informations climatiques mensuelles pour chaque saison. Ces données incluent la température moyenne annuelle (T), la température maximale moyenne annuelle (TM), la température minimale moyenne annuelle (Tm), la précipitation totale annuelle de pluie et/ou de neige fondue en mm (PP), la vitesse moyenne annuelle du vent en Km/h (V), le nombre total de jours de pluie durant l'année (RA), le nombre total de jours de neige durant l'année

(SN) (Voir La **Figure III.3**). Les données agricoles comprennent les superficies dédiées à la culture du blé ainsi que les productions et les rendements de cette céréale (Voir La **Figure III.4**).

Tableau III.2. *Détails sur la base de données*

Variables	Source de données	Unité	Max	Min	Moy
T	Tutie po	°C	24	4.3	14.15
TM	Tutie po	°C	22	8	15
Tm	Tutie po	°C	10.9	0	5.45
PP	Tutie po	mm	39.88	0	19.94
RA	Tutie po	jour	27	0	0.5
V	Tutie po	Km/h	34.8	8.3	21.55
SN	Tutie po	Jour	13	0	0.5
Superficies blé	DGAA	Ha	16511,5	13087	14.799
Production blé	DGAA	Qx	480780	233270	357.025
Rendement blé	DGAA	%	33.8	16.3	25.05

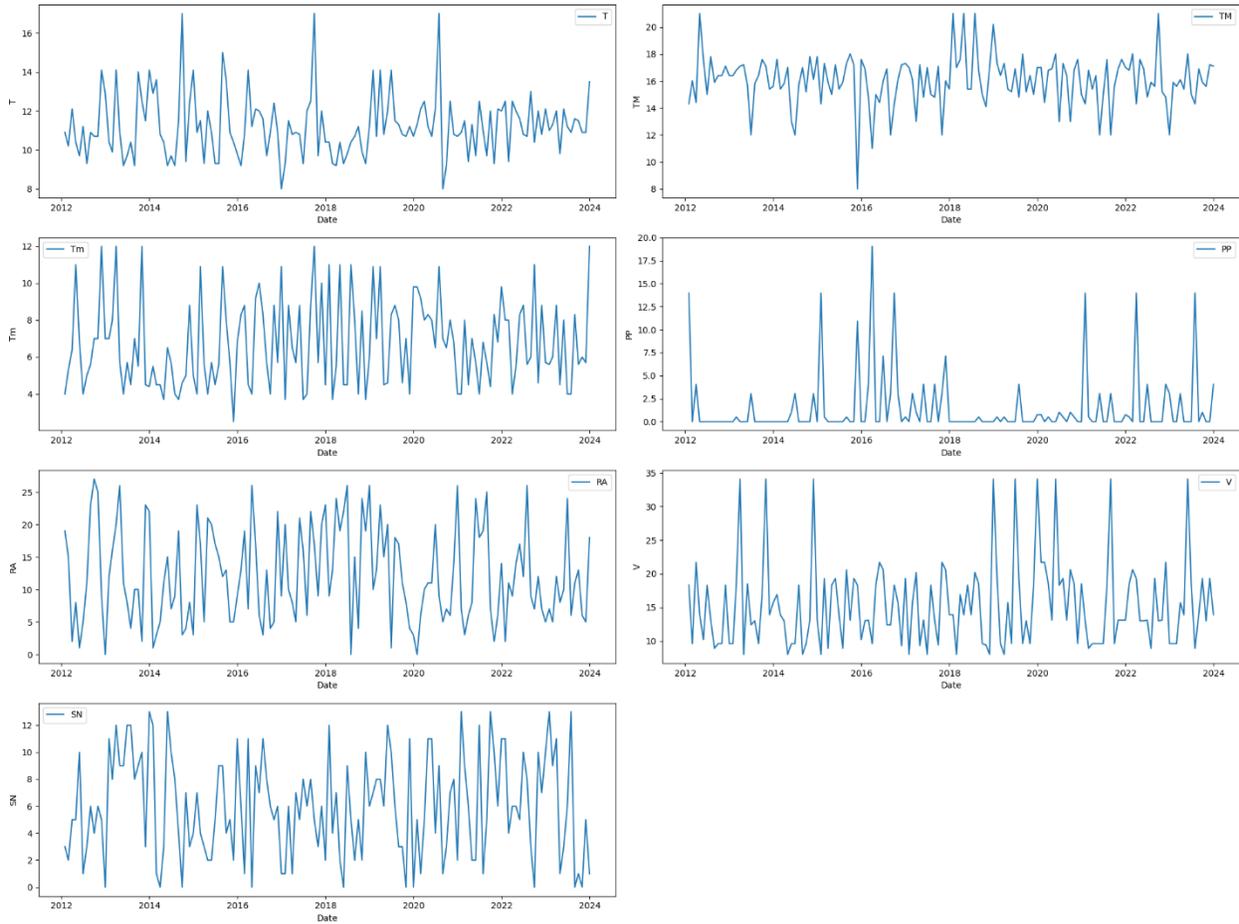


Figure III.3. Ensemble de données climatiques de Annaba

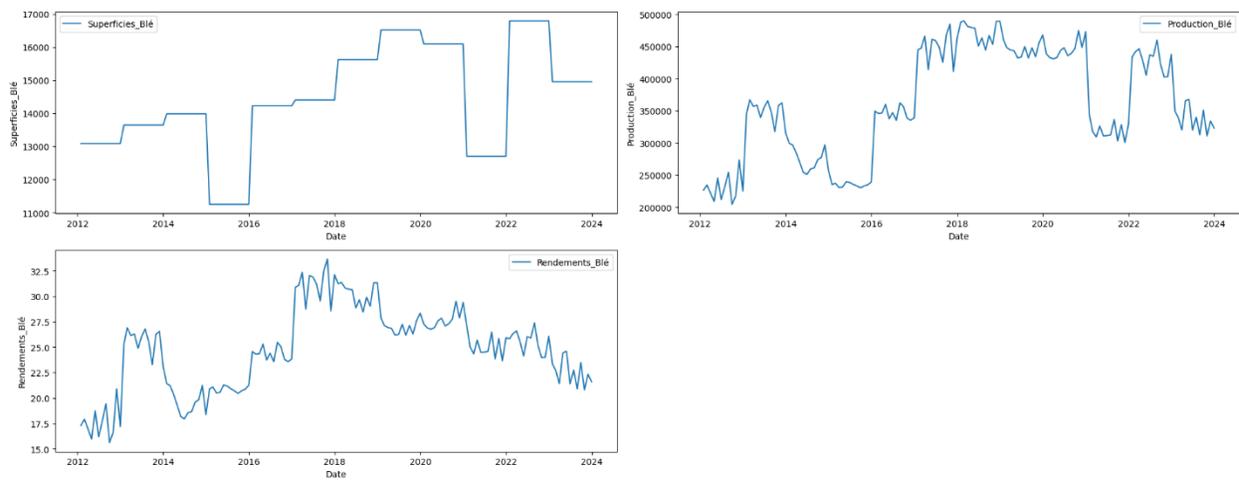


Figure III.4. Ensemble de données agricoles de Annaba

III.5 Entraînement du modèle

Les données ont été divisées en un ensemble d'entraînement, un ensemble de validation et un ensemble de test. Les ensembles de données d'entraînement et de test sont utilisés pour entraîner le modèle et évaluer les performances du réseau formé, respectivement. L'ensemble de test est ignoré pendant l'entraînement et la sélection des hyperparamètres, et ses erreurs déterminent la capacité du modèle à se généraliser aux données non vues. Les données de validation sont utilisées pour comparer différents modèles, sélectionner des hyperparamètres et prévenir le surapprentissage pendant la phase d'entraînement.

Dans ce travail, 30% des données ont été sélectionnées comme ensemble de test, et le reste a été divisé en 60-10% comme ensembles d'entraînement et de validation, respectivement. Le modèle de prédiction du blé a été formé sur 7 échantillons (saison) et testé sur 3 échantillons.

Lors de l'entraînement d'un réseau neuronal, certaines valeurs spécifiques doivent être initialisées, que l'on appelle hyperparamètres. Les hyperparamètres (tels que le nombre de couches cachées, la fonction d'activation, etc.) déterminent la structure d'un modèle et les hyperparamètres (tels que le taux d'apprentissage, le temps de déclin, etc.) déterminent comment le modèle est formé.

Les résultats de plusieurs simulations du même algorithme avec différents hyperparamètres varieront. Les hyperparamètres considérés dans ce travail sont le nombre de couches, le nombre d'unités cachées, la taille du lot, la taille de l'abandon, le taux d'apprentissage le déclin du taux d'apprentissage et le nombre d'époques.

