

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة أبي بكر بلقايد - تلمسان -

Université Aboubakr Belkaïd – Tlemcen –

Faculté des Sciences économiques, Commerce et de Gestion



Thèse

Présenté en vue de l'obtention du diplôme de *DOCTORAT en Techniques Quantitatives Appliquées en Economie*

Thème

Prévision utilisant l'intelligence artificielle : une étude standard des fluctuations des prix du pétrole pour le panier de l'OPEP au cours de la période 2020-2024

Présenté par :

Mme. HAMDANI Souad

Sous la direction :

Pr. BENAMAAR Soumia

Soutenu publiquement, le 14 / 10 / 2025 .

Devant le jury composé de :

Mr. BETTAHAR Samir	Professeur	Univ. Tlemcen	Président
Mme. BENAMAR Soumia	Professeur	Univ. Tlemcen	Directeur de thèse
Mme. HALIMI Wahiba	Professeur	Univ. Tlemcen	Examineur 1
Mr. METAACH Mohamed Amine	Professeur	Ecole supérieur de management. Tlemcen	Examineur 2
Mr. INAL Amina	Docteur	Univ. Sidi Belabess	Examineur 3
Mme. RAHALI Amina Souad	Docteur	Ecole supérieur de management. Tlemcen	Examineur 4

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Remerciement

En premier lieu, je rends grâce à Dieu, le Tout-Puissant, pour m'avoir accordé la patience, la force et la persévérance nécessaires pour mener à bien cette thèse. Sans Sa bénédiction, rien de tout cela n'aurait été possible.

Je remercie particulièrement mon directeur de thèse Pr. BENAMAR Soumia, pour sa disponibilité, ses conseils précieux et son accompagnement tout au long de cette recherche.

Je tiens également à exprimer ma profonde gratitude envers ma famille, pour leur soutien, leur amour et leur compréhension tout au long de ce parcours. Leur encouragement constant a été une source inestimable de motivation dans les moments difficiles. Trop nombreux sont ceux que je n'ai pu nommer. Dans ces années de recherche j'ai vécu une aventure passionnante, Ce temps dédié à la recherche m'a beaucoup apporté tant sur le plan scientifique que sur le plan humain. (الحمد لله)

Enfin, je souhaite remercier toutes les personnes qui, d'une manière ou d'une autre, ont contribué à la réalisation de ce travail, que ce soit par leurs conseils, leur assistance ou leur soutien moral.

***Que Dieu bénisse tous ceux qui ont été présents dans ce
parcours.***

ملخص:

في سياق يتسم بزيادة عدم اليقين، تلعب التنبؤات دورًا مركزيًا في توجيه القرارات الاستراتيجية. فهي تعتمد على التنبؤ بالتطورات المستقبلية استنادًا إلى البيانات التاريخية. ومع ذلك، فإن الأساليب التقليدية للتنبؤ مثل نماذج ARIMA و GARCH، على الرغم من فعاليتها في التعامل مع البيانات الخطية أو البسيطة، فإنها تصل إلى حدودها عند مواجهة الظواهر غير الخطية أو السلاسل التي تتسم بالتقلبات العالية.

لقد حولت تقنيات الذكاء الاصطناعي (AI) بشكل عميق مجال التنبؤ. تساعد الأساليب مثل التعلم العميق، وبخاصة الشبكات العصبية الاصطناعية، والنماذج الهجينة التي تجمع بين الذكاء الاصطناعي والتقنيات التقليدية في التغلب على العديد من هذه القيود. تقدم هذه الأساليب تنبؤات أكثر دقة وموثوقية، حتى في سياقات تتسم بعدم اليقين أو وجود اضطرابات في البيانات.

ومع ذلك، فإن تقنيات الذكاء الاصطناعي ليست خالية من التحديات: فالقابلية للتفسير للنماذج، وجودة البيانات المدخلة، وثقة المستخدمين تظل قضايا حاسمة. من خلال الجمع بين قوة الأساليب التقليدية وقوة الذكاء الاصطناعي، يصبح من الممكن تحسين جودة التنبؤات مع معالجة هذه التحديات. يفتح هذا المزيج آفاقًا واعدة للتطبيقات الاستراتيجية، وخاصة في إدارة تقلبات السلاسل المالية.

في هذه الرسالة، نستكشف نهجًا هجينًا يجمع بين النماذج التقليدية وتقنيات الذكاء الاصطناعي للتنبؤ بالسلاسل الزمنية. هدفنا هو استغلال التكامل بين الأساليب الإحصائية والشبكات العصبية الاصطناعية لتحسين دقة وقوة التنبؤات. كمثال توضيحي، قمنا بتطوير نموذج هجيني GARCH-NN تطبيقًا لتنبؤ أسعار النفط، مما يبرز قدرته على التقاط الديناميكيات المعقدة والخصائص المتقلبة للأسواق المالية.

الكلمات المفتاحية: الأساليب التقليدية، تقنيات الذكاء الاصطناعي، التقلبات، التنبؤ، GARCH، NAR، GARCH-NN، هجيني، XAI.

Résumé

Dans un contexte marqué par une incertitude croissante, la prévision occupe une place centrale pour orienter les décisions stratégiques. Elle repose sur l'anticipation des évolutions futures à partir de données historiques. Toutefois, les approches classiques de prévision, comme les modèles ARIMA et GARCH, bien qu'efficaces pour des données linéaires ou peu complexes, atteignent leurs limites face à des phénomènes non linéaires ou des séries caractérisées par une forte volatilité.

L'émergence des techniques d'intelligence artificielle (IA) a profondément transformé le domaine de la prévision. Des méthodes comme le Deep Learning, notamment les réseaux de neurones artificiels, et les modèles hybrides associant IA et techniques traditionnelles, permettent de surmonter plusieurs de ces limitations. Ces approches offrent des prédictions plus précises et robustes, même dans des contextes marqués par l'incertitude ou des perturbations dans les données.

Cependant, les techniques d'IA ne sont pas exemptes de défis : l'interprétabilité des modèles, la qualité des données d'entrée, et la confiance des utilisateurs demeurent des enjeux cruciaux. En combinant les forces des approches traditionnelles et celles de l'IA, il devient possible d'améliorer la qualité des prévisions tout en répondant à ces défis. Cette combinaison ouvre des perspectives prometteuses pour des applications stratégiques, notamment dans la gestion de la volatilité des séries financières.

Dans cette thèse, nous explorons une approche hybride qui associe des modèles traditionnels et des techniques d'intelligence artificielle pour la prévision de séries temporelles. Notre objectif est d'exploiter la complémentarité entre les méthodes statistiques et les réseaux de neurones artificiels afin d'améliorer à la fois la précision et la robustesse des prédictions. À titre illustratif, nous avons développé un modèle hybride GARCH-NN appliqué à la prédiction des prix du pétrole, démontrant sa capacité à capturer les dynamiques complexes et les caractéristiques volatiles des marchés financiers.

Mots Clé : Techniques Traditionnelles, Techniques de l'intelligence Artificielle, Volatilité, Prévision, GARCH, NAR, GARCH-NN, Hybride, XAI.

Abstract

In a context marked by increasing uncertainty, forecasting plays a central role in guiding strategic decisions. It relies on the anticipation of future developments based on historical data. However, traditional forecasting approaches, such as ARIMA and GARCH models, while effective for linear or less complex data, reach their limits when faced with nonlinear phenomena or series characterized by high volatility.

The emergence of artificial intelligence (AI) techniques has profoundly transformed the field of forecasting. Methods such as Deep Learning, particularly artificial neural networks, and hybrid models combining AI with traditional techniques, help overcome many of these limitations. These approaches offer more precise and robust predictions, even in contexts marked by uncertainty or data disturbances.

However, AI techniques are not without challenges: model interpretability, the quality of input data, and user trust remain crucial issues. By combining the strengths of traditional approaches with those of AI, it becomes possible to improve the quality of forecasts while addressing these challenges. This combination opens up promising perspectives for strategic applications, particularly in managing the volatility of financial time series.

In this thesis, we explore a hybrid approach that combines traditional models and artificial intelligence techniques for time series forecasting. Our goal is to leverage the complementarity between statistical methods and artificial neural networks to improve both the accuracy and robustness of predictions. As an illustrative example, we have developed a GARCH-NN hybrid model applied to oil price forecasting, demonstrating its ability to capture the complex dynamics and volatile characteristics of financial markets.

Keywords : Traditional Techniques, Artificial Intelligence Techniques, Volatility, Forecasting, GARCH, NAR, GARCH-NN, Hybrid, XAI.

Tables des Matières

INTRODUCTION GENERALE.....	1
PARTIE I : THEORIE	7
SECTION 1 : L'ECONOMETRIE ET L'INTELLIGENCE ARTIFICIELLE.....	7
INTRODUCTION.....	7
1 L'ECONOMETRIE	8
1.1 Les normes de la pratique économétrique	10
1.2 Méthodes traditionnelles de prévision.....	11
1.3 Le Big Data et l'économétrie	11
1.4 Algorithmes d'apprentissage automatique (Machine Learning ML) et l'économétrie.....	12
1.4.1 Applications de l'IA dans les prévisions.....	12
2 LES LIMITES DE L'INTELLIGENCE ARTIFICIEL POUR L'ECONOMETRIE	13
3 LES PERSPECTIVES FUTURES DE L'IA DANS LES RECHERCHES ECONOMETRIQUES	14
CONCLUSION	14
SECTION 2 : TECHNIQUES TRADITIONNELLES DE PREVISION	17
INTRODUCTION.....	17
1 LES SERIES TEMPORELLES (SERIES CHRONOLOGIQUES).....	17
2 LES MODELES STOCHASTIQUE DE PREVISION	18

2.1	Modèle autorégressif (AR).....	19
2.2	Processus de moyenne mobile (MA).....	19
	Un bruit blanc	19
2.3	Le processus ARMA	20
2.4	Les Modèle ARIMA	20
3	LA PREVISION PAR UN MODELE ARMA(P,Q)	21
4	LES MODELE ARCH LINEAIRE UNIVARIEE	21
4.1	Le modèle ARCH	22
4.2	Le modèle GARCH	23
	4.2.1 La modélisation ARMA en présence d'effet GARCH	24
	4.2.1.1 Le test d'asymétries	24
	4.2.2 Les modèle GARCH Multivariés	25
4.3	Le modèle IGARCH.....	25
4.4	Le modèle EGARCH.....	25
4.5	Le modèle GJR-GARCH.....	26
	4.5.1 Explication du modèle.....	26
5	LA PREVISION PAR UN MODELE GARCH (1.1).....	26
	CONCLUSION	27
	SECTION 3 : LES TECHNIQUES DE L'INTELLIGENCE ARTIFICIELLE (DEEP LEARNING)	30
	INTRODUCTION.....	30
1	DOMAINES D'APPLICATION	30

2	DEEP LEARNING POUR LES SERIES TEMPORELLES.....	31
2.1	Les réseaux de neurones artificiels.....	31
2.1.1	Mise en œuvre d'un réseau de neurone.....	32
2.1.2	Architecture d'un réseau neurone	32
2.1.3	Les réseaux de neurones non bouclés	32
2.1.4	Réseaux de neurones bouclés.....	33
2.1.5	Exemple de modélisation d'un seul neurone.....	34
2.1.6	L'apprentissage du réseau neurone.....	34
2.1.6.1	Algorithme d'apprentissage.....	35
2.1.6.2	Fonctionnement	36
3	LES ALGORITHMES GENETIQUES AGS	37
3.1	Présentation des algorithmes génétiques.....	37
3.2	Fonctionnement d'un algorithme génétique.....	38
3.3	Séries temporelles et algorithmes génétiques.....	39
4	LA MODELISATION DES SERIES TEMPORELLES BASEE SUR LES RESEAUX DE NEURONE ARTIFICIELLE.....	40
4.1	Réseaux de neurones artificiels autorégressif non-linéaire NAR.....	41
4.1.1	Réseau neurone autorégressif NAR	41
4.1.1.1	Série financière.....	43
4.1.1.2	Les caractéristiques des séries financières	43
4.2	Modèles hybridés	45
	CONCLUSION	47
	PARTIE II : PRATIQUE	48
	INTRODUCTION.....	49

SECTION 1 : LES TECHNIQUES DE L'INTELLIGENCE ARTIFICIELLE UTILISEES DANS LA PREVISION DES PRIX DU PETROLE	53
INTRODUCTION.....	53
1 HISTORIQUE DES FLUCTUATIONS DES PRIX 2020-2022	54
2 LA VOLATILITE DES PRIX DU PETROLE	55
3 LA PREVISION DES PRIX DU PETROLE	56
4 L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE POUR PREDIRE LA VOLATILITE DES PRIX DU PETROLE	57
1- Modèles ARIMA-GARCH	58
2- Modèles SVM-ANN	59
3- Modèles VARX-MLP	60
4- Modèles LSTM-CNN	60
5- Les modèles GARCH-NN	61
6- Le modèle NN-GARCH	62
CONCLUSION	64
SECTION 2 : LA MODELISATION DE LA VOLATILITE DES PRIX DU PETROLE PAR LE MODELE GARCH	67
INTRODUCTION.....	67
1 ANALYSE DESCRIPTIVES ET EXPLORATION DES DONNEES.....	67
1.1 Outils Utilisés pour l'Implémentation et la Préparation des Données	68
1.1.1 Langage de programmation MATLAB R2024a :	68
1.1.2 Toolboxes Principales	68

1.1.2.1 Econometrics Toolbox :	68
1.1.2.2 Statistics and Machine Learning Toolbox:.....	68
1.1.2.3 Financial Toolbox :	68
1.1.2.4 System Identification Toolbox :	69
1.1.3 L'environnement de travail Econometric Modeler	69
1.1.3.1 Fonctionnalités de l'Econometric Modeler.....	69
1.2 Etude graphique.....	69
1.3 Etude de la stationnarité des prix du pétrole	75
1.3.1 Stationnarité	75
1.4 Décomposition de la série des prix du pétrole :	76
2 MODELISATION DE LA SERIE DES PRIX DU PETROLE.....	80
2.1 La Modélisation avec un GARCH	80
2.1.1 Identification du modèle GARCH.....	81
2.1.2 La spécification d'un modèle GARCH.....	82
2.1.3 L'estimation des paramètres du modèle	83
2.1.4 Validation du modèle GARCH (1.1).....	83
2.1.4.1 Vérification des Résidus Standardisés	84
2.1.4.2 Autocorrélation des Résidus Standardisés et le Carré des résidus standardisés.....	84
2.1.4.3 Test ARCH sur les Résidus Standardisés.....	85
3 LA PREVISION DE LA VOLATILITE DES PRIX DU PETROLE PAR UN MODELE GARCH (1.1).....	86
CONCLUSION.....	87
SECTION 3 : LA MODELISATION DE LA VOLATILITE DES PRIX DU PETROLE PAR LE MODELE NAR.....	90
INTRODUCTION.....	90
1 LA STRUCTURE DU MODELE NAR.....	90

2	LA MODELISATION DES PRIX DU PETROLE PAR UN MODELE NAR	90
a.	L'architecture du réseau neurone	92
b.	Entraînement du modèle NAR.....	93
c.	Test de la performance du réseau de neurone NAR	93
3	LA PREVISION DES PRIX DU PETROLE PAR LE NAR	99
d.	Prévision muti-step	99
e.	Prévision un pas Step-Ahead Prediction Network.....	100
SECTION 4 : LA MODELISATION DE LA VOLATILITE DES PRIX DU PETROLE PAR LE MODELE GARCH-NN		105
INTRODUCTION.....		105
1	LA PREVISION PAR LE MODELE GARCH-NN	105
a.	La modélisation GARCH.....	105
b.	Extraction des résidus et des volatilités conditionnelles.....	105
c.	Conception du Réseau de Neurones	106
2	VALIDATION DU MODELE GARCH-NN (1.1) (1 :1) (10).....	109
3	LA PREVISION DE LA VOLATILITE DES PRIX DU PETROLE.....	112
CONCLUSION		115
CONCLUSION GENERALE		119
BIBLIOGRAPHIE		122

Tables des Figures

FIGURE 1:RESEAU NEURONE A COUCHE	33
FIGURE 2:RESEAU NEURONE BOUCLE.....	33
FIGURE 3:RESEAU NEURONE AUTOREGRESSIF NAR	41
FIGURE 4:EVOLUTION DES PRIX DU PETROLE ENTRE 02/01/2020 ET 29/01/2024 ..	70
FIGURE 5:LE GRAPHE SAISONNIER DES PRIX DU PETROLE ENTRE 03/01/2020 ET 29/12/2023.	72
FIGURE 6:LE GRAPHE MOYENNE MENSUELLES DES PRIX DU PETROLE ENTRE 03/01/2020 ET 29/01/2024.	73
FIGURE 7:LE GRAPHE DE LA VARIANCE MENSUELLES DES PRIX DU PETROLE ENTRE 03/01/2020 ET 29/01/2024.	73
FIGURE 8:L'EVOLUTION DES PRIX DU PETROLE ET LA LIGNE TREND / MOYENNE.....	77
FIGURE 9:EVOLUTION DES RESIDUS OBTENUS DE L'ESTIMATION DE LA FONCTION DE LA TENDANCE DE LA SERIE PT.....	78
FIGURE 10:LA SERIE DES PRIX DU PETROLE DIFFERENCIER ΔP_T	81
FIGURE 11:PREVISIONS DE LA VOLATILITE CONDITIONNELLE A L'HORIZON DE 30JOURS.	86
FIGURE 12:LA VOLATILITE GLISSANTE DES PRIX DU PETROLE.....	91
FIGURE 13:L'ARCHITECTURE DU NAR (1 :2) (10) DES PRIX DU PETROLE.	92
FIGURE 14:PERFORMANCE DU RESEAU DE NEURONE.....	95
FIGURE 15:PERFORMANCE DU RESEAU DE NEURONE NAR (1 :1)(10).....	96
FIGURE 16:PERFORMANCE DU RESEAU DE NEURONE NAR (1 :1)(14).....	98
FIGURE 17:LE RESEAU EN BOUCLE FERME (CLOSE LOOP)	100
FIGURE 18:LE RESEAU EN BOUCLE OUVERTE (OPEN LOOP)	100
FIGURE 19:RESPONSE DES ELEMENTS DE SORTIE.....	101
FIGURE 20:ARCHITECTURE DU NAR DES RESIDUS STANDARDISES.	106
FIGURE 21:ARCHITECTURE DU RESEAU NE NEURONE AJUSTE.....	108
FIGURE 22:LA PERFORMANCE DU RESEAU DE NEURONE DES RESIDUS STANDARDISES.....	109
FIGURE 23:RESPONSE OF OUTPUT ÉLÉMENT 1 FOR TIME SERIES.....	111

FIGURE 24:RESEAU EN BOUCLE FERMEE.....	112
FIGURE 25:RESEAU EN BOUCLE OUVERTE.....	113
FIGURE 26:COMPARAISON ENTRE LES VALEURS REELLES ET LES VALEURS PREVUS.....	114
TABLEAU 1:REGRESSION LINAIRE DE LA VARIANCE MENSUELLES SUR LES MOYENNES MENSUELLES DES PRIX DU PETROLE.....	74
TABLEAU 2:DIVISION DES DONNEES UTILISEES	74
TABLEAU 3:LES TEST DE DICKEY-FULLER AUGMENTES DES PRIX DU PETROLE	76
TABLEAU 4:REGRESSION LINAIRE DE LA TENDANCE.....	78
TABLEAU 5:TEST DE DFA DE LA SERIE DES PRIX DIFFERENCIER.....	80
TABLEAU 6:TEST DE ENGLE'S ARCH.....	81
TABLEAU 7:MODELE GARCH RETENUS.....	82
TABLEAU 8:ESTIMATION DU MODELE GARCH (1.1).....	83
TABLEAU 9:TEST LJING-BOX DES RESIDUS STANDARDISES.....	85
TABLEAU 10:TEST ENGLE'S ARCH SUR LES RESIDUS STANDARDISES.....	86
TABLEAU 11:RESULTAT D'ENTRAINEMENT ET LES STATISTIQUES D'AJUSTEMENT DU RESEAU DE NEURONE.....	93
TABLEAU 12:RESULTAT D'ENTRAINEMENT ET LES STATISTIQUES D'AJUSTEMENT DU RESEAU DE NEURONE NAR (1:1)(10).....	95
TABLEAU 13:RESULTAT D'ENTRAINEMENT ET LES STATISTIQUES D'AJUSTEMENT DU RESEAU DE NEURONE NAR (1:1)(14).....	97
TABLEAU 14:RESULTATS D'ENTRAINEMENT DES RESIDUS STANDARDISES (1 :2) (10).....	106
TABLEAU 15:RESULTATS D'ENTRAINEMENT DES RESIDUS STANDARDISES (1 :1) (10).....	107
TABLEAU 16:COMPARAISON DES MSE ET RMSE DES RESIDUS REELS ET AJUSTES.....	114
TABLEAU 17:COMPARAISON ENTRE LES PERFORMANCES DES TROIS MODELES.	115

Liste des Corrélogrammes et Histogramme

Corrélogrammes

CORRELOGRAMME 1:LES AUTOCORRELATION SIMPLE ET PARTIELLE DE LA SERIE DES PRIX DU PETROLE.	75
CORRELOGRAMME 2:LES AUTOCORRELATIONS SIMPLE ET PARTIELLES DES RESIDUS DE LA REGRESSION TREND.	79
CORRELOGRAMME 3:LES AUTOCORRELATIONS SIMPLES ET PARTIELLES DE LA NOUVELLE SERIE DES PRIX DU PETROLE.	79
CORRELOGRAMME 4:RESIDUS STANDARDISES ET LE CARRE DES RESIDUS STANDARDISES.....	85
CORRELOGRAMME 5:LES RESIDU DU MODELE GARCH-NN (1.1) (1 :1) (10).....	118

Histogrammes

HISTOGRAMME 1:RESIDUS STANDARDISES DE L'ESTIMATION GARCH (1.1).....	84
HISTOGRAMME 2:HISTOGRAMME DES ERREURS DU MODELE NAR (1 :1)(14). ...	98
HISTOGRAMME 3:ERREUR DE PREVISION DU MODELE.....	110
HISTOGRAMME 4:RESIDUS DU MODELE GARCH -NN (1.1) (1 :1) (10).....	117

Introduction Générale

Introduction Générale

Introduction Générale

Le progrès de l'intelligence artificielle (IA) a ouvert de nouvelles opportunités dans le domaine de la prévision économétrique à cause de leurs techniques qui offrent des capacités de modélisation et d'apprentissage avancées qui permettent de relever les défis complexes et les limites auxquels sont confrontées les méthodes de prévision traditionnelles.

Les méthodes de prévision traditionnelles, telles que les modèles de régression linéaire, les modèles ARIMA et les méthodes de GARCH étaient utilisées pour prévoir les phénomènes économiques, financiers et sociaux. Bien que ces approches aient démontré leur utilité, elles présentent des limites lorsqu'il s'agit de capturer les dynamiques non linéaires et les interactions complexes dans les données.

L'utilisation de l'IA est devenue très populaire dans différents sujets tel que l'économie et les finances à cause de sa capacité de minimiser les travaux effectués par l'être humain ce qui conduit à une efficacité et une productivité dans la résolution des différents problèmes.

L'efficacité démontrée par les algorithmes automatiques, tels que les réseaux de neurones artificiels, les machines à vecteurs de support pour capturer la non linéarité et les interactions complexes existant dans les données les rend adaptés à la prévision des phénomènes économiques, financiers et sociaux qui présentent souvent des comportements non linéaires.

De plus, l'apprentissage profond (Deep Learning) a révolutionné de nombreux domaines de la prévision. Les architectures de réseaux de neurones, comme les réseaux récurrents et les réseaux convolutifs, ont permis d'améliorer les applications telles que la prévision des séries temporelles, la prévision de la demande et la prévision des prix.

Bien que l'IA offre de nombreux avantages comme l'amélioration de la précision des prévisions, le traitement efficace de grands volumes de données, la possibilité d'automatiser et d'accélérer les processus de prévision, ce qui permet une rapidité et une bonne prise de décision.

L'intégration de l'IA dans les processus de prévision nécessite une compréhension approfondie des techniques, une gestion rigoureuse des données et une évaluation attentive des performances des modèles. L'objectif est de bénéficier de la puissance de l'IA tout en maintenant des pratiques de prévision fiables.

Introduction Générale

Ce travail concentrera sur l'exploration des différentes techniques de prévision, à la fois les méthodes traditionnelles et les approches basées sur l'intelligence artificielle ce qui fait l'objet de la première partie du travail. Une attention particulière sera accordée à l'étude du modèle hybride GARCH-NN (Generalized Autoregressive Conditional Heteroskedasticity - Neural Network), qui combine le modèle GARCH pour la modélisation de la volatilité et réseaux de neurones pour capturer les schémas non linéaires et complexes. Ce modèle sera appliqué à la prévision de la volatilité des prix du pétrole, un domaine où les techniques de prévision traditionnelles ont montré leurs limites dans la précision et la qualité des prévisions, où l'IA pourrait apporter des améliorations significatives et cela fait l'objet de la deuxième partie du travail.

Objectifs de la recherche :

L'objectif principal de cette étude sur la prévision par les techniques de l'intelligence artificielle est de développer et d'appliquer des modèles capables de prédire des événements futurs ou des résultats en se basant sur des données historiques. Cette démarche vise à améliorer la précision des prévisions en utilisant des algorithmes avancés. La combinaison entre l'IA et un modèle traditionnel GARCH-NN est pour objectif d'examiner comment cette technique peut améliorer la précision et de tester l'approche hybride pour vérifier les gains en précision et en robustesse par rapport à l'approches classiques GARCH et par rapport à une modélisation par un réseau neurone autorégressive seul passant par les étapes de modélisation de chaque approche. Discuter l'amélioration apporter par le modèle hybride GARCH_NN dans la prévision de la volatilité des prix du pétrole. Enfin, cette approche est conçue pour résoudre des problématiques spécifiques dans des domaines variés tels que la finance, l'énergie, répondant ainsi à des besoins concrets dans des contextes divers.

La problématique de la recherche :

Face à la complexité croissante de la modélisation des données financière très volatile, les méthodes traditionnelles de prévision atteignent leurs limites, notamment en termes de précision, de gestion des non-linéarités et d'adaptabilité à des contextes dynamiques. Dans ce contexte, les techniques d'intelligence artificielle, telles que les réseaux neurones artificielles, offrent des opportunités prometteuses et considéré excellent comme technique dans ces scénarios mais manquent souvent d'interprétabilité et nécessitent des ressources de très bonne qualité. L'émergence des approches hybrides, combinant la rigueur des modèles traditionnels

Introduction Générale

avec la flexibilité et la puissance des techniques d'IA, promet de pallier ces limites. Cela nous conduit à poser deux problématiques pour notre recherche :

- Aux quel point les techniques de l'IA peuvent améliorer la qualité de la prévision par rapport aux techniques traditionnelles ?
- Est-ce que les modèles hybrides combinant les deux approches peuvent remplacer l'utilisation exclusives d'une de ces approches ?

Pour bien répondre à notre problématique nous allons examiner les trois hypothèses suivantes :

- **H₁** : Les techniques d'intelligence artificielle surpassent les méthodes traditionnelles de prévision en termes de précision.
- **H₂** : les techniques de l'IA peuvent mieux capturer les tendance non linaires et les anomalies dans les séries temporelles.
- **H₃** : Les approches hybrides combinant IA et méthodes traditionnelles offrent une meilleure robustesse que l'utilisation exclusive de l'une ou l'autre.