Quatre modèles distincts ont été sélectionnés pour notre étude, à savoir LSTM, BiLSTM, GRU et RNN. Chacun de ces modèles a été élaboré en deux configurations différentes : une avec deux couches et l'autre avec trois couches, afin de permettre une comparaison exhaustive. De plus, nous avons apporté plusieurs ajustements au niveau des hyperparamètres.

III.6 Résultats et discussion

Afin de déterminer lequel parmi les quatre modèles est le plus robuste et efficace pour notre base de données, nous allons effectuer des modifications au niveau des hyperparamètres. Nous avons défini des valeurs différentes pour le dropout, le taux d'apprentissage initial, le taux

d'apprentissage et la décroissance afin de comprendre leur impact sur nos modèles. De plus, nous testerons les quatre modèles une fois avec deux couches et une fois avec trois couches, pour comprendre l'effet du nombre de couches sur la qualité des modèles. Nous avons également ajouté une couche de dropout, une couche dense et une couche de normalisation par lots après chaque couche dans les modèles.

a. Résultats du modèle RNN

Tableau III.3. Résultats du modèle RNN utilisant l'optimiseur Adam en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	4.88	4.90	4.99	5.13	4.96	4.90	4.98	5.14
	0.2	5.71	5.06	4.94	4.99	5.71	5.88	5.66	5.34
	0.4	4.96	5.13	5.03	5.31	5.45	5.39	5.36	5.31
Initial Learning Rate	10^{-2}	4.79	5.03	5.43	5.30	5.81	5.03	5.58	5.30
	10^{-3}	4.71	4.97	5.13	5.28	5.71	5.23	5.91	5.28
	10^{-4}	12.87	8.88	5.53	5.23	13.07	8.77	5.90	5.95
Learning Rate	10^{-3}	4.63	4.41	5.43	5.18	5.63	4.41	5.34	5.18
	10^{-4}	5.04	4.79	5.26	6.13	5.04	4.79	6.23	6.13
	10^{-5}	4.93	4.44	5.34	6.05	5.35	4.44	6.12	5.05
Decay	0.25	5.71	5.13	5.31	4.96	5.71	5.13	5.45	5.96
	0.50	5.28	4.98	5.22	4.88	5.28	4.98	5.33	5.88
	0.75	4.90	4.92	4.96	5.70	4.89	5.33	5.15	5.70

Le **Tableau III.3** représente les performances du modèle de prédiction basé sur RNN en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.3**, la meilleure performance du modèle basé sur RNN a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0, Initial Learning Rate= 10^{-3} , Learning Rate= 10^{-3} et Decay=0.75.

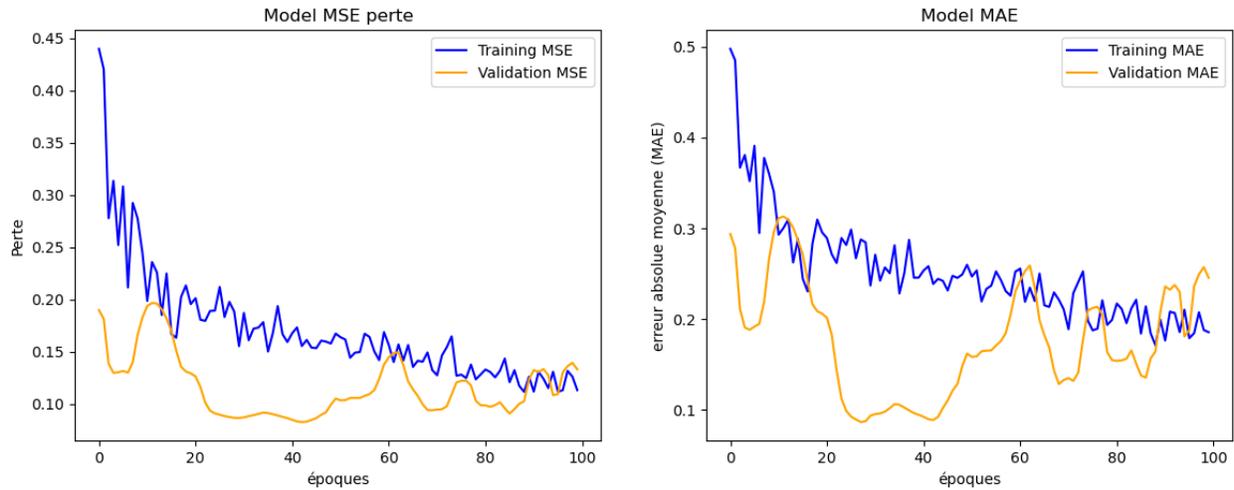


Figure III.5. *Modèle de perte pendant l'apprentissage de modèle RNN (Adam)*

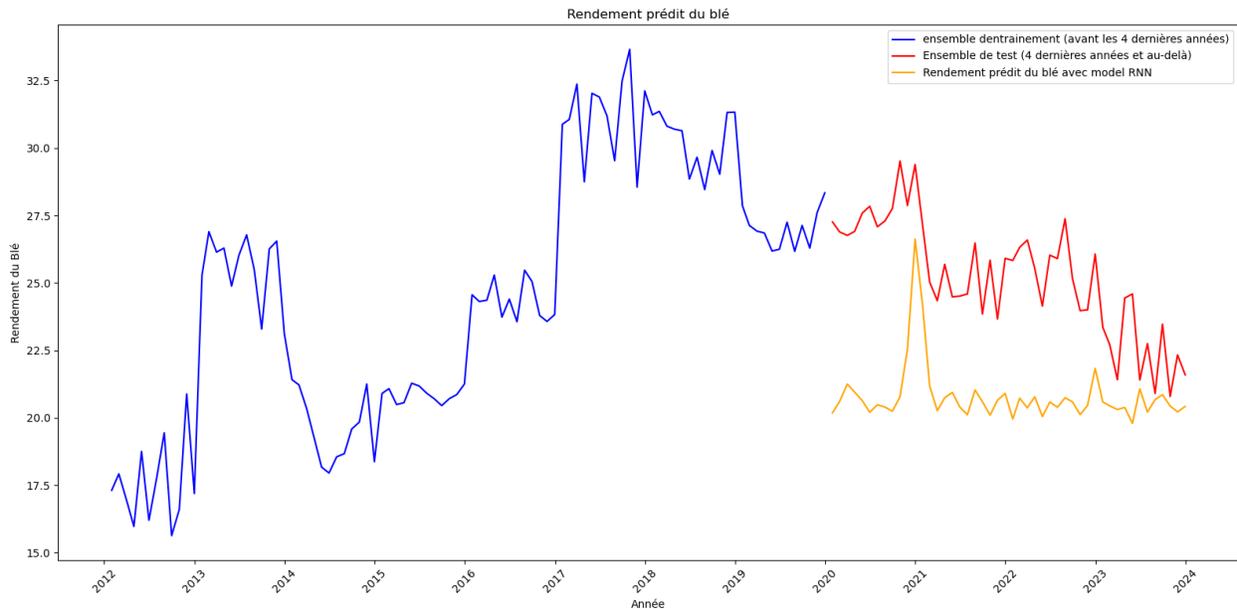


Figure III.6. *Résultat de prédiction des rendements du blé basé sur le modèle RNN utilisant l'optimiseur Adam*

Tableau III.4. Résultats du modèle RNN utilisant l'optimiseur RMSprop en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	4.44	3.87	4.04	3.80	4.44	3.87	4.04	3.80
	0.2	5.71	4.88	4.66	4.34	5.71	4.88	4.66	4.34
	0.4	4.41	4.39	4.36	4.31	4.41	4.39	4.36	4.31
Initial Learning Rate	10^{-2}	5.79	5.03	4.58	4.30	5.79	5.03	4.58	4.30
	10^{-3}	5.71	5.23	4.91	4.28	5.71	5.23	4.91	4.28
	10^{-4}	12.87	8.88	5.87	4.23	12.87	8.88	5.87	4.23
Learning Rate	10^{-3}	4.63	4.41	4.34	4.18	4.63	4.41	4.34	4.18
	10^{-4}	5.04	4.79	4.23	4.13	5.04	4.79	4.23	4.13
	10^{-5}	4.93	4.44	4.12	4.05	4.93	4.44	4.12	4.05
Decay	0.25	5.71	5.13	4.45	3.96	5.71	5.13	4.45	3.96
	0.50	5.28	4.98	4.33	3.88	5.28	4.98	4.33	3.88
	0.75	4.50	4.33	4.15	3.70	3.23	4.33	4.15	3.70

Le **Tableau III.4** représente les performances du modèle de prédiction basé sur RNN en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.4**, la meilleure performance du modèle basé sur RNN a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-4} et Decay=0.75.

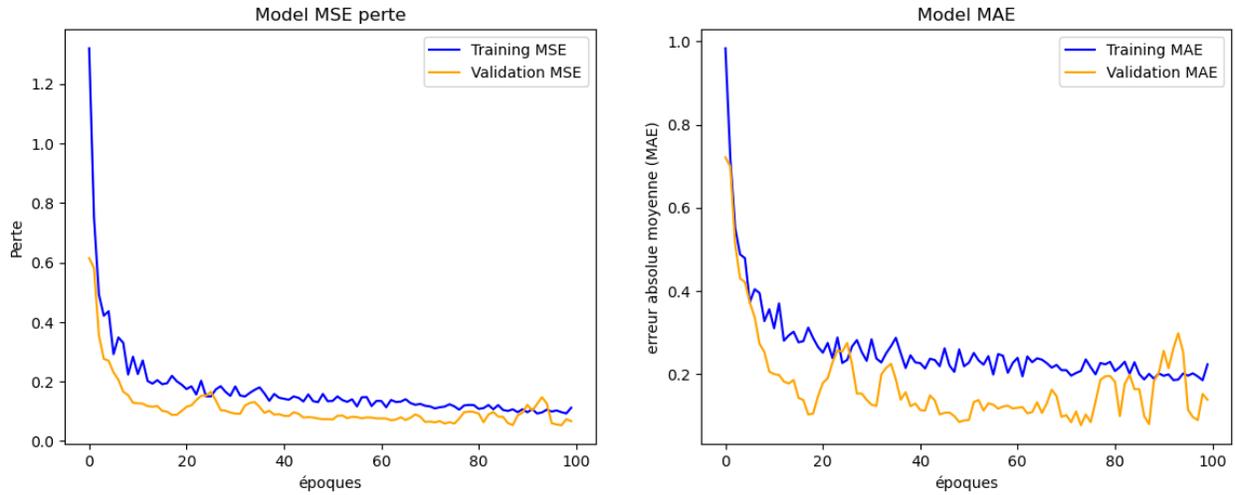


Figure III.7. Modèle de perte pendant l'apprentissage de modèle RNN (RMSprop)

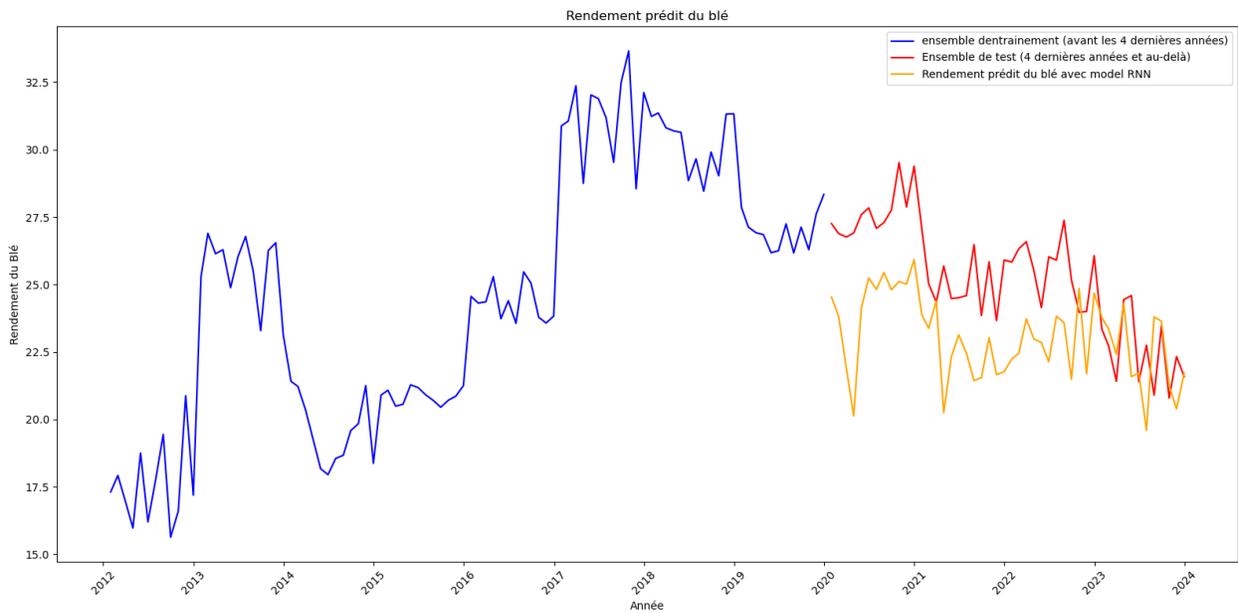


Figure III.8. Résultat de prédiction des rendements du blé basé sur le modèle RNN utilisant l'optimiseur RMSprop

b. Résultats du modèle LSTM

Tableau III.5. Résultats du modèle LSTM utilisant l'optimiseur Adam en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	4.55	4.39	4.36	4.34	3.34	3.32	3.30	3.26
	0.2	4.41	4.35	4.30	4.21	3.31	3.26	3.24	3.20
	0.4	4.39	4.29	4.28	4.18	3.23	3.22	3.21	3.19
Initial Learning Rate	10^{-2}	4.34	4.25	4.21	4.13	3.20	3.16	3.14	3.13
	10^{-3}	4.29	4.21	4.18	4.07	3.19	3.15	3.06	3.02
	10^{-4}	4.24	4.18	4.16	4.04	3.13	3.14	2.99	2.98
Learning Rate	10^{-3}	4.21	4.13	4.07	3.94	3.09	3.07	2.94	2.91
	10^{-4}	4.18	4.09	4.01	3.88	3.02	2.99	2.88	2.79
	10^{-5}	4.13	4.06	3.96	3.84	2.99	2.91	2.80	2.84
Decay	0.25	4.09	4.01	3.88	3.79	2.90	2.88	2.77	2.70
	0.50	4.01	3.84	3.70	3.67	2.86	2.79	2.70	2.66
	0.75	3.84	3.77	3.66	3.55	2.79	3.75	2.69	2.61

Le **Tableau III.5** représente les performances du modèle de prédiction basé sur LSTM en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.5**, la meilleure performance du modèle basé sur LSTM a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-4} et Decay=0.75.