PARTIE I

Théorie

Section 1
L'économétrie et l'Intelligence
Artificielle

Sommaire

SECTION 1 : L'ECONOMETRIE ET L'INTELLIGENCE ARTIFICIELLE.....	7
INTRODUCTION.....	7
1 L'ECONOMETRIE	8
1.1 Les normes de la pratique économétrique.....	10
1.2 Méthodes traditionnelles de prévision.....	11
1.3 Le Big Data et l'économétrie	11
1.4 Algorithmes d'apprentissage automatique (Machine Learning ML) et l'économétrie.....	12
1.4.1 Applications de l'IA dans les prévisions.....	12
2 LES LIMITES DE L'INTELLIGENCE ARTIFICIEL POUR L'ECONOMETRIE	13
3 LES PERSPECTIVES FUTURES DE L'IA DANS LES RECHERCHES ECONOMETRIQUES	14
CONCLUSION	14

Section 1 : L'économétrie et l'Intelligence Artificielle

Partie I : Théorie

Section 1 : L'économétrie et l'Intelligence Artificielle

Introduction

Pour introduire le sujet de l'utilisation des techniques de l'intelligence artificielle (IA) dans la prévision qui est considéré comme un objectif de base de l'économétrie il faut donner quelques définitions et une projection sur l'historique de l'assemblage des deux discipline économétrie et intelligence artificielle (IA). En générale, l'économétrie est l'une des branches de la science économique qui assemble de l'économie et les statistiques elle joue un rôle central dans la compréhension et la quantification des complexités des phénomènes économiques. Alors que l'intelligence artificielle (IA) et l'apprentissage automatique le machine Learning (ML) et la science des données sont développés rapidement et deviennent de plus en plus influents dans plusieurs secteurs et domaines, l'économétrie se retrouve à l'avant-garde, fournissant un champ très puissant pour l'analyse des données économiques et la prise des décisions éclairées. Dans ce chapitre, ont approfondi le monde de l'économétrie dans le contexte de l'intelligence artificielle (IA) et le Machine Learning (ML) en explorant ses origines, ses applications, ses cas d'utilisation, ses perspectives de carrière et ses meilleures pratiques.

Section 1 : L'économétrie et l'Intelligence Artificielle

1 L'économétrie

L'économétrie est une branche de la science économique qui fait appelle conjointement à plusieurs disciplines comme la théorie économique, la statistique et les mathématiques et de plus en plus l'informatique. Cette conjointent de ces quatre disciplines qui va définir l'économétrie et pour donner une définition plus précise on va noter la définition originale élaborer par **Rognaros FRISCH** 1933, « l'économétrie n'est pas du tout la même chose que la statistique économique, elle ne doit pas être non plus identifier avec ce que l'on appelle la théorie économique générale bien qu'une partie considérable de cette théorie est un caractère résolument quantitatif. L'économétrie ne doit pas non plus être considéré comme synonyme de l'application des mathématiques à l'économie.

L'expérience a montré que chacun de ces trois points de vue donc ce lui des statistique, théorie économique et de celui des mathématiques est une condition nécessaire mais non suffisante elle-même à une compréhension réelle des relations quantitatives dans la vie économique et donc c'est l'unification des trois qui est puissante et c'est cette unification qui va définir l'économétrie »

D'un point de vue plus pratique l'économétrie va consister simplement l'étude et la mesure de phénomènes économiques à partir de l'observation des grandeurs pertinentes qui vont décrire ces phénomènes. L'économétrie cherche à quantifier des relations causales entre des variables économiques « L'économétrie ni une méthode de découvert ni une méthode de critique c'est un moyen de donnée une précision quantitative à ce que nous savons déjà en terme qualitative » KEYNES donc l'économétrie va servir à quantifier des relations guidées par la théorie qualitative.

L'économétrie continu à se développer et ce développement s'est accéléré en particulier avec l'avènement des ordinateurs et grâce à l'application des big data qui a permis aussi de faire évoluer l'économétrie permettant l'analyse de grands ensembles de données et de modèles complexes.

Aujourd'hui, l'économétrie constitue la base de la recherche économique, de l'élaboration des politiques et de la prise de décision dans le monde universitaire et industriel. A l'ère de l'intelligence artificielle (IA) et le Machine Learning (ML) l'économétrie est considéré comme une boîte à outils précieuse pour comprendre les phénomènes économiques et exploiter les données pour prendre des décisions importantes. Alors que la science des données se concentre sur l'extraction d'informations à partir des données à l'aide des techniques

Section 1 : L'économétrie et l'Intelligence Artificielle

informatiques et statistiques. L'économétrie apporte un point de vue économique à cette opération en intégrant la théorie et la modélisation économiques dans l'analyse et permet aux économistes et aux data scientists d'estimer et de quantifier les relations entre les variables économiques et de comprendre la causalité entre ces variables et prévenir les résultats qui est le processus le plus important dans l'étude économétrique selon FRIDMEN « un bon modèle dit bonne prédiction ».

L'économétrie offre des perspectives dans le monde universitaire et industriel, en particulier dans le domaine de l'intelligence artificielle et le Machine Learning. La maîtrise de l'économétrie donne aux professionnels les compétences nécessaires pour analyser des données économiques complexes et de créer des modèles robustes et tirer des informations significatives. Plusieurs domaines utilisent l'économétrie où on peut citer quelques exemples de cette utilisation :

- **La modélisation financière** : où les modèles économétriques sont utilisés pour prévoir les cours des actions et analyser les risques de rendement aussi pour prédire les tendances du marché et évaluer l'impact des politiques financières. Par exemple, les modèles de séries temporelles tels que les modèles de moyenne mobile intégrée autorégressive (ARIMA) qui sont couramment utilisés pour prédire la volatilité et les prix des marchés boursiers et pétroliers.
- **Évaluation des politiques** : l'économétrie est utilisée pour évaluer l'impact des politiques et des interventions publiques. En comparant les résultats avant et après la mise en œuvre des politiques, les modèles économétriques peuvent mesurer les effets causals et informer les décideurs politiques sur l'efficacité des diverses interventions.
- **Le marketing** : l'économétrie joue un rôle très important dans l'analyse du comportement des consommateurs et de la demande du marché et des stratégies de tarification. En combinant des données sur la démographie des consommateurs, les habitudes d'achat et les campagnes marketing, les modèles économétriques peuvent estimer l'efficacité des initiatives marketing et guider la prise de décision.
- **Économie de la santé** : l'économétrie est utilisée pour étudier la relation entre l'utilisation des soins de santé et les coûts et les résultats de santé. En analysant un grand ensemble de données sur les soins de santé, dans ce domaine les modèles économétriques peuvent être utilisés à identifier les facteurs qui affectent les résultats des soins de santé et à éclairer les décisions d'allocation des ressources.

Section 1 : L'économétrie et l'Intelligence Artificielle

- **Économie de l'environnement** : l'économétrie joue un rôle essentiel dans l'analyse de l'impact des politiques environnementales en estimant les effets de ces politiques sur les émissions, la consommation d'énergie et les résultats économiques, les modèles économétriques peuvent guider les décideurs politiques dans la conception d'interventions environnementales efficaces.

1.1 Les normes de la pratique économétrique

Pour garantir une analyse économétrique robuste et fiable, il est essentiel de suivre les meilleures pratiques et de respecter les normes établies.

1-Le prétraitement des données	Pour une analyse économétrique précise il est important de travailler avec des données de haute qualité en transformant les variables, gérant les valeurs manquantes et en résolvant les problèmes de multi-colinéarité.
2-Sélection et spécification du modèle	Le choix des modèles économétriques appropriés en fonction de la question de recherche et des caractéristiques des données. Prendre en considération les facteurs tels que les hypothèses du modèle, l'endogénéité, l'hétéroscédasticité et la corrélation en série lors de la sélection et de la spécification des modèles.
3-Estimation	L'estimation économétrique nécessite l'utilisation des techniques d'estimation appropriées, telles que l'estimation des moindres carrés ordinaires (OLS), l'estimation du maximum de vraisemblance (MLE) ou l'estimation de la variable instrumentale (IV). Utiliser des statistiques inférentielles appropriées pour tester les hypothèses et évaluer l'adéquation du modèle.
4-Validation et robustesse	La validation d'un modèle économétrique s'effectue à l'aide de techniques de tests hors échantillon et l'évaluation de la robustesse des résultats s'effectue en analysant la sensibilité et en explorant des spécifications de modèles alternatives.
5-Interprétation des résultats	L'interprétation des résultats de l'analyse économétrique pour fournir un contexte approprié des résultats de l'analyse économétrique et souligner les limites de l'analyse.

L'économétrie se présente comme une discipline indispensable à la croisée de l'économie et de la science des données. L'économétrie continue à se développer en particulier

Section 1 : L'économétrie et l'Intelligence Artificielle

grâce aux données aussi à l'apparition des big data, l'intelligence artificielle (IA) et l'apprentissage automatique (ML) ces technologies sont de plus en plus utilisées pour améliorer les prévisions et les analyses économiques, ce qui permet d'obtenir des prévisions plus précises et de meilleures décisions.

1.2 Méthodes traditionnelles de prévision

Les méthodes traditionnelles de prévision et d'analyse économiques s'appuient sur des modèles économétriques. Ces modèles utilisent des équations mathématiques et des méthodes statistiques pour estimer les relations de causalités et les corrélations entre les variables économiques. Ces modèles traditionnels présentent certaines limites, car ils ne tiennent pas compte à des relations plus complexes entre les variables économiques. De plus, les méthodes traditionnelles nécessitent de nombreuses données et certaines hypothèses sur la structure du modèle, ce qui peut affecter leur adaptabilité à l'évolution des conditions économiques. L'utilisation de l'intelligence artificielle et le Machine Learning en économie peut aider à surmonter ces problèmes et à améliorer la précision des prévisions sur les tendances futures.

1.3 Le Big Data et l'économétrie

Le big data à une histoire récente c'est un terme utilisé pour décrire des ensembles de données volumineux et complexes, le terme big data désigne les ressources d'informations dont les caractéristiques en termes de volume et de variété imposent l'utilisation de technologies et de méthodes analytiques particulières, et qui dépassent en général les capacités d'une seule et unique machine. il est difficiles à les traiter et à l'analyser avec des méthodes traditionnelles où l'identification et la prédiction de ces données nécessite une technique de l'analyse prédictive.

Les avantages de l'utilisation du Big Data et de l'analyse prédictive pour l'analyse économique sont nombreux. Il permet de prévoir plus précisément et plus rapidement les tendances économiques en reconnaissant rapidement des corrélations qui pourraient ne pas être apparentes avec les méthodes traditionnelles. De plus, l'analyse prédictive donne un aperçu des risques et opportunités économiques potentiels, et permettant ainsi aux décideurs politiques de prendre des décisions intelligentes.

Section 1 : L'économétrie et l'Intelligence Artificielle

1.4 Algorithmes d'apprentissage automatique (Machine Learning ML) et l'économétrie

Les algorithmes d'apprentissage automatique sont devenus de plus en plus populaires ces dernières années en raison de leur puissante capacité à traiter et analyser de grands ensembles de données. Ces algorithmes sont conçus pour identifier des modèles et des relations significatifs dans les données qui pourraient ne pas être détectés par les modèles économétriques traditionnels, qui reposent sur des hypothèses a priori sur le lien entre les variables et leur capacité de gérer les relations de données non linéaires, les modèles d'apprentissage automatique peuvent être utilisés pour de nombreuses tâches de prévision et d'analyse économiques, Les algorithmes d'apprentissage automatique peuvent également être appliqués à des ensembles de données volumineuses.

Les avantages des algorithmes d'apprentissage automatique sont nombreux notamment à l'amélioration de la précision et l'efficacité par rapport aux modèles économétriques traditionnels, ils présentent également certains inconvénients, notamment le besoin de données d'entraînement de haute qualité et les difficultés d'interprétation des modèles. Ils doivent donc être utilisés avec prudence lorsqu'ils sont appliqués à l'analyse économique.

1.4.1 Applications de l'IA dans les prévisions

Nombreux travaux de l'applications d'intelligence artificielle dans les prévisions et analyses économiques. Dans le tableau suivant quelques exemples de la manière dont cette technologie peut être utilisées pour les indicateurs économique et l'analyse financière et dans la gestion des risques :

La prévision des indicateurs économiques	L'analyse des marchés financiers	Gestion des risques
Les techniques de l'IA et les algorithmes de ML peuvent détecter des modèles à partir de données historiques, permettant des prévisions plus précises du PIB, de l'inflation et des taux de chômage.	Les techniques de l'IA et les algorithmes de ML peuvent passer au traitement des données des marchés financiers pour prédire les cours des actions, les taux de change et d'autres indicateurs.	Les techniques de l'IA et les algorithmes de ML peuvent identifier les risques économiques possibles, tels que les récessions ou les crises financières, ainsi que les risques de crédit et de marché dans les institutions financières.

Section 1 : L'économétrie et l'Intelligence Artificielle

L'utilisation des techniques de l'intelligence artificiel et les algorithmes du Machine Learning peuvent être utilisés dans plusieurs domaines que les 3 domaines situés dans le tableau ci-dessus. L'utilisation des techniques de l'IA et les algorithmes du ML pour les prévisions économiques fournit des informations indispensables sur les relations complexes tout en tenant compte des limitations potentielles telles que la qualité ou l'interprétabilité des données.

2 Les limites de l'intelligence artificiel pour l'économétrie

L'utilisation des techniques de l'IA et le ML dans le domaine de l'économétrie a apporté des bénéfices pour les prévisions et l'analyse économiques, plusieurs applications ont montré l'amélioration des résultats obtenus par ces techniques par rapport aux techniques traditionnelles, Aussi ces applications ont montré plusieurs défis et limitations qu'on doit prendre en compte. Dans les lignes suivantes on va cerner quelques limitations de l'IA et ML :

- 1- La qualité des données : la précision des modèles d'IA et de ML dépend fortement de la qualité des données utilisées pour les entraîner. Si les données sont incomplètes, erronées ou biaisées, les résultats risquent de ne pas correspondre avec précision aux facteurs sous-jacents entre les variables.
- 2- L'interprétabilité : les algorithmes d'IA et de ML peuvent identifier des modèles qui ne sont pas forcément visibles ou connus. Contrairement aux modèles économétriques créés sur la base de relations économiques connues, Il est donc difficile de comprendre ce qui motive les prédictions ou de détecter l'erreurs dans le modèle.
- 3- Le surajustement : si un algorithme est entraîné avec un ensemble de données trop petit ou spécifique, il peut finir par ajuster le bruit au lieu des modèles réels, ce qui entraîne des prévisions inexactes et limite son applicabilité générale à d'autres ensembles de données.
- 4- Les modèles de boîte noire : certains modèles d'IA et de ML fournissent des explications limitées sur la manière dont ils atteignent leurs résultats ; cela limite leur utilisation puisque l'interprétabilité peut être cruciale dans des applications spécifiques.
- 5- Le biais humain : si les données utilisées pour entraîner les algorithmes d'IA et de ML reflètent des préjugés humains (soit en raison de décisions prises par ceux qui créent les données), alors ces algorithmes sont susceptibles de perpétuer ces biais.
- 6- Les limites technologiques : l'utilisation de l'IA et du ML nécessite des ressources informatiques et une expertise importante ; cela pourrait limiter l'accès aux entreprises ne disposant pas de ressources suffisantes.

Section 1 : L'économétrie et l'Intelligence Artificielle

L'utilisation des techniques de l'IA et ML à des fins de prévision et d'analyse économiques, consiste à tenir compte des limites potentiels. Une réflexion approfondie sur ces questions peut contribuer à garantir que les modèles développés sont fiables, non influencés par les préjugés et utiles à la prise de décision.

3 Les perspectives futures de l'IA dans les recherches économétriques

De nombreuses opportunités d'innovation et d'amélioration pour l'utilisation de l'IA et le ML dans les prévisions et l'analyse économiques, voici quelques implications et évolutions à surveiller :

- Les données en temps réel deviennent de plus en plus disponibles ;
- Les algorithmes d'IA et de ML peuvent être utilisés plus efficacement et permettant des prévisions plus précises et des réponses plus rapides aux changements des conditions économiques ;
- Des recherches sont en cours sur ces techniques visant à rendre les algorithmes d'IA et de ML plus faciles à interpréter cela pourrait accroître la confiance dans ces modèles et les rendre plus utiles lors de la prise de décisions ;

L'avenir de l'IA et du ML en économétrie sera probablement caractérisé par une disponibilité accrue des données, une interprétabilité améliorée et une plus grande collaboration entre le monde universitaire et industriel, ce qui devrait se traduire par une meilleure précision lors de la prévision des tendances économiques et offrir de nouvelles possibilités.

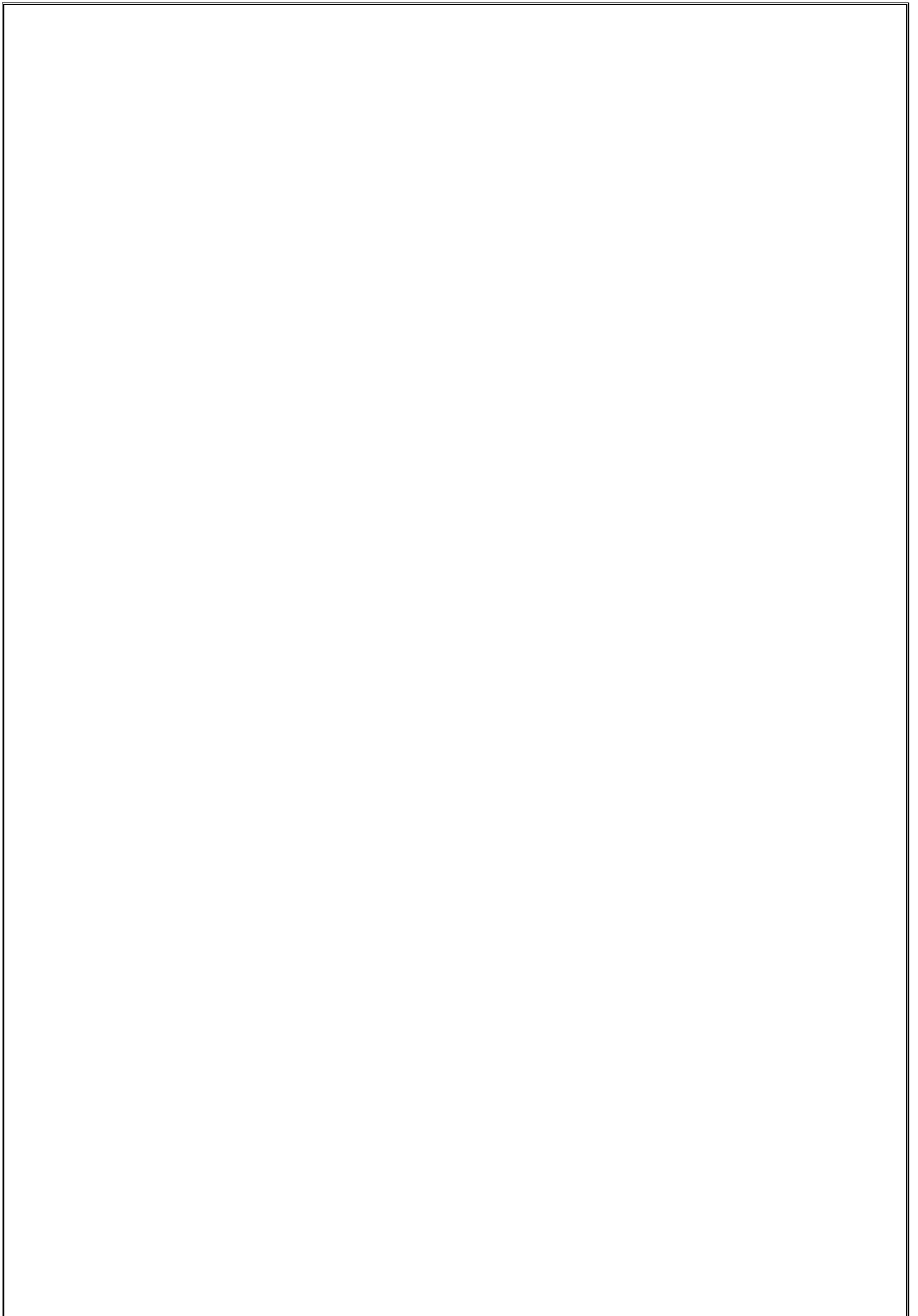
Conclusion

Malgré les défis à relever, l'avenir de l'IA et du ML dans les prévisions et analyses économiques semble prometteur. À mesure que de nouvelles techniques sont développées et intégrées à d'autres domaines de recherche, nous nous attendons à une précision et une utilité accrue des prévisions économiques. Les entreprises et les décideurs politiques devraient suivre de près ces avancées, en tenant compte des avantages et des limites potentiels lorsqu'ils prennent des décisions concernant l'avenir.

Section 2
Les Techniques Traditionnelles
de Prévision

Sommaire

SECTION 2 : TECHNIQUES TRADITIONNELLES DE PREVISION	17
INTRODUCTION.....	17
1 LES SERIES TEMPORELLES (SERIES CHRONOLOGIQUES).....	17
2 LES MODELES STOCHASTIQUE DE PREVISION	18
2.1 Modèle autorégressif (AR).....	19
2.2 Processus de moyenne mobile (MA).....	19
Un bruit blanc	19
2.3 Le processus ARMA	20
2.4 Les Modèle ARIMA	20
3 LA PREVISION PAR UN MODELE ARMA(P,Q)	21
4 LES MODELE ARCH LINEAIRE UNIVARIEE	21
4.1 Le modèle ARCH	22
4.2 Le modèle GARCH	23
4.2.1 La modélisation ARMA en présence d'effet GARCH	24
4.2.1.1 Le test d'asymétries	24
4.2.2 Les modèle GARCH Multivariés	25
4.3 Le modèle IGARCH.....	25
4.4 Le modèle EGARCH.....	25
4.5 Le modèle GJR-GARCH.....	26
4.5.1 Explication du modèle.....	26
5 LA PREVISION PAR UN MODELE GARCH (1.1).....	26
CONCLUSION	27



Section 2 : Les Techniques Traditionnelles de Prévision

Section 2 : Techniques Traditionnelles de Prévision

Introduction

Les techniques de prévision traditionnelles ont été largement utilisées et étudiées depuis des décennies. Ces méthodes se basent sur l'analyse statistique et économétrique des données historiques afin d'identifier les tendances, les cycles et les relations entre les variables économiques. Parmi les approches les plus courantes, on peut citer les modèles de régression linéaire, les modèles ARIMA, les modèles GARCH et les modèles à correction d'erreur. Ces techniques permettent de capturer la dynamique linéaire des séries temporelles et de produire des prévisions à court et moyen terme. Cependant, elles présentent certaines limites, notamment lorsqu'il s'agit de modéliser des phénomènes économiques complexes et non linéaires. L'objectif de cette section est d'explorer les différents modèles basés sur l'approche ARMA et GARCH.

1 Les séries temporelles (séries chronologiques)

Les séries temporelles concise une étude qui correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Une série temporelle est défini comme une suite finie de données indexées par le temps (x_1, x_2, \dots, x_n) nous identifions ces données en tant que « Dynamiques » car ces données sont indexées sur la base d'un attribut de temps. L'indice temps peut être selon les cas : la minute, l'heure, le jour, l'année etc. La discipline qui traite ce type de quantités est L'économétrie qui est une discipline quantitative dédiée à l'économie. L'analyse des séries temporelles en question consiste en opération de prédiction, de modélisation et caractérisation. L'objectif de prédiction est de prévoir avec précision l'évolution à cours termes de séries temporelles, et la modélisation vise la détermination des caractéristiques du comportement à long terme, la caractérisation à comme but de déterminer les propriétés fondamentales d'une série temporelle. La prévision est le sujet de ce travail. Nous rappelons que la prévision des séries temporelles a principalement débuté dans les années 70, de nombreuses techniques de prévision ont été développées pour des domaines variés que l'agriculture, la météorologie, l'automatique, la finance, etc.

Une spécificité liée à l'utilisation des séries temporelles pour la prévision est d'assurer la qualité des données fournies, remplir les données manquantes.

Section 2 : Les Techniques Traditionnelles de Prévision

Une série temporelle est constituée de quatre composantes : la tendance – le cycle – la saisonnalité – fluctuations irrégulières (Bruce L. Bowerman, 1993). L'extraction de ces composantes peut s'effectuer par la décomposition STL (Seasonal and Trend decomposition using Loess).

Pour la majorité des séries temporelles, il existe souvent une dépendance temporelle entre les observations, ce qui a des modélisations de type autorégressif : on utilise le passé pour expliquer le présent et prédire le futur. La modélisation et prédiction d'une série temporelle suppose, dans la plupart des cas, de faire des hypothèses sur son comportement ; Puisqu'il s'agit de séries non-déterministes, il va falloir que la composante aléatoire ' varie, mais pas trop'. Cette condition se traduira par la « stationnarité ». La stationnarité d'une série temporelle est une propriété importante, c'est une série temporelle qu'elles ne contiennent pas de tendance (trend) ou de saisonnalité.

« La stationnarité fait référence à une des propriétés des séries temporelles qui affiche des modèles périodiques se répétant à une fréquence constante »¹.

Soit X_t une série temporelle. X_t est dite stationnaire (stationnarité de second degré) si :

- $E[X_i] = \mu \rightarrow \forall i = 1, 2, \dots, t$
- $Var(X_i) = \sigma^2 \neq \infty \rightarrow \forall i = 1, 2, \dots, t$
- $Cov(X_i, X_{i-j}) = f_j \rightarrow \forall i = 1, 2, \dots, t; \forall j = 1, 2, \dots, t$

Si une de ces conditions n'est pas vérifiée, on parle donc d'une non stationnarité dans le temps. Une faible stationnarité est une hypothèse courante dans l'analyse des séries temporelles.

2 Les Modèles stochastique de prévision

Les modèles utilisés dans la prévision des séries temporelles linéaires sont variés et représentent différents processus stochastiques telles que les modèles autorégressifs (AR), les modèles intégrés (I) et les modèles de moyenne mobile (MA), qui dépendent linéairement de leurs points de données antérieurs. Les combinaisons de ces modèles produisent de nouveaux modèles comme moyenne mobile autorégressive (ARMA) et moyenne mobile intégrée autorégressive (ARIMA).

¹ Valentin NOËL, « Séries temporelles et réseaux de neurones récurrents », CSI 05/09/2022.

Section 2 : Les Techniques Traditionnelles de Prévision

Et parmi les autres types de modèles de séries temporelles non linéaires, il existe certains modèles pour désigner les changements de variance au cours du temps. Les modèles ARCH (AutoRegressive Conditionnel Heteroskedastic) et GARCH (Generalized AutoRegressive Conditionnel Heteroskedastic), introduits par Engle (1982) et Bollerslev (1986).

2.1 Modèle autorégressif (AR)

Une simple définition d'un modèle d'ordre autorégressif AR (p) un processus où les valeurs actuelles de la série temporelle sont exprimées en fonction linéaire des valeurs passées. P représente l'ordre du modèle autorégressif le nombre de termes retardés pris en compte. Le modèle est défini par l'équation suivante :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t \dots (1)$$

On note le modèle autorégressif d'ordre p AR(p). Où X_t est la valeur de la série à l'instant t , p est l'ordre du modèle, les ϕ_i sont les paramètres autorégressifs et ε_t représente le bruit blanc.