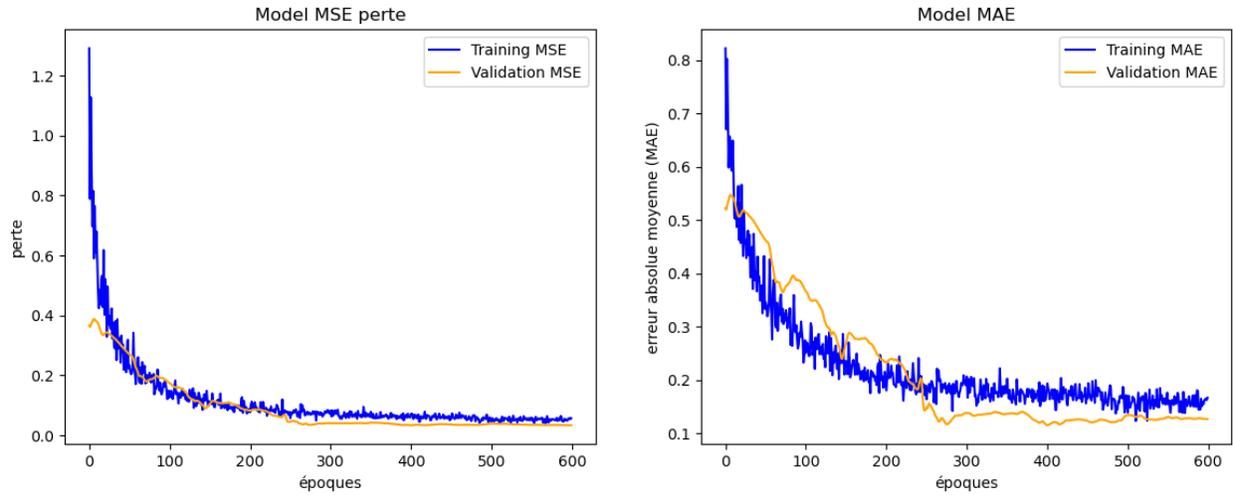


Figure III.9. *Modèle de perte pendant l'apprentissage de modèle LSTM (Adam)*

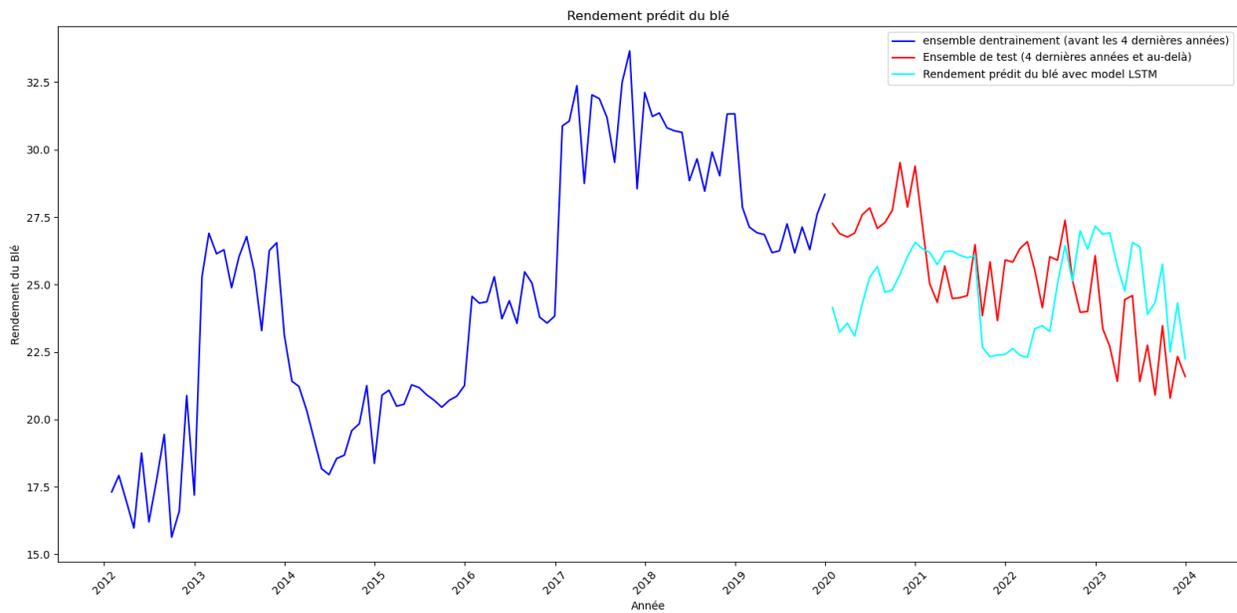


Figure III.10. *Résultat de prédiction des rendements du blé basé sur le modèle LSTM utilisant l'optimiseur Adam*

Tableau III.6. Résultats du modèle LSTM utilisant l'optimiseur RMSprop en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	3.56	3.33	3.21	3.13	2.98	2.77	2.70	2.65
	0.2	3.48	3.27	3.20	3.08	2.88	2.65	2.66	2.61
	0.4	3.40	3.14	3.03	2.98	2.81	2.63	2.63	2.54
Initial Learning Rate	10^{-2}	3.34	3.03	2.96	2.94	2.77	2.59	2.59	2.49
	10^{-3}	3.12	2.96	2.93	2.90	2.70	2.55	2.51	2.44
	10^{-4}	3.07	2.90	2.85	2.86	2.66	2.51	2.49	2.40
Learning Rate	10^{-3}	2.99	2.86	2.80	2.79	2.60	2.49	2.44	2.36
	10^{-4}	2.90	2.80	2.77	2.74	2.52	2.44	2.39	2.34
	10^{-5}	2,83	2.79	2.71	2.66	2.46	2.41	2.35	2.29
Decay	0.25	2.74	2.66	2.62	2.59	2.41	2.36	2.30	2.23
	0.50	2.70	2.62	2.55	2.51	2.37	2.30	2.24	2.20
	0.75	2.66	2.61	2.54	2.49	2.31	2.29	2.20	2.14

Le **Tableau III.6** représente les performances du modèle de prédiction basé sur LSTM en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.6**, la meilleure performance du modèle basé sur LSTM a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-5} et Decay=0.75.

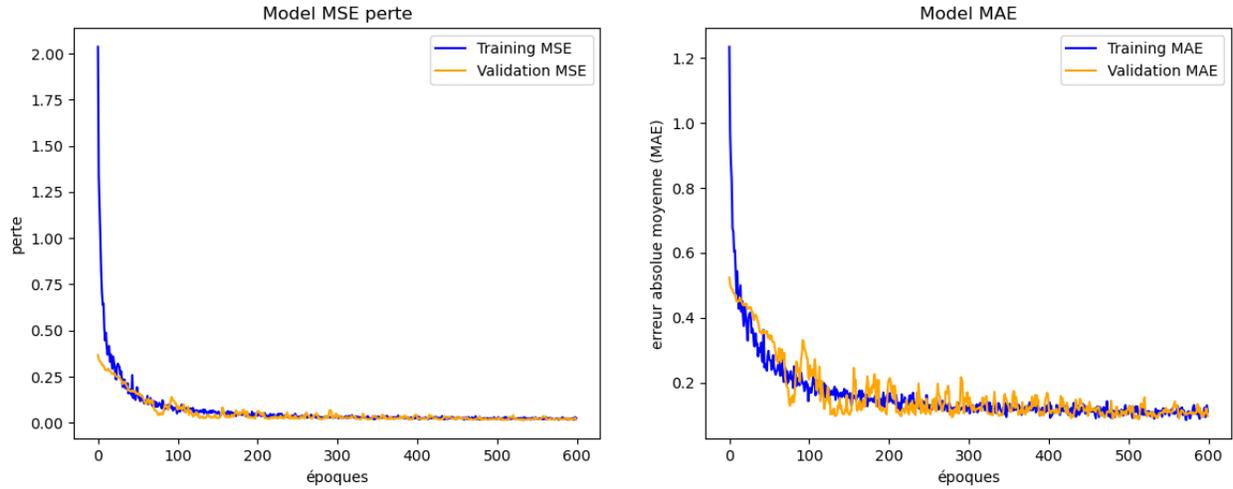


Figure III.11. Modèle de perte pendant l'apprentissage de modèle LSTM (RMSprop)

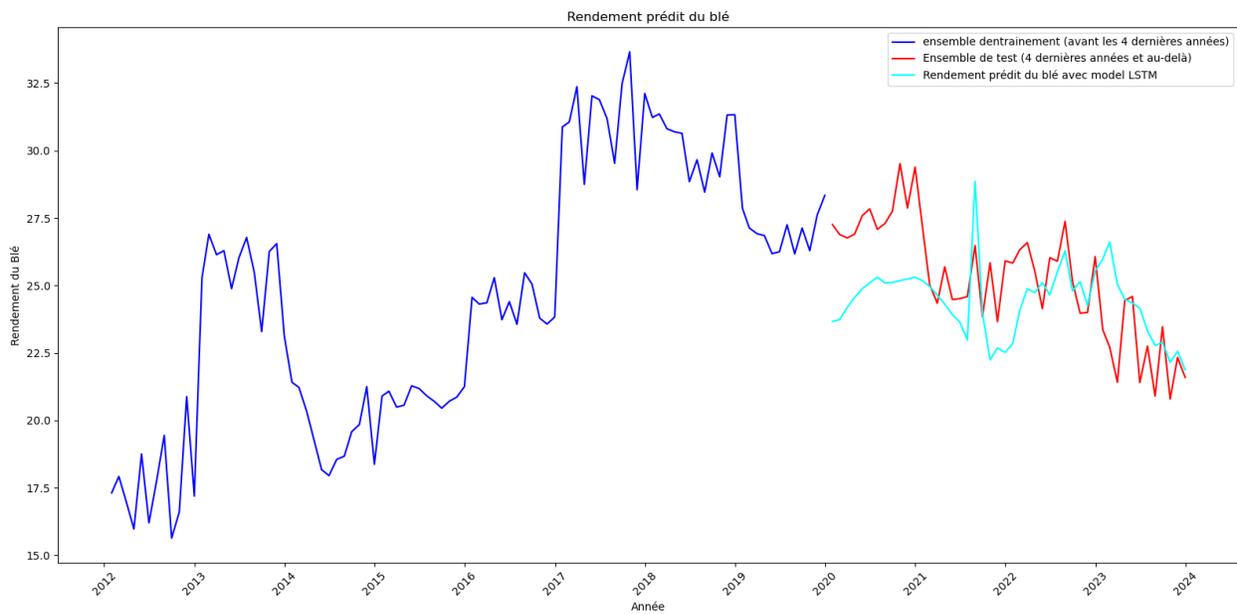


Figure III.12. Résultat de prédiction des rendements du blé basé sur le modèle LSTM utilisant l'optimiseur RMSprop

c. Résultats du modèle GRU

Tableau III.7. Résultats du modèle GRU utilisant l'optimiseur Adam en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	3.79	3.61	3.55	3.50	3.31	3.26	3.20	3.16
	0.2	3.71	3.59	3.54	3.47	3.30	3.23	3.17	3.12
	0.4	3.64	3.54	3.53	3.41	3.29	3.19	3.14	3.08
Initial Learning Rate	10^{-2}	3.59	3.49	3.46	3.39	3.27	3.14	3.12	3.04
	10^{-3}	3.53	3.45	3.44	3.33	3.20	3.09	3.04	3.01
	10^{-4}	3.44	3.41	3.39	3.30	3.19	3.02	2.94	2.96
Learning Rate	10^{-3}	3.41	3.39	3.33	3.29	3.14	2.98	2.92	2.92
	10^{-4}	3.37	3.33	3.28	3.21	3.08	2.94	2.88	2.86
	10^{-5}	3.31	3.28	3.23	3.19	3.05	2.93	2.85	2.83
Decay	0.25	3.27	3.23	3.16	3.15	2.99	2.90	2.80	2.77
	0.50	3.21	3.20	3.09	3.06	2.90	2.80	2.75	2.70
	0.75	3.19	3.15	3.03	2.95	2.82	2.77	2.70	2.69

Le **Tableau III.7** représente les performances du modèle de prédiction basé sur GRU en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.7**, la meilleure performance du modèle basé sur GRU a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-5} et Decay=0.75.

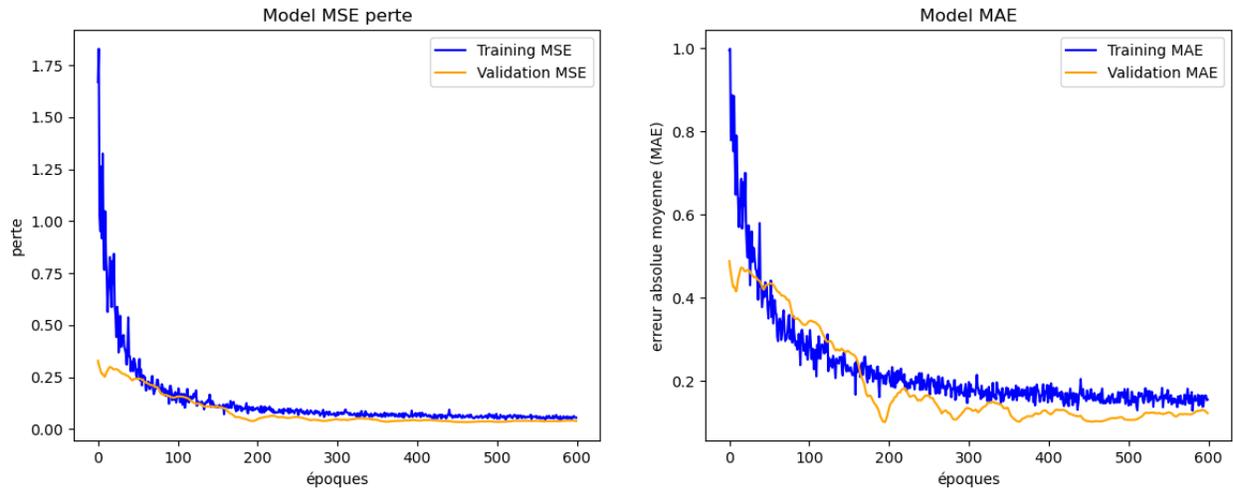


Figure III.13. Modèle de perte pendant l'apprentissage de modèle GRU (Adam)

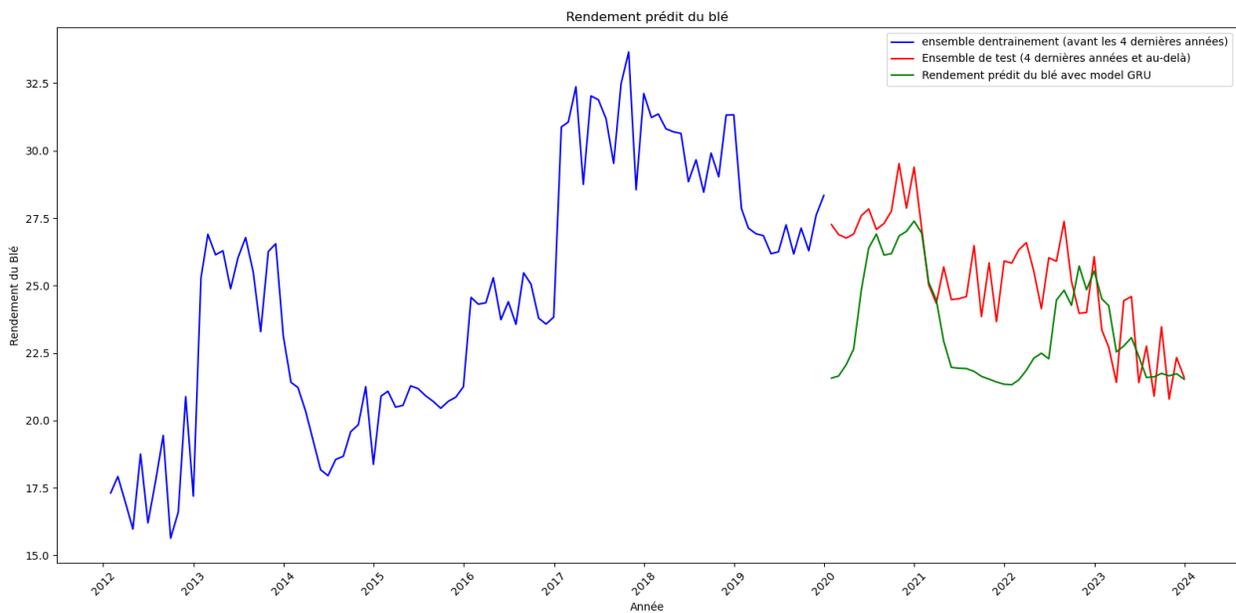


Figure III.14. Résultat de prédiction des rendements du blé basé sur le modèle GRU utilisant l'optimiseur Adam

Tableau III.8. Résultats du modèle GRU utilisant l'optimiseur RMSprop en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	3.57	3.51	3.45	3.41	3.24	3.15	3.01	2.99
	0.2	3.52	3.49	3.44	3.38	3.22	3.04	2.98	2.96
	0.4	3.46	3.44	3.39	3.32	3.19	2.97	2.94	2.88
Initial Learning Rate	10^{-2}	3.42	3.41	3.31	3.27	3.14	2.94	2.90	2.82
	10^{-3}	3.39	3.37	3.27	3.21	3.02	2.89	2.84	2.76
	10^{-4}	3.33	3.32	3.23	3.15	2.95	2.83	2.77	2.74
Learning Rate	10^{-3}	3.29	3.28	3.19	3.13	2.90	2.77	2.71	2.69
	10^{-4}	3.24	3.21	3.13	3.06	2.84	2.74	2.69	2.61
	10^{-5}	3.19	3.16	3.08	3.01	2.77	2.68	2.62	2.58
Decay	0.25	3.13	3.09	3.01	2.95	2.71	2.62	2.55	2.51
	0.50	3.08	3.01	2.97	2.90	2.63	2.55	2.51	2.49
	0.75	3.01	2.96	2.90	2.86	2.57	2.51	2.46	2.41

Le **Tableau III.8** représente les performances du modèle de prédiction basé sur GRU en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.8**, la meilleure performance du modèle basé sur GRU a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-5} et Decay=0.75.