2.2 Processus de moyenne mobile (MA)

On appelle moyenne mobile d'ordre q , notée MA(q), exprime la valeur actuelle de la série comme une combinaison linéaire des erreurs passées, q est l'ordre du modèle de moyenne mobile le nombre de termes d'erreur passés utilisés. Le modèle défini par l'équation :

$$X_t = \sum_{i=1}^q \vartheta_i \varepsilon_{t-i} \dots (2)$$

Où les ϑ_i sont les paramètres à moyenne mobiles et les ε_{t-1} sont les bruits blancs.

Un bruit blanc

Un bruit blanc est un type de variation aléatoire dans une série, caractérisé par des valeurs qui ne sont pas corrélées avec une distribution de moyenne nulle. En d'autres termes, chaque point de bruit blanc est indépendant des autres et ne présente aucune structure ou tendance prévisible. Dans le contexte des séries temporelles, un bruit blanc est considéré comme un signal aléatoire qui n'a pas de pattern identifiable et ne contient aucune information utile pour les prédictions.

Section 2 : Les Techniques Traditionnelles de Prévision

Une série temporelle présente un bruit blanc discret si :

- Les X_t sont répartis de manière identique ;
- $cov(t_1, t_2) = 0$ lorsque $t_1 \neq t_2$;
- $cov(t, t) = \sigma_{(t,t)}^2 = \sigma^2$. où $0 < \sigma^2 < \infty$.

2.3 Le processus ARMA

Le processus Autorégressif Moyenne mobile de (Box & Jenkins, 1976), ARMA (p, q), est une combinaison des deux processus AR (q) autorégressif et Moyenne mobile (p). Comme le montre l'équation ci-dessous :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \vartheta_i \varepsilon_{t-i} + \varepsilon_t \dots (3)$$

Où les paramètres ϕ_i et ϑ_i sont des réels, et les ε_t sont les résidus (Bruit blanc) d'une variance σ^2 .

2.4 Les Modèle ARIMA

C'est un modèle de la moyenne mobile intégrée autorégressive ARIMA (p, d, q), introduit par (Salas et al. 1988), Le but de ce modèle est de capturer les relations passées (AR), les variations non stationnaires (I) et les erreurs passées (MA). Le processus des ordres p, d et q est un processus, X_t dont les différences $(1 - \beta)^d X_t$ satisfont un modèle ARMA (p, q) qui est un modèle stationnaire dans lequel d est un nombre de différenciation. Nous utilisons la notation suivante :

$$\phi_p(\beta)(1 - \beta)^d X_t = \vartheta(\beta)\varepsilon_t \dots (4)$$

Où

$$\phi_p(\beta) = 1 - \phi_1 - \phi_2 - \phi_p \dots (4.1)$$

Et

$$\vartheta_q(\beta) = 1 - \vartheta_1 - \vartheta_2 - \vartheta_q \dots (4.2)$$

Le paramètre d dans le modèle ARIMA représente la différence première de X_t pour trouver une série temporelle stationnaire comme suit :

$$X_t = \mu + X_{t-1} \dots (5)$$

Section 2 : Les Techniques Traditionnelles de Prévision

Les modèles présentés ci-dessus sont des modèles classiques d'estimation supposent que les séries temporelles ont une volatilité constante au fil du temps. Plusieurs recherches effectuées sur les séries temporelles mettent en évidence des périodes de fortes fluctuations de cours suivies de périodes un peu stables.

C'est pourquoi, les modèles établis sous les conditions de la linéarité et de la normalité et la stabilité de la variance dans le temps sont inadaptés, et cela parce qu'elles négligent la partie des informations contenues dans les séries temporelles. Une première modélisation introduite de la variance dépendante du temps est due aux travaux de Engel (1982). Et par la suite plusieurs modèles ont été élaborés.

3 La prévision par un modèle ARMA(p,q)

Le modèle ARMA est principalement utilisé pour capturer les tendances courtes et les fluctuations dans des données stationnaires. Il est efficace dans les prévisions à court terme. Pour un point futur X_{t+h} basé sur les informations disponibles jusqu'à t peut être exprimée comme suit :

On retient l'équation 3 du modèle ARMA_(p,q) :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \vartheta_i \varepsilon_{t-i} + \varepsilon_t \dots (3)$$

Prévision future de X_{t+h} :

$$\hat{X}_{t+h} = \phi_1 \hat{X}_{t+h-1} + \phi_2 \hat{X}_{t+h-2} \dots + \phi_p \hat{X}_{t+h-p} + \theta_1 \hat{\varepsilon}_{t+h-1} + \theta_2 \hat{\varepsilon}_{t+h-2} \dots + \theta_q \hat{\varepsilon}_{t+h-q}$$

ϕ_p Sont des coefficients autoregressifs ;

θ_q Sont des coefficients Moyennes mobiles ;

$\hat{\varepsilon}_t$ Sont les erreurs résiduelles (inconnues dans le futur, elles sont supposées être nulles ou remplacées par les valeurs estimées).

4 Les modèle ARCH linéaire univariée

Le modèle ARCH (AutoRegressive Conditionnel Heteroskedastic) est un modèle économétrique utilisé dans la modélisation de la volatilité d'une série temporelle. Il permet de prendre en compte l'hétéroscédasticité conditionnelle (la variance de la série n'est pas constante dans le temps), mais dépend de l'information passée. Où son idée de base est qu'une série peut

Section 2 : Les Techniques Traditionnelles de Prévision

être non corrélée ou avec un bas ordre de corrélation. L'utilisation des modèles d'hétéroscédasticité permet de capturer toute dépendance dans la série. On considère pour une série (y_t) la moyenne conditionnelle :

$$\mu_t = E(Y_t, F_{t-1}) \dots (6)$$

Où F_{t-1} , désigne l'ensemble des informations du passé jusqu'à la date $t-1$.

Et la variance conditionnelle :

$$\sigma_t^2 = var(y_t, F_{t-1}) = E[(y_t - \mu_t)^2, F_{t-1}] \dots \dots (7)$$

Si nous faisons l'hypothèse que la moyenne conditionnelle μ_t de la série y_t suit un modèle ARMA_(p,q), c'est-à-dire :

$$y_t = \mu_t + a_t, \mu_t = \varphi_0 + \sum_{i=1}^p \varphi_i y_{t-i} - \sum_{i=1}^q \theta_i a_{t-i} \dots \dots (8)$$

En combinant les équations (7-8) nous avons :

$$\sigma_t^2 = var(y_t, F_{t-1}) = var(a_t, F_{t-1}) \dots (9)$$

L'étude des modèles d'hétéroscédasticité conditionnelle est dont le but d'observer l'évolution de la variance σ_t^2 , l'évolution de cette variance distingue un modèle d'un autre pour trouver l'équation dynamique qui gouverne l'évolution temporelle de la variance conditionnelle du choc ou de l'innovation a_t .

4.1 Le modèle ARCH

La plus importante amélioration qu'a apporté l'économétrie classique en matière de prévision était due à l'utilisation de la moyenne conditionnelle.

Le modèle ARCH introduit par Engle (1982) est utilisé dans l'estimation de la volatilité quand la variance des erreurs suit un processus autorégressif. Pour illustrer ceci à travers un modèle AR (1). Soit le modèle :

$$y_t = \phi y_{t-1} + \varepsilon_t \text{ avec } \varepsilon_t \rightarrow N(0, \sigma^2)$$

On a alors

$$\varepsilon_t = u_t h_t \text{ où } u_t \rightarrow N(0,1)$$

Section 2 : Les Techniques Traditionnelles de Prévision

$$h_t^2 = a_0 + \sum_{i=0}^p a_i \varepsilon_{t-i}^2 = a_0 + a(B) \varepsilon_t^2 \dots (10)$$

Tel que

$$a_0 > 0, a_i > 0 \forall_i \text{ et } [B] = a_1 B + a_2 B^2 + \dots + a_p B^p$$

h_t^2 est appelé processus ARCH d'ordre p il est noté ARCH_(p)

Les variances conditionnelles et non conditionnelles sont indépendantes du temps, selon Engle (1982), il est possible pour des processus réels qu'on ait des meilleurs intervalles de prévision si les informations passées sont prises en considération dans la variance de prévision. Pour une modélisation à variance dépendante du temps ; il le définit sous l'équation suivante :

$$\begin{cases} a_t = \varepsilon_t \sigma_t \\ \sigma_t^2 = a_0 + a_1 a_{t-1}^2 \text{ avec } a_0 > 0, a_1 \geq 0 \\ \varepsilon_t \rightarrow N(0,1) \end{cases}$$

On dit que a_t suit un processus ARCH (autoregressive conditional heteroscedasticity) d'ordre 1. La variance conditionnelle dans ce modèle est une fonction de la dernière valeur retardée du carré du processus.

4.2 Le modèle GARCH

L'application des modèles ARCH sur certaine série temporelle a montré que cette modélisation requiert dans certain cas la prise en compte de retard de plus en plus éloignés et cela pourrait conduire à la violation de la condition de non négativité des coefficients et demande l'estimation de plusieurs paramètres. Bollerslev (1986) a développé le modèle ARCH en une nouvelle classe appelée GARCH (Generalized Autoregressive conditional heteroscedasticity), c'est une extension du modèle ARCH. Un modèle GARCH(p,q) où q est un ordre de ARCH et p est ordre de GARCH .

Ce modèle se présente comme suit :

$$y = aX + \varepsilon$$

Avec

$$\varepsilon_t = u_t \times h_t, u_t \rightarrow N(0,1) \text{ et } h_t^2 = \hat{a}_0 + a(D) \varepsilon_t^2 + \beta(D) h_t^2$$

Section 2 : Les Techniques Traditionnelles de Prévision

$$h_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 \dots (11)$$

Avec

$p > 0, q \geq 0, a_0 > 0, a_i \geq 0$ et $\beta_i \geq 0$; on dit que le processus suit un GARCH s'ordre p et q
ou $\text{GARCH}_{(p,q)}$

Si $p=0$ se confond avec ARCH (q) et si de plus $q=0$ alors a_i est un bruit blanc de variance a_0 .

4.2.1 La modélisation ARMA en présence d'effet GARCH

La présence d'un effet ARCH ou GARCH dans les perturbations d'un modèle de régression s'interprète comme un témoignage d'une mauvaise spécification selon Engle(1982).

L'introduction de la modélisation d'un modèle ARMA avec bruits GARCH est effectuée par Weiss (1984) en donnant les conditions de consistance et de normalité asymétrique des estimateurs des équations ARMA et GARCH.

Il applique le modèle sur les résidus a_t pour une série de T observations de y_t non pas sur le processus y_t . Weiss (1984) a défini le modèle comme suit :

$$\left\{ \begin{array}{l} \phi(B)(y_t - \mu) = \theta(B)a_t \\ E(a_t^2 / F_{t-1}) = h_t = a_0 + \sum_{i=1}^p a_i a_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i} \end{array} \right.$$

ϕ et θ Sont respectivement les polynômes autorégressifs d'ordre s et moyenne mobile d'ordre m de l'équation ARMA (s,m) .

4.2.1.1 Le test d'asymétries

Selon Bakhtiar et al (2020), le test d'asymétrie est une propriété qui montre un déséquilibre de certaine condition. Pour une série temporelle, les propriétés asymétriques sont appelées une forte volatilité. Pour déterminer la nature de l'asymétrie il faut voir le skewness et kurtosis. Ce test est dans le but de tester le degré du déséquilibre de distribution. Le test d'asymétrie peut être effectuer en utilisant l'intersection entre la corrélation des résidus ε_t et le carré des résidus ε_t^2

Section 2 : Les Techniques Traditionnelles de Prévision

4.2.2 Les modèle GARCH Multivariés

Suite à l'article pionnier d'Engel, plusieurs variantes du modèle ARCH(p) ont été proposées afin de donner une meilleure description et prévision de la volatilité. Et pour améliorer cette approche les chercheurs exploitent principalement deux dimensions pour améliorer les modèles GARCH. En premier lieu, ils s'intéressent aux autres distributions que la loi normale pour les innovations. Ainsi, ils recherchent des modèles plus flexibles pour mieux expliquer l'évolution de la volatilité observée dans les modèles. En règle générale, la classe des modèles ARCH peut être divisée en deux sous-ensembles : les modèle ARCH linéaires et les modèles ARCH non linéaires. Les premiers regroupent les processus ARCH (p) et GARCH (p,q) décrits précédemment. D'autres variantes comme modèle GARCH intégré (IGARCH) Engel et Bollerslev (1986), EGARCH Nelson (1991), GJR-GARCH Glosten Jagannathan et Runkle (1993).

4.3 Le modèle IGARCH

Le modèle GARCH applique les deux processus autorégressif et moyenne mobile pour la variance σ_t^2 . Utilisé pour étudier les propriétés de mémoire longue des séries temporelles financières. Il peut donner de meilleurs résultats qu'un modèle GARCH unique. Le modèle GARCH intégré est défini comme suite :

$$\varepsilon_t = \sigma_t h_t; \sigma^2 = \omega + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_j^q \beta_j \sigma_{t-j}^2 \dots (12)$$

La somme des coefficients est limitée à 1. la variable exogène peut être facilement reflétée dans les différentes spécifications du modèle GARCH seulement par l'addition du $y_t \bar{\beta}$.

4.4 Le modèle EGARCH

C'est un outil puissant pour la modélisation et la prévision de la volatilité financière introduit par Nelson (1991), il a introduit une méthode alternative en modifiant le modèle GARCH a un modèle exponentiel EGARCH. Le modèle EGARCH est spécifié comme suite :

$$\varepsilon_t = \sigma_t h_t; \ln \sigma^2 = \omega + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \ln \sigma_{t-j}^2 \dots (13)$$

Ce modèle est différent de la structure de la variance à cause du $\log \sigma^2$.