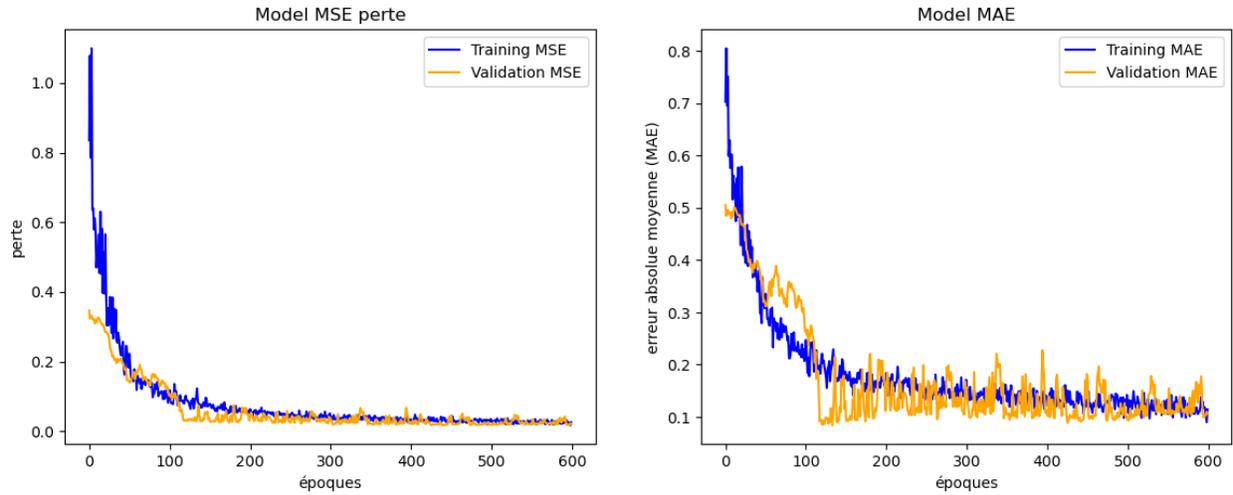


Figure III.15. *Modèle de perte pendant l'apprentissage de modèle GRU (RMSprop)*

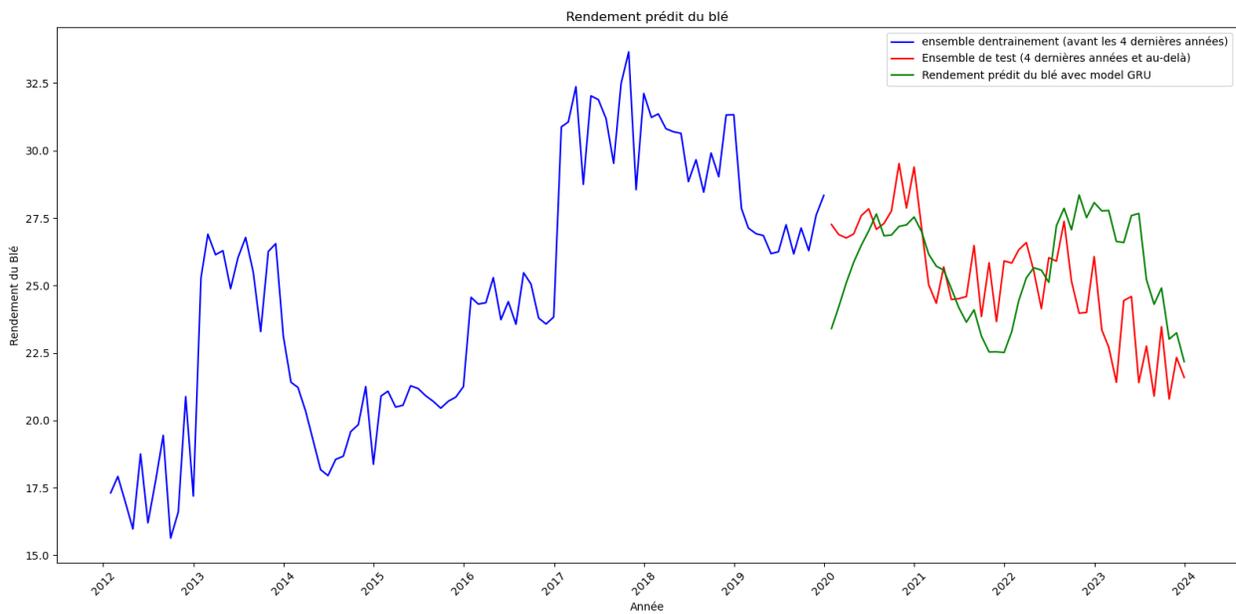


Figure III.16. *Résultat de prédiction des rendements du blé basé sur le modèle GRU utilisant l'optimiseur RMSprop*

d. Résultats du modèle BiLSTM

Tableau III.9. Résultats du modèle BiLSTM utilisant l'optimiseur Adam en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	3.44	3.39	3.33	3.29	2.93	2.90	2.88	2.85
	0.2	3.42	3.33	3.30	3.24	2.91	2.88	2.87	2.84
	0.4	3.37	3.30	3.29	3.19	2.90	2.86	2.83	2.80
Initial Learning Rate	10^{-2}	3.33	3.25	3.21	3.15	2.89	2.83	2.80	2.79
	10^{-3}	3.29	3.21	3.18	3.10	2.88	2.80	2.78	2.74
	10^{-4}	3.24	3.19	3.11	3.05	2.83	2.79	2.74	2.71
Learning Rate	10^{-3}	3.19	3.15	3.06	2.99	2.80	2.77	2.70	2.66
	10^{-4}	3.11	3.11	3.01	2.94	2.77	2.74	2.69	2.64
	10^{-5}	3.07	3.04	2.99	2.90	2.72	2.70	2.66	2.61
Decay	0.25	3.00	2.98	2.91	2.86	2.69	2.67	2.60	2.58
	0.50	2.99	2.91	2.88	2.80	2.63	2.60	2.55	2.52
	0.75	2.98	2.88	2.80	2.73	2.60	2.59	2.51	2.49

Le **Tableau III.9** représente les performances du modèle de prédiction basé sur BiLSTM en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.9**, la meilleure performance du modèle basé sur BiLSTM a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-5} et Decay=0.75.

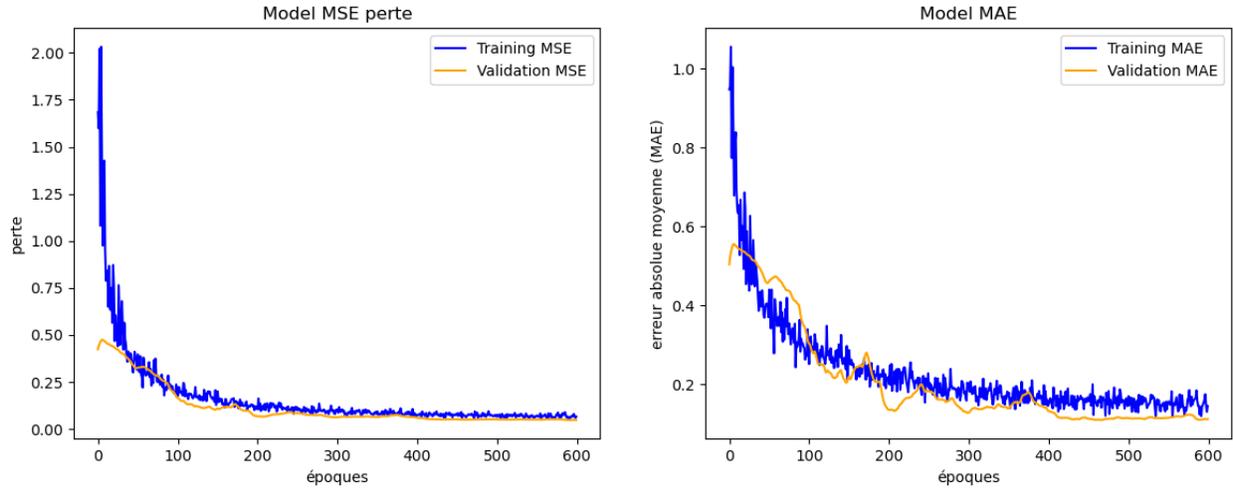


Figure III.17. *Modèle de perte pendant l'apprentissage de modèle BiLSTM (Adam)*

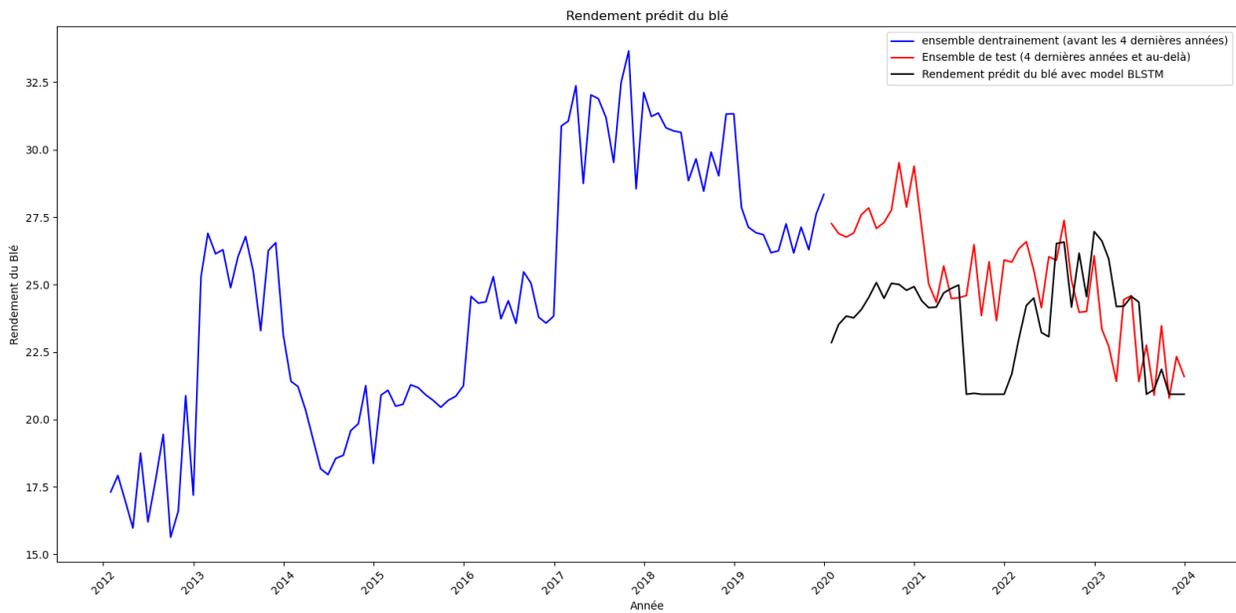


Figure III.18. *Résultat de prédiction des rendements du blé basé sur le modèle BiLSTM utilisant l'optimiseur Adam*

Tableau III.10. Résultats du modèle BiLSTM utilisant l'optimiseur RMSprop en terme de RMSE

Paramètre	Variables	2 couches				3 couches			
		Epoch				Epoch			
		100	300	500	600	100	300	500	600
Dropout	0	2.85	2.77	2.70	2.66	2.66	2.62	2.59	2.56
	0.2	2.82	2.66	2.69	2.62	2.64	2.60	2.58	2.53
	0.4	2.81	2.61	2.63	2.59	2.61	2.59	2.56	2.51
Initial Learning Rate	10^{-2}	2.78	2.59	2.60	2.55	2.59	2.56	2.53	2.49
	10^{-3}	2.76	2.51	2.56	2.50	2.55	2.53	2.49	2.46
	10^{-4}	2.74	2.49	2.53	2.48	2.52	2.49	2.46	2.41
Learning Rate	10^{-3}	2.66	2.45	2.49	2.41	2.47	2.45	2.41	2.36
	10^{-4}	2.61	2.42	2.44	2.39	2.41	2.40	2.36	2.32
	10^{-5}	2.53	2.48	2.41	2.34	2.39	2.36	2.31	2.28
Decay	0.25	2.48	2.46	2.35	2.30	2.37	2.31	2.24	2.21
	0.50	2.41	2.33	2.30	2.24	2.33	2.27	2.21	2.15
	0.75	2.39	2.31	2.24	2.20	2.30	2.23	2.15	2.10

Le **Tableau III.10** représente les performances du modèle de prédiction basé sur BiLSTM en termes de RMSE. Quatre (4) hyperparamètres ont été modifiés pour tester leur impact sur le modèle de prédiction. Adam a été choisi comme optimiseur tout au long de ces tests. D'après le **Tableau III.10**, la meilleure performance du modèle basé sur BiLSTM a été obtenue pour les valeurs des hyperparamètres suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-5} et Decay=0.75.

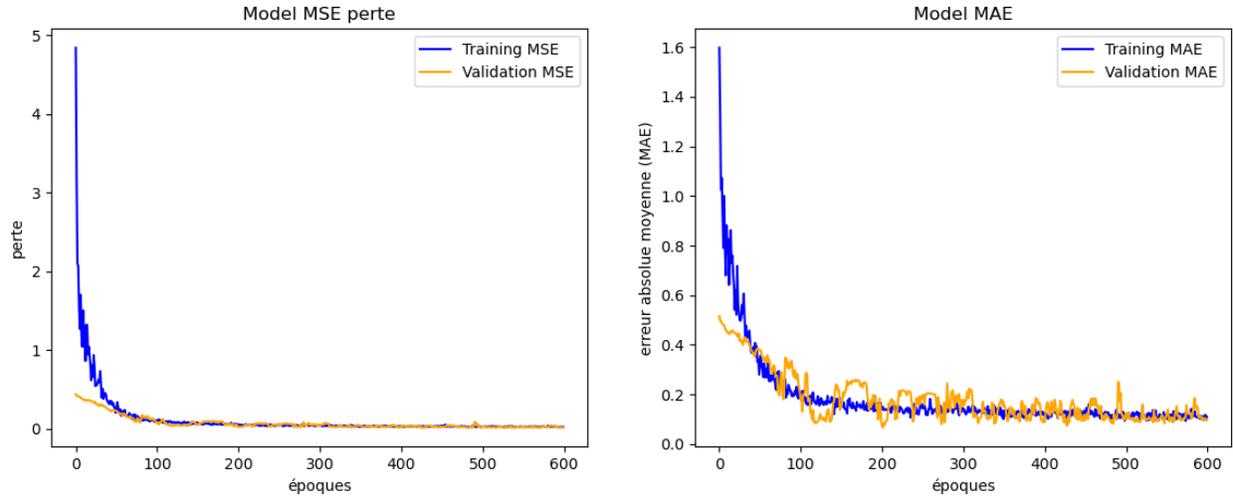


Figure III.19. Modèle de perte pendant l'apprentissage de modèle BiLSTM (RMSprop)

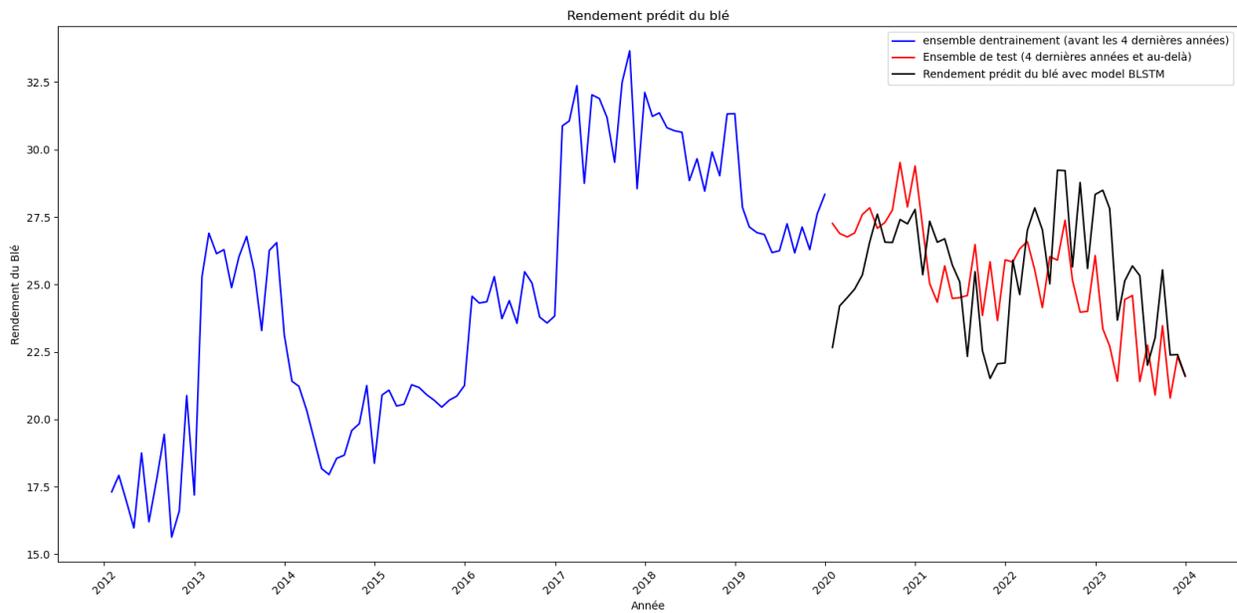


Figure III.20. Résultat de prédiction des rendements du blé basé sur le modèle BiLSTM utilisant l'optimiseur RMSprop

Le **Tableau III.11** illustre les valeurs de RMSE pour les quatre meilleurs résultats, qui ont été obtenus à partir de quatre modèles distincts, en se basant sur les résultats précédemment acquis.

Tableau III.11. *La valeur RMSE pour les meilleurs quatre résultats obtenus à partir des quatre modèles*

Le modèle	La valeur RMSE
BiLSTM RMSprop 3 Couches	2.10
LSTM RMSprop 3 Couches	2.14
GRU RMSprop 3 Couches	2.41
LSTM Adam 3 couches	2.61

En fin de compte, nous avons choisi un modèle BiLSTM à trois couches, avec 600 époques, pour la prédiction du rendement du blé. Les hyperparamètres optimaux pour ce modèle sont les suivants: Dropout=0.4, Initial Learning Rate= 10^{-4} , Learning Rate= 10^{-5} et Decay=0.75 , et l'optimiseur RMSprop.

L'ajout des couches Dense, Dropout, et Batch Normalization après les couches BiLSTM a significativement amélioré les résultats obtenus.

III.7 Conclusion

Dans ce chapitre, nous avons exploré comment un système de prédiction des rendements agricoles peut être mis en place pour la région d'Annaba. Nous avons compilé un ensemble de données robuste, comprenant des informations sur les températures et les précipitations.

Nous avons analysé et comparé trois types de modèles de réseaux neuronaux récurrents (RNN) : le RNN traditionnel, la LSTM, la BiLSTM, et la GRU.

Nous avons également étudié l'impact des hyperparamètres sur ces modèles, ce qui nous a permis de mieux comprendre comment optimiser ces modèles pour des tâches de prédiction complexes.

Les résultats de nos expérimentations ont démontré la performance optimale du modèle BiLSTM avec 3 couches en utilisant l'optimiseur RMSprop avec un RMSE égal 2.10.

Enfin, notre analyse a révélé que le modèle BiLSTM est non seulement le meilleur parmi les modèles que nous avons étudiés, mais il est aussi plus efficace par rapport aux autres modèles tels que le RNN traditionnel, la LSTM et la GRU.