Section 2 : Les Techniques Traditionnelles de Prévision

4.5 Le modèle GJR-GARCH

Glosten Jagannathan et Runkle (1993), introduit le modèle asymétrique GARCH dite modèle GJR-GARCH. L'avantage du modèle GJR-GARCH est la possibilité de mesurer la volatilité due aux effets des différents informations (mauvaises et bonne nouvelles) c'est-à-dire chocs positif et négatif intervenus sur la série étudié. Il est différent du modèle GARCH par la présentation des propriété asymétriques. Le modèle GJR-GARCH se présente sous l'expression suivante :

$$\delta_t^2 = \omega + \sum_{i=1}^p a_i \varepsilon_t^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \sum_{i=1}^p \gamma_i I_{t-i} a_{t-i}^2 \dots (14)$$

Quand

$$I_{t-i} \begin{cases} 1 & \text{si } \varepsilon_{t-i} < 0 \\ 0 & \text{si } \varepsilon_{t-i} \geq 0 \end{cases}$$

4.5.1 Explication du modèle

a_i Est le paramètre du modèle ARCH, β_j est le paramètre du modèle GARCH jusqu'à l'observation (j) et le γ_i est le paramètre de l'effet de levier (effet asymétrique dans la dynamique de la volatilité). I_{t-i} Présente la variable muette qui signifie l'indice fonctionnel égal à 0 quand a_{t-i} est positif et égal à 1 quand a_{t-i} est négatif.

Si le paramètre $\gamma_i > 0$ donc la négativité des erreurs ne fonctionne pas et cela signifie que l'influence des mauvaises nouvelles sera supérieure à celle des bonnes nouvelles (Dritsaki, 2017). La procédure d'estimation du modèle GJR-GARCH est comme suite :

Effectuer une estimation du modèle GARCH à partir de la série temporelle étudié et puis effectuer le test de la présence des effets ARCH sur les résidus obtenus du modèle GARCH et puis réaliser un test de diagnostic pour observer la pertinence du modèle puis un test d'effet asymétrique si on trouve un effet d'asymétrie on peut donc l'utiliser pour la prévision.

5 La prévision par un modèle GARCH (1.1)

Dans un modèle GARCH, nous prévoyons la volatilité future (variance conditionnelle) à partir de la variance passée et des erreurs passées. La prévision de la variance au temps $t+h$ peut être exprimée comme suite :

L'équation du modèle GARCH (1.1) :

$$h_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Section 2 : Les Techniques Traditionnelles de Prévision

σ_{t-1}^2 La variance conditionnelle au temps t ;

ε_{t-1}^2 Le carré de l'erreur au temps $t-1$;

a_1 et β_1 Sont les coefficients du modèle GARCH.

La prévision future de la variance pour $t+h$:

$$\hat{h}_{t+h}^2 = a_0 + a_1 \hat{\varepsilon}_{t+h-1}^2 + \beta_1 \hat{\sigma}_{t+h-1}^2$$

$\hat{\varepsilon}_{t+h-1}^2$ La prévision de l'erreur au temps $t+h-1$, souvent prise comme nulle dans les prévisions futures.

$\hat{\sigma}_{t+h-1}^2$ La variance conditionnelle prédite au temps $t+h-1$.

Conclusion

Les modèles traditionnels de prévision, comme ARIMA, GARCH et ces extensions, sont des outils robustes et fiables pour l'analyse et la prédiction des séries temporelles. Les modèles classiques reposent sur des hypothèses de la stationnarité, la linéarité et la dépendance des données passées, ce qui les rend adaptés aux séries temporelles simples avec des tendances ou des cycles réguliers. Cependant, ces modèles peuvent être limités dans leur capacité de capturer les relations complexes, les non-linéarités, ou des dépendances à long terme présentes dans des séries plus complexes. De plus, ces modèles nécessitent souvent un ajustement manuel des paramètres. Malgré ces limitations, leur simplicité et leur transparence en font des solutions efficaces dans de nombreux contextes, notamment pour des prévisions à court terme ou lorsque les données sont bien structurées.

Section 3

Les Techniques de l'Intelligence Artificielle (Deep Learning)

Sommaire

SECTION 3 : LES TECHNIQUES DE L'INTELLIGENCE ARTIFICIELLE (DEEP LEARNING)	30
INTRODUCTION.....	30
1 DOMAINES D'APPLICATION	30
2 DEEP LEARNING POUR LES SERIES TEMPORELLES.....	31
2.1 Les réseaux de neurones artificiels.....	31
2.1.1 Mise en œuvre d'un réseau de neurone	32
2.1.2 Architecture d'un réseau neurone	32
2.1.3 Les réseaux de neurones non bouclés	32
2.1.4 Réseaux de neurones bouclés	33
2.1.5 Exemple de modélisation d'un seul neurone.....	34
2.1.6 L'apprentissage du réseau neurone.....	34
2.1.6.1 Algorithme d'apprentissage.....	35
2.1.6.2 Fonctionnement	36
3 LES ALGORITHMES GENETIQUES AGS	37
3.1 Présentation des algorithmes génétiques	37
3.2 Fonctionnement d'un algorithme génétique.....	38
3.3 Séries temporelles et algorithmes génétiques.....	39
4 LA MODELISATION DES SERIES TEMPORELLES BASEE SUR LES RESEAUX DE NEURONE ARTIFICIELLE.....	40
4.1 Réseaux de neurones artificiels autorégressif non-linéaire NAR.....	41
4.1.1 Réseau neurone autorégressif NAR	41
4.1.1.1 Série financière.....	43
4.1.1.2 Les caractéristiques des séries financières	43
4.2 Modèles hybridés	45
CONCLUSION	47

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

Introduction

Le Deep Learning ou l'apprentissage profond (dans ce travail on va adopter la désignation Deep Learning) est l'application du Machine Learning en utilisant des algorithmes complexe et plusieurs niveaux de calculs mathématiques qui permet l'apprentissage au logiciel à effectuer des taches plus complexes à partir d'une quantité de data (donnés) sans être guidés par l'homme.

Aussi, on peut le définir comme le système qui englobe des méthodes d'apprentissage avec des architectures complexe pour modéliser des données en combinant diverses transformations non - linaire.

$$X \rightarrow f(X, \theta) \text{ telle que } y \approx f(X, \theta)$$

1 Domaines d'application

A nos temps, le Deep Learning est considéré comme l'avenir unique de l'intelligence artificielle à cause de leurs meilleurs résultat obtenus dans plusieurs applications dans nombreux domaine il est aussi appelé à investir de plusieurs aspects de notre vie quotidienne, on va citer quelques domaines d'utilisation du Deep Learning :

- 1- Le Deep Learning est fortement présent dans la traduction et rédaction automatique à partir de fichiers d'image, la langue d'un texte en assurant une rédaction grammaticalement et orthographiquement correct.
- 2- Il est utilisé dans le domaine des finances et le marketing, ventes ; comme capteur et détecteur d'anomalie dans le domaine sensible des transactions financières, études des marchés.
- 3- La conduite automatique
- 4- Robots industriels et maintenance
- 5- La médecine

On peut noter que parmi les domaines qui utilise le Deep Learning, le domaine de l'économétrie, et cela à cause des meilleurs résultats obtenus dans l'analyse et l'études effectuer

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

par plusieurs chercheurs et économètre qui ont intégré les techniques du Deep Learning dans leurs études beaucoup plus dans l'étude des séries temporelles.

On peut faire dans cette section une simple projection sur quelques techniques du Deep Learning utilisé dans la prédiction et la modélisation des séries temporelles.

2 Deep Learning pour les séries temporelles

Depuis les années 1990, les techniques du Deep Learning font leurs entrés dans la science économique et de gestion dans le but de renforcer le processus de prise de décision chez les décideurs.

Dans cette partie de notre travail, on va présenter trois outils intelligents du Deep Learning fortement utiliser dans l'étude des séries temporelles. On premier, on va présenter les Réseaux de neurones artificiels, puis faire une brève description sur les algorithmes génétiques et finalement une présentation du modèle NAR (Autoregressif Neural Network) et son utilisation dans la modélisation des séries temporelles.

2.1 Les réseaux de neurones artificiels

Les réseaux neurones représentent la première catégorie de méthodes qui sera explorée dans cette thèse. Les réseaux de neurones sont un outil issu de l'intelligence artificielle qui a reçu une très grande attention, ils sont les bienvenus dans les applications économiques et financières, les chercheurs du domaine des fiance ont commencé d'utiliser cette technique moderne dans le but de réaliser des prévisions, des classifications et une reconnaissance générale des formes. et aussi, les réseaux de neurones sont devenus préférable pour raison qu'aucune hypothèse est t'exigé sur les variables et leurs distributions, et la possibilité d'estimer sans éliminer les composantes de la séries étudier (saisonnalité, la linéarité de la fonction de la tendance).

Un réseau de neurones est essentiellement un opérateur mathématique, dont la valeur numérique est déterminée par quelques lignes de code (programme). Il s'agit d'un ensemble de cellules, appelées neurones, interconnectées et organisées en couches. Chaque neurone représente une unité de calcul autonome, connectée aux couches précédentes.

La première couche est appelée couche d'entré (input) elle reçois des informations du monde extérieur,

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

La couche finale est appelée couche de sortie (output) contient les résultats du calcul effectué par le réseau, entre ces deux couches il y a la couche caché (hidden) qui est en fait comme une boîte noire et là où, les neurones n'ont aucun contact avec l'extérieur, et où s'effectuent tous les calculs intermédiaires et transformation.

2.1.1 Mise en œuvre d'un réseau de neurone

Trois étapes sont nécessaires pour la réalisation d'un réseau de neurone

- 1- Le choix de l'architecture ou le nombre de neurones dans la couche cachée doit être équilibré : le réseau de neurones ne doit ni être trop flexible, ni trop rigide.
- 2- Les poids sont ajustés en minimisant l'erreur d'approximation sur les données de l'ensemble d'apprentissage, de façon à ce que le réseau accomplisse la tâche souhaitée.
- 3- Le modèle est validé sur un ensemble dédié pour vérifier sa généralisation, puis testé sur des données inédites afin d'évaluer ses performances et son aptitude à une utilisation réelle.

2.1.2 Architecture d'un réseau neurone

Les réseaux neurones sont présentés par diverses architectures, chacune définissant une manière spécifique de connecter les neurones au sein des couches et les couches entre elles. Selon la littérature, on distingue deux principales structures de réseaux, basées sur le type de leurs connexions :

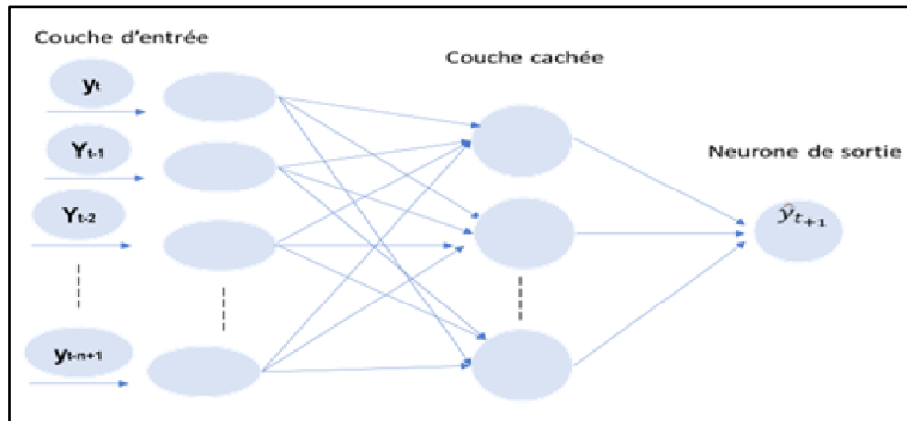
- Les réseaux de neurones statiques (ou non bouclés / non récurrents) ;
- Les réseaux de neurones dynamiques (ou bouclés/ récurrents) ;

2.1.3 Les réseaux de neurones non bouclés

Ce type d'architecture dite **feed-forward**, sont des réseaux où les connexions entre les neurones ne forment pas de cycles, permettant une propagation unidirectionnelle des informations, de l'entrée vers la sortie il est présenté par l'architecture ci-dessous :

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

FIGURE 1: Réseau Neurone à Couche



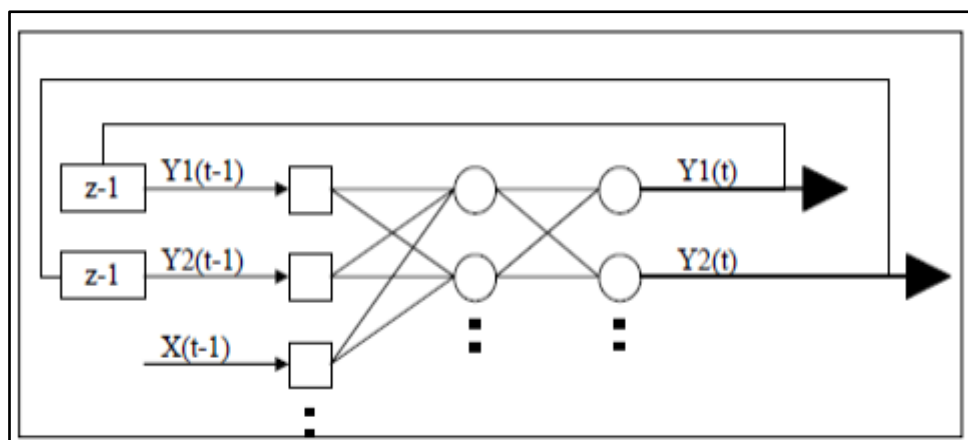
Source : Réalisée par l'auteur.