Conclusion Générale

En conclusion, ce mémoire a mis en évidence l'importance cruciale de l'agriculture dans l'économie mondiale et la nécessité d'améliorer la précision des prévisions des rendements agricoles. En se concentrant sur la région d'Annaba, une zone agricole clé en Algérie, il a exploré l'utilisation des réseaux neuronaux récurrents (RNN) et leurs variantes, comme les LSTM et les GRU, pour relever ce défi.

Le mémoire a souligné les défis uniques auxquels est confrontée l'agriculture à Annaba, notamment les variations climatiques, les techniques agricoles traditionnelles et les contraintes socio-économiques. Il a également proposé un système de prédiction des rendements agricoles en utilisant un ensemble de données robustes pour la région d'Annaba et en analysant l'impact des hyperparamètres sur les performances des modèles. Les résultats des expérimentations ont démontré la performance optimale du modèle de prédiction basé BiLSTM avec 3 couches en utilisant l'optimiseur RMSprop.

En fin de compte, ce travail vise à combler les lacunes des méthodes traditionnelles de prévision des rendements agricoles en utilisant des approches modernes basées sur les RNN. Il offre ainsi des solutions potentielles pour améliorer la précision des prévisions agricoles dans la région d'Annaba et potentiellement dans d'autres régions agricoles similaires. Il s'agit d'une contribution significative qui pourrait avoir un impact positif sur la gestion des ressources, la planification des opérations agricoles et la prise de décisions éclairées pour assurer la sécurité alimentaire.

Annexe

Le **Tableau A.1** représente l'architecture de model BiLSTM utilisé pour la prédiction des rendements agricoles.

Notre Base de données en GitHub : https://github.com/Moiz1500/PFE_Database

Tableau A.1. *Architecture de BiLSTM utilisée pour prédire le rendement Blé*

Layer (type)	Output Shape	Param #
bidirectional_1	(None, 12, 300)	186000
dropout_1	(None, 12, 300)	0
batch_normalization_1	(None, 12, 300)	1200
bidirectional_2	(None, 12, 200)	320800
dropout_2	(None, 12, 200)	0
batch_normalization_2	(None, 12, 200)	800
bidirectional_3	(None, 100)	100400
dropout_3	(None, 100)	0
batch_normalization_3	(None, 100)	400
dense_3	(None, 62)	6262
dropout_4	(None, 62)	0
dense_4	(None, 16)	1008
dropout_5	(None, 16)	0

Bibliographie

1. *Seif-El-Islam, R. (2020). Héritage industriel en Algérie, entre reconnaissance et abandon : Le cas du patrimoine minier et agricole de la ville d'Annaba. Ethnologies, 42(1), 167-185.*
2. *Autret, C. (2020). Production et consommation d'huile d'olive en Cilicie Trachée occidentale: état des recherches. Anatolia Antiqua. Revue internationale d'archéologie anatolienne, (XXVIII), 69-87.*
3. *Muruganantham, P., Wibowo, S., Grandhi, S., Samrat, N. H., & Islam, N. (2022). A systematic literature review on crop yield prediction with deep learning and remote sensing. Remote Sensing, 14(9), 1990.*
4. *Kheir, A. M., Alkharabsheh, H. M., Seleiman, M. F., Al-Saif, A. M., Ammar, K. A., Attia, A., ... & Schillaci, C. (2021). Calibration and validation of AQUACROP and APSIM models to optimize wheat yield and water saving in arid regions. Land, 10(12), 1375.*
5. *Wolanin, A., Mateo-García, G., Camps-Valls, G., Gómez-Chova, L., Meroni, M., Duveiller, G., ... & Guanter, L. (2020). Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. Environmental research letters, 15(2), 024019.*
6. *Meghraoui, K., Sebari, I., Pilz, J., Ait El Kadi, K., & Bensiali, S. (2024). Applied Deep Learning-Based Crop Yield Prediction: A Systematic Analysis of Current Developments and Potential Challenges. Technologies, 12(4), 43.*
7. *Burdett, H., & Wellen, C. (2022). Statistical and machine learning methods for crop yield prediction in the context of precision agriculture. Precision agriculture, 23(5), 1553-1574.*

8. Cen, H., & Wan, L. (2023). *Crop Yield Estimation and Prediction*. In *Encyclopedia of Smart Agriculture Technologies* (pp. 1-13). Cham: Springer International Publishing.
9. Banerjee, R., Bharti, Das, P., & Khan, S. (2024). *Crop Yield Prediction Using Artificial Intelligence and Remote Sensing Methods*. In *Artificial Intelligence and Smart Agriculture: Technology and Applications* (pp. 103-117). Singapore: Springer Nature Singapore.
10. Maheswary, A., Nagendram, S., Kiran, K. U., Ahammad, S. H., Priya, P. P., Hossain, M. A., & Rashed, A. N. Z. (2024). *Intelligent Crop Recommender System for Yield Prediction Using Machine Learning Strategy*. *Journal of The Institution of Engineers (India): Series B*, 1-9.
11. Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). *Quo vadis artificial intelligence?*. *Discover Artificial Intelligence*, 2(1), 4.
12. Dorr, L. (2022). *Types of Artificial Intelligence, Explained*.
13. Morales, E. F., & Escalante, H. J. (2022). *A brief introduction to supervised, unsupervised, and reinforcement learning*. In *Biosignal processing and classification using computational learning and intelligence* (pp. 111-129). Academic Press.
14. Zhang, Z., & Zhang, Z. (2018). *Artificial neural network. Multivariate time series analysis in climate and environmental research*, 1-35.
15. Janiesch, C., Zschech, P., & Heinrich, K. (2021). *Machine learning and deep learning*. *Electronic Markets*, 31(3), 685-695
16. Patterson, J., & Gibson, A. (2017). *Deep learning: A practitioner's approach*. " O'Reilly Media, Inc."

17. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical evaluation of gated recurrent neural networks on sequence modeling*. *arXiv preprint arXiv:1412.3555*.
18. Lynn, H. M., Pan, S. B., & Kim, P. (2019). *A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks*. *IEEE Access*, 7, 145395-145405.
19. Hu, L., Wang, J., Guo, Z., & Zheng, T. (2023). *Load Forecasting Based on LVMD-DBFCM Load Curve Clustering and the CNN-IVIA-BLSTM Model*. *Applied Sciences*, 13(12), 7332.
20. Gal, Y., & Ghahramani, Z. (2016). *A theoretically grounded application of dropout in recurrent neural networks*. *Advances in neural information processing systems*, 29.
21. Chandriah, K. K., & Naraganahalli, R. V. (2021). *RNN/LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting*. *Multimedia Tools and Applications*, 80(17), 26145-26159.
22. Ye, Y., Huang, Q., Rong, Y., Yu, X., Liang, W., Chen, Y., & Xiong, S. (2023). *Field detection of small pests through stochastic gradient descent with genetic algorithm*. *Computers and Electronics in Agriculture*, 206, 107694.
23. Elshamy, R., Abu-Elnasr, O., Elhoseny, M., & Elmougy, S. (2023). *Improving the efficiency of RMSProp optimizer by utilizing Nesterov in deep learning*. *Scientific Reports*, 13(1), 8814.
24. Karaman, Y., Akdeniz, F., Savaş, B. K., & Becerikli, Y. (2022, October). *A comparative analysis of svm, lstm and cnn-rnn models for the bbc news classification*. In *The Proceedings of the International Conference on Smart City Applications* (pp. 473-483). Cham: Springer International Publishing.

25. Ren, J., Zhang, M., Yu, C., & Liu, Z. (2022). *Balanced mse for imbalanced visual regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7926-7935).*

26. Chicco, D., Warrens, M. J., & Jurman, G. (2021). *The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peerj computer science, 7, e623.*

Webographie

1. <https://fr.wikipedia.org/wiki/Annaba>
2. <https://annaba.mta.gov.dz/fr/notre-wilaya/>
3. <https://airalgerie.dz/en/blog/wilaya/annaba/>
4. <https://www.algerie360.com/de-grandes-potentialits-agricoles-annaba-a-traine/>
5. [World hunger is still not going down after three years and obesity is still growing – UN report \(who.int\)](#)
6. [Crop Yield Prediction - SATPALDA](#)
7. [64. Use of Remote Sensing in Crop Yield Forecasting.pdf \(justagriculture.in\)](#)
8. [BatchNormalization layer \(keras.io\)](#)
9. [Dense layer \(keras.io\)](#)
10. [A Gentle Introduction To Sigmoid Function - MachineLearningMastery.com](#)
11. [Tanh Activation Function for Deep Learning: A Complete Guide • datagy](#)
12. [A Gentle Introduction to the Rectified Linear Unit \(ReLU\) - MachineLearningMastery.com](#)