Dans un réseau de neurones, les informations sont stockées dans les poids reliant les neurones. Chaque neurone reçoit les données des neurones de la couche précédente, calcule son potentiel d'activation à l'aide d'une fonction, puis transmet une impulsion aux neurones de la couche suivante. La sortie de chaque neurone est déterminée par une fonction mathématique appliquée à la somme des entrées pondérées.

2.1.4 Réseaux de neurones bouclés

Un réseau dynamique, ou récurrent, est similaire à un réseau multicouche, mais avec des connexions qui renvoient des informations en arrière. Ces connexions peuvent relier tous les neurones ou seulement certains d'entre eux.

FIGURE 2: Réseau Neurone Bouclé



Source : Réalisée par l'auteur.

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

2.1.5 Exemple de modélisation d'un seul neurone

La première modélisation a été présentée par Mc Culloch et Walter Pitts, inspirée du fonctionnement d'un neurone biologique. Le principe de la modélisation : il existe une unité de calcul connectée à une ou plusieurs entrées grâce à des liens appelés synapses ou poids. Il produit une seule sortie, calculée à partir de la somme des entrées pondérées par une fonction mathématique.

On trouve dans le modèle les entrées qui seront notés Y qui est un N -vecteurs colonne tel que $Y = (y_1, y_2, \dots, y_N)^t$ Elles peuvent être discrètes, avec les valeurs $\{0,1\}$ ou $\{-1,0,1\}$, ou continues avec des valeurs comprises dans un intervalle donné $[0,1]$ ou $[-1,1]$. La sortie, notée O , est déterminée par le poids W associé aux connexions. Un niveau d'activation, noté S , est calculé pour le neurone et sert d'entrée à la fonction d'activation suivante :

$$S = \sum_{j=1}^N w_j y_j - b \quad b \text{ est le biais du neurone.}$$

$$O = f(S)$$

2.1.6 L'apprentissage du réseau neurone

C'est une phase essentielle qui permet le développement du réseau neurone durant la modification du comportement du réseau jusqu'à l'obtention du comportement désirées, l'algorithme d'apprentissage consiste à ajuster les poids synaptiques du réseau neurone d'une manière à produire correctement la sortie associée à chaque point de l'espace d'entrée.

Afin de découvrir la forme de la relation entre variables ou autrement dit l'apprentissage de corrélation formulée par Donald Hebb en 1949, on distingue deux types d'apprentissage :

- 1- **L'apprentissage supervisé** : Ce type d'apprentissage repose sur l'utilisation d'un échantillon d'apprentissage où le réseau apprend à reconnaître des formes. Les paramètres du réseau sont ajustés pour minimiser une mesure spécifique de l'erreur de sortie.
- 2- **L'apprentissage non-supervisé** : ce type d'apprentissage consiste à présenter au réseau un échantillon contenant des données sans informations ou catégories clairement indiquées. Le réseau organise les entrées en classes de manière autonome, en cherchant à minimiser un critère de performance.

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

3- **Apprentissage par renforcement** : C'est une méthode d'apprentissage qui est proche de l'apprentissage supervisé, mais au lieu de fournir des résultats précis, le réseau reçoit un score ou une récompense qui évalue sa performance après plusieurs essais. Les algorithmes de renforcement sont particulièrement utilisés dans les systèmes de contrôle et la prise de décision.

2.1.6.1 Algorithme d'apprentissage

L'algorithme d'apprentissage est une technique mathématique utilisée dans l'ajustement des paramètres du modèle, afin qu'il apprenne à partir des données d'entraînement et optimise sa capacité à prédire ou à classer correctement cela via :

- La modification des poids de connexions pour trouver une solution permettant le réseau à accomplir la tâche désirée.
- L'identification et l'optimisation des valeurs des poids du réseau de manière à minimiser l'erreur, améliorant ainsi la précision des prédictions du modèle.
- Permettre la prédiction de la valeur cible étant donnée une ou des valeurs d'entrées.

L'apprentissage vise à déterminer S neurones ou vecteurs codes qui résument les informations de l'ensemble des entrées, pour chaque entrée un vecteur code le plus proche y est associé. D'où, pour deux entrées associées au même vecteur code sont regroupées dans une même classe. Les entrées sont donc réparties en S classes. Un algorithme de Kohonen peut être décrit comme suit :

- Initialiser les poids $W_i(t)$ par des petites valeurs aléatoires
- Fixer $\sigma(0)$, h_0 et t_{\max}
- $t \leftarrow 0$ — Répéter tant que $t \leq t_{\max}$
- Choix aléatoire d'une entrée $Y(t)$
- Déterminer le neurone vainqueur u_c à l'aide de l'équation
- Mis à jour des poids selon l'équation
- $t \leftarrow t + 1$

Parmi les algorithmes d'apprentissage les plus utilisés :

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

- **Levenberg–Marquardt algorithm** : c'est une méthode d'optimisation utilisée principalement pour résoudre des problèmes de moindres carrés non linéaires. Cet algorithme combine les avantages de la méthode de Gauss-Newton et de la descente de gradient, offrant un compromis entre la vitesse de convergence et la stabilité. Il est souvent utilisé pour l'entraînement, notamment dans les situations où la fonction d'erreur est non linéaire. Il est très efficace dans les données qui présente un bruit ou lorsque la solution optimale doit être trouvée rapidement.
- **Bayesian Regularization algorithm** : L'algorithme de régularisation bayésienne est une méthode utilisée pour améliorer les performances des réseaux de neurones en évitant le surapprentissage. Il modifie la fonction d'erreur en ajoutant un terme de pénalisation basé sur les poids, ce qui permet de trouver un équilibre entre l'ajustement précis aux données et la simplicité du modèle.
- **Scaled conjugate gradient algorithm** : L'algorithme du gradient conjugué à échelle (Scaled Conjugate Gradient, SCG) est une technique d'optimisation pour l'entraînement des réseaux de neurones. Il optimise la descente de gradient classique en adaptant la taille des pas de manière dynamique et en contournant le calcul direct de la matrice Hessienne, ce qui le rend plus rapide et plus performant.

2.1.6.2 Fonctionnement

Pour réaliser un réseau de neurones, il est nécessaire de :

- Identifier les entrées et les sorties clés qui ont un impact significatif sur le phénomène à modéliser.
- Rassembler les données requises pour l'entraînement et l'évaluation des performances du réseau de neurones.
- Définir le nombre de neurones cachés requis pour assurer une approximation satisfaisante.
- Réaliser l'étape de l'apprentissage
- Évaluer les performances du réseau de neurones après l'entraînement.

Les réseaux de neurones artificiels, sont utilisés dans plusieurs applications, ils sont considérés comme un outil incontournable qui a prouvé son efficacité en pratique, grâce à ses capacités de généralisation et sa robustesse face au bruit, séduisant les chercheurs en quête de perfection.

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

3 Les algorithmes génétiques AGs

Une voie de recherche prometteuse porte sur l'utilisation des algorithmes génétiques pour résoudre des problèmes économétriques et financiers. Cette méthode intelligente est utilisée pour : trouver les modèles économétriques les mieux adaptés à la modélisation des séries temporelles, estimer les valeurs des coefficients et des paramètres de régression non linéaire, ainsi qu'optimiser les décisions financières. Les Algorithmes génétiques ont été développés par John Holland en 1975. Leur popularisation a été après leur publication au livre de Goldberg en 1989. Les algorithmes génétiques ont un large éventail d'applications : en économie pour l'optimisation des fonctions [De Jong (1980)], en finance [Pereira (2000)], et en théorie du contrôle optimal [Krishnakumar et Goldberg (1992), Michalewicz, Janikow et Krawczyk (1992), et Marco et al. (1996)]. Leur popularité et leur efficacité dans de nombreux domaines sont dues à la simplicité de leur mécanisme et à la puissance de leurs applications.

3.1 Présentation des algorithmes génétiques

Selon Lerman et Ngouenet (1995) un algorithme génétique est défini par :

- **Individu/chromosome/séquence** : désigne une réponse possible qui pourrait résoudre le problème, et qui est examinée pour évaluer son adéquation à l'objectif visé ;
- **Population** : représente un groupe de candidats possibles, chacun étant une solution potentielle au problème à résoudre, explorée dans le cadre de l'algorithme génétique ;
- **Environnement** : l'espace de recherche ;
- **Fonction de fitness** : la fonction positive que nous cherchons à maximiser.

À la différence d'autres méthodes, les algorithmes génétiques fonctionnent sur une population de solutions, et non sur une seule solution, tout en maintenant une taille de population constante. Cette approche repose sur le codage des paramètres plutôt que sur les paramètres eux-mêmes. De plus, les algorithmes génétiques utilisent directement les valeurs de la fonction à optimiser, sans se baser sur ses dérivées (Vallée & Yildiwoglu, 2001).

La population utilisée par l'algorithme génétique est composée d'individus, où chaque individu représente le codage d'une solution possible au problème à résoudre. Cette solution est exprimée sous la forme d'une chaîne de caractères, également appelée chromosome (ou individu/séquence). Chaque caractère de la chaîne représente un gène, et chaque lettre de

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

l'alphabet correspond à un allèle. Un algorithme génétique se base sur les phases suivantes (Vallée & Yildizoglo, 2001) :

- 1- **La phase d'initialisation** : cette phase s'agit de générer aléatoirement une population d'individus de taille donnée ;
- 2- **La phase d'évaluation** : c'est la phase où chaque chromosome est décodé puis évalué ;
- 3- **La phase de sélection** : cette phase est pour but de créer une nouvelle population de N chromosomes ;
- 4- **Reproduction** : Il s'agit de combiner deux individus sélectionnés lors de la phase précédente pour générer deux nouveaux individus. Cette recombinaison peut inclure des processus de mutation ou de croisement au sein de la nouvelle population, afin d'introduire de la diversité et d'explorer de nouvelles solutions possibles.
- 5- **La phase de retour** : Lors de la phase de décodage et d'évaluation des chromosomes, chaque individu est examiné pour déterminer la qualité de sa solution en fonction de la fonction de fitness. Ce processus continue jusqu'à ce que le critère d'arrêt soit atteint, comme un nombre d'itérations prédéfini ou une solution optimale trouvée.

3.2 Fonctionnement d'un algorithme génétique

La réussite d'algorithme génétique s'arrête sur les trois principaux opérateurs suivants :

- 1- **Opérateur de sélection** : La sélection des chromosomes est la première étape d'un algorithme génétique. La méthode la plus utilisée, la roue de loterie biaisée (Roulette Wheel), sélectionne les chromosomes en fonction de leur valeur d'adaptation. Plus un chromosome est adapté, plus il a de chances d'être choisi pour former la nouvelle population.
- 2- **Opérateur de Croisement** : L'opérateur de croisement permet de créer de nouveaux individus en échangeant des informations entre les chromosomes (individus). Un individu sélectionné pour la reproduction ne subit pas toujours un croisement, car celui-ci a lieu avec une probabilité déterminée. Plus cette probabilité est élevée, plus la population subira des changements.

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

3- Opérateur de Mutation : Cet opérateur est essentiel car il a pour but de modifier de manière aléatoire, avec une certaine probabilité, la valeur d'un composant de l'individu, introduisant ainsi de la diversité dans la population.

3.3 Séries temporelles et algorithmes génétiques

Les algorithmes génétiques présentent un grand potentiel pour résoudre des problématiques en économétrie. Leur efficacité a été comparée à celle d'autres méthodes de résolution numérique à travers divers modèles non-linéaires. Les économètres ont démontré que ces algorithmes sont tout aussi performants, tout en nécessitant moins d'informations analytiques.

Les recherches dans ce domaine où cette technique est utilisée sont très diverses on peut noter comme exemple :

- **Pan et al** en 1995, utilisent les algorithmes génétiques afin de trouver les paramètres optimaux de régression non-linéaires ;
- **Boné et al** en 1998, utilisent les algorithmes génétiques pour trouver la forme économétrique théorique la mieux adaptée à la modélisation de séries temporelles ;
- **Slimane et al** en 1998, utilisent un modèle d'algorithme génétique hybride, appelé **GHOSP**, qui incorpore une recherche de type gradient au sein de la boucle principale de l'algorithme génétique.
- **Weiss** en 1999 et **Weiss et Hirsh** 1998,1999,2000, utilisent les algorithmes génétiques afin de prévoir des événements rares.

Les algorithmes génétiques non restreints, également appelés programmation génétique, offrent un grand potentiel pour rechercher des formes fonctionnelles qui reproduisent au mieux une série de données. Au lieu de se limiter à trouver les coefficients optimaux d'un modèle linéaire pour ajuster une série empirique, ces algorithmes permettent une exploration plus flexible des solutions.

Leur efficacité peut être considérablement améliorée en intégrant un mécanisme inductif qui guide leur exploration en fonction de la représentation de l'environnement. Un tel mécanisme inductif pourrait être fourni par un réseau de neurones artificiels, qui offre à l'algorithme génétique une représentation du problème à résoudre ou de l'environnement d'apprentissage de l'agent.

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

4 La modélisation des séries temporelles basée sur les réseaux de neurone artificielle

Avec l'évolution et la variété des recherches dans le domaine de prévision des séries temporelles où le lien entre la variable cible et les variables prédictives est modélisé d'une façon non-linéaire.

Tous les modèles précédemment introduit sont devenus limités par rapport à l'innovation scientifique et les nouvelles technologies intelligentes, plusieurs études ont été effectuées pour trouver une solution à ce problème de limitation des modèles traditionnels de prévision des séries temporelles, les Réseaux de neurones artificiels figurent parmi les approches les plus efficaces proposées dans la littérature. Depuis quelques années, les réseaux de neurones artificiels ont attiré une attention grandissante de la part des économistes et des financiers en raison de leur potentiel de détection et de reproduction des relations linéaires et non linéaires, ainsi que pour des applications telles que les diagnostics économiques et financiers, la prévision et la modélisation des marchés, la sélection d'investissements, etc.

Cette méthodologie est valable dans la modélisation et la prévision des séries temporelles caractérisées par une forte volatilité, elle permet une estimation avec une préservation d'une longue mémoire de la série étudiée et une fiabilité en matière de prévision à moins de paramètres.

Le principal avantage des réseaux de neurones artificiels réside dans leur capacité à apprendre les dépendances entre les variables à partir d'un nombre limité d'observations. Cette caractéristique leur permet de modéliser des relations complexes sans nécessiter d'hypothèses préalables sur la distribution des variables.

Plusieurs études ont utilisé les **réseaux de neurones artificiels** pour la prévision de séries temporelles, avec de grands succès dans des domaines tels que la prévision des prix des actions en bourse (Refenes et al. 1994). Ils ont également été appliqués à la prévision de la demande en eau (Cubero 1991), des précipitations (French et al. 1992), de la qualité de l'eau (Maier et Dandy 1996), du débit (Karunanithi et al., 1994 ; Hsu et al. 1995 ; Smith et Eli, 1995 ; Zhang et Trimble, 1996), ainsi qu'à la gestion en temps réel d'un réservoir (Sakakima et al. 1992).

Dans toutes ces applications, un **réseau de neurones perceptron** à une couche intermédiaire a été utilisé. Lapedes et Farber (1987) ont démontré que ce type de réseau est un

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

modèle autorégressif non linéaire. Des architectures plus complexes, appelées **réseaux récurrents** (Williams 1990), constituent une classe de modèles autorégressifs non linéaires de type ARMA.net. Nous commencerons par détailler ce premier type de réseau, avant de discuter de leur calibration.

4.1 Réseaux de neurones artificiels autorégressif non-linéaire NAR

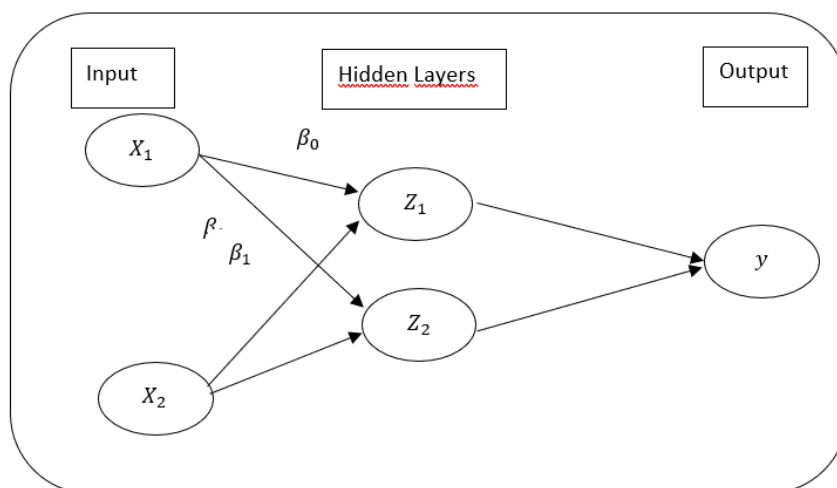
Le modèle réseau neurone artificiel autorégressif noté dans la suite du travail par AR.net est considéré comme réseau neurone récurrent fait l'objet d'un modèle autorégressif. Lapedes et Farber (1987) ont démontré que le réseau perceptron à une couche intermédiaire constitue un modèle autorégressif non linéaire. De plus, Chakraborty et al. (1992) ont prouvé que, pour des séries réelles, cette approche pouvait surpasser les modèles linéaires AR et ARMA dans le cadre multivarié.

4.1.1 Réseau neurone autorégressif NAR

On peut considérer le réseau neurone comme modèle autorégressif non-linéaire contient des variables entrantes (input) : $y_{t+1}, y_{t+2}, \dots, y_{t+p}$. Et valeurs de variable sortante (output) : $x_{t+1}, x_{t+2}, \dots, x_{t+h}$. Et les erreurs.

La figure (03) réseau neurone autorégressif.

FIGURE 3:Réseau Neurone Autorégressif NAR



Source : Réalisée par l'auteur.

- X sont les variables indépendantes (input) ;
- Z les couches cachés où l'apprentissage du réseau est effectué ;

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

- Y variable dépendante (output).

On peut considérer le réseau neurone comme modèle autorégressive :

$$Y = \beta X + \varepsilon \dots (1)$$

Où

Y= vecteur des observations de la variable dépendante de (n×1) degré

X= la matrice des observations de la variable explicative de (n×K) degré de nombre K-1

β = vecteur des paramètres de nombre K

ε = vecteur des erreurs aléatoires de (n×1)

Où $E(\varepsilon/X) = 0$, $g(X, W)$ est l'approximation de la fonction Y où W sont le vecteur des poids.

La prévision de la valeur Z dans la présence de la fonction $g(X, W)$.

On peut déterminer deux types d'erreurs : erreurs aléatoire et l'erreur d'approximation, qui représente la différence entre les valeurs réelle et valeur estimée.

Les paramètres de la fonction $g(X, W)$ constitue l'architecture du réseau neurone qui sont la combinaison entre la fonction linéaire et non- linéaire. Le modèle AR.net :

$Z = g(X, W) + \varepsilon$ Où X : sont les Input du réseau, Y : sont les Output du réseau avec une seul couche caché.

On peut écrire le modèle linéaire :

$$Y = Z \cdot W_2 + \varepsilon_2$$

Où

$$W_2 = \begin{bmatrix} \beta_4 \\ \beta_5 \end{bmatrix}$$

$$Y = Z \cdot W_1 + \varepsilon_1$$

$$W_1 = [\beta_0 \beta_1 \beta_2 \beta_3],$$

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

β, W_1, W_2 Sont les poids du réseau

On met la fonction Y à la place de Z on obtient :

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

$$Y = (XW_1 + \varepsilon_1)W_2 + \varepsilon_2$$

$$Y = XW_2W_1 + \varepsilon_1W_2 + \varepsilon_2 \text{ Avec } \beta = W_2W_1, \varepsilon = \varepsilon_1W_2 + \varepsilon_2$$

On constat que : $Y = X\beta + \varepsilon$

Ainsi, nous avons obtenu le modèle général de régression multiple, où on peut ajouter plusieurs couches cachées pour obtenir l'architecture précédente de la figure (1).

Plus de modèles basés sur les réseaux de neurones sont proposé par les chercheurs qui ont prouvé leur efficacité en prévision, avant d'approfondir dans le sujet on va donner quelques définitions :

4.1.1.1 Série financière

Une série financière est une séquence de données, indexées par le temps, telles que la minute, l'heure, le jour ou l'année. L'objectif de la prédiction est de prévoir avec précision l'évolution à court terme de cette série, tandis que la modélisation cherche à déterminer son comportement à long terme. La caractérisation, quant à elle, vise à identifier les propriétés fondamentales d'une série temporelle. La prévision de l'évolution d'un marché financier à partir de l'historique des cours de Bourse nécessite des techniques robustes pour gérer la forte volatilité de ces séries.

4.1.1.2 Les caractéristiques des séries financières

La plupart des modèles de séries financières supposaient que les résidus pouvaient être considérés comme un bruit blanc gaussien. Cependant, il est désormais reconnu que cette hypothèse est incorrecte. En réalité, les séries financières présentent plusieurs caractéristiques distinctes du bruit blanc gaussien. Parmi les plus marquantes, on retrouve :

1- **Leptokurtique** : ce terme signifie qu'il y a un coefficient d'aplatissement plus important que celui du bruit blanc gaussien.

2- **Hétéroscédasticité** : La variance des résidus qui varie au cours du temps.

3- **Autocorrélation** : Les résidus ne montrent qu'une faible auto -corrélacion ce qui signifie qu'il n'y a pas de relation significative entre les valeurs successives des résidus, tandis que les carrés des résidus sont auto-corrélés à un niveau significatif.

Les séries financières sont caractérisées par une volatilité importante qui est un indicateur de l'incertitude liée à la rentabilité de l'actif. Cette volatilité joue un rôle clé au sein

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

des modèles de risque. Mais, la volatilité n'est pas directement observée, avant d'estimer une quelconque modèle l'exemple de la crise financière de 2007-2008 n'était pas prévue par la plupart des modèles de risque en raison d'estimations inexactes produites par les modèles de prévision de la volatilité. Il est donc nécessaire de sélectionner un proxy (un réseau informatique) de volatilité. La méthodologie proposée et les principales familles de modèles de prévision de la volatilité sont les modèles stochastique GARCH car cette famille de modèles est probablement la plus utilisée dans la littérature en raison de sa capacité à s'adapter à la volatilité observée empiriquement dans les séries financières (Mandelbrot 1963) et la deuxième méthodologie est les techniques de l'intelligence artificielle.

La première méthodologie, Les modèles GARCH classiques se sont révélés trop rigides pour ajuster des séries de rendements, en particulier sur une longue période, car la persistance estimée des variances conditionnelles sont proches de 1. Par conséquent, des modèles GARCH plus flexibles ont été développés afin de surmonter ce problème EGARCH, GIR-GARCH, ZD-GARCH...

La deuxième méthodologie présentée est le Machine Learning, qui comprend un ensemble de techniques utilisé pour analyser l'évolution future des cours boursiers et de la volatilité : Support Vector Machine (Gestel et al. 2001), hidden Markov models (Gupta and DHINGA 2012 and Dias et al. 2019), Réseaux neurones artificielle RN. Ces algorithmes essayez d'apprendre automatiquement et de reconnaître des modèles dans une grande quantité de données (Krollner et coll. 2010). De plus, les RN ont été appliqués avec succès à des séries financières volatile.

Malgré les performances élevées des RN, les prédictions dérivées de l'utilisation de cet algorithme pourraient être inexactes lorsque les cours boursiers évoluent fortement (Patel et Yalamalle. 2014). Pour résoudre ce problème, les Réseaux de Neurones ont été combinés avec d'autres modèles statistiques (Kristjanpoller et al. 2014) créant ce que l'on appelle les modèles hybrides. L'hybridation peut être défini comme une approche dans laquelle plusieurs modèles sont fusionnés pour former un nouveau modèle amélioré afin de produire des meilleurs résultats de prévision. Un modèle hybride est donc une combinaison des techniques d'intelligence artificielle avec certains composants d'un modèle de prévision traditionnels comme ceux présentés au sein de la famille GARCH.

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

Récemment, plusieurs études ont été faites sur l'efficacité de la combinaison entre les réseaux de neurones et les modèles traditionnelles comme GARCH dans l'étude de la volatilité observée dans la plupart des séries financières. Ce type de modèles surmontent les inconvénients des modèles GARCH comme la linéarité, les modèles de réseaux neurones peuvent capter la non-linéarité et effectuent une prévision sans la nécessité de cette hypothèse.

Plusieurs chercheurs ont confirmé que les modèles hybridés améliorent la précision de la prévision et ils ont vérifié leurs capacités de prévoir la volatilité des séries financières. Dans la suite du travail, on va présenter la combinaison d'un RN et le modèle GARCH.

4.2 Modèles hybridés

On considère trois types de modèles de la famille GARCH hybridés le modèle EGARCH et GJR-GARCH et FIGARCH les derniers modèles intègrent une fonctionnalité importante pour l'analyse des séries temporelles à savoir la longue mémoire dans la variance conditionnelle. On a déjà vu les modèles EGARCH et le GJR-GARCH dans la section précédente et pour le modèle FIGARCH développé par **Baillie et al en 1996** pour couvrir l'hypothèse de la positivité et la stationnarité des paramètres du modèle GARCH qui suppose seulement une volatilité à court-terme.

Le modèle développé est appelé Fractionally integrated GARCH (FIGARCH) permet aux ordres fractionnaires d'intégration de varier entre zéro et un (0,1).

Dans notre travail, on va parler des deux premiers types : le premier type un modèle consiste à établir le résultat de la modélisation GARCH dans le modèle du réseau neurone comme une variable d'entrée (input) appelé ANN-GARCH dans l'intention d'augmenter la performance de la prévision de la volatilité. Et le deuxième type considère la couche de sortie du réseau neurone (output) comme une variable du modèle GARCH appelé le GARCH-ANN cela peut mieux comporter dans la prédiction de la volatilité.

Le modèle GARCH-NN est exprimé Comme suite :

Le modèle stochastique X_t avec un effet GARCH (p,q) si

$$X_t = \sigma_t Z_t$$

Où
$$\sigma_t = \sqrt{\omega + \sum_{i=1}^p \phi_i X_{t-1}^2 + \sum_{i=1}^q \theta_i \sigma_{t-1}^2}$$

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

L'écart type conditionnelle standard du modèle stochastique X_t donne les valeurs passées du processus σ_t qui sont injecté dans les valeurs présentes et Z_t présente un bruit blanc gaussien par unité de variance.

Le modèle NN est représenté par l'équation suivante :

$$a_j = \sum_{i=1}^d w_{ji} x_i$$

J c'est l'unité des couches cachés, on transforme la somme linéaire de l'équation (N) en utilisant la fonction logique d'activation

$$g(a): h_j = g\left(\sum_{i=1}^d w_{ji} x_i\right)$$

Où $g(a) = \frac{\exp(a)}{(1+\exp(a))}$

Le neurone de la couche de sortie (output) est donnée par :

$$y_1 = g\left(\sum_{j=1}^3 w_{1j} g\left[\sum_{i=1}^d w_{ji} x_i\right]\right)$$

Donc le modèle GARCH-NN s'écrit comme suite :

$$\sigma_t^2 = \omega + \sum_{i=1}^q \phi_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \varphi_j \sigma_{t-j}^2 + \sum_h^d \xi_h y_t$$

- **Un modèle EGARCH-NN :**

$$\log(h_t) = \omega + \sum_{j=1}^q \phi_j \log(h_{t-i}) + \sum_{i=1}^p \varphi_i \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| + \sum_{k=1}^r \gamma_k \frac{\varepsilon_{t-k}}{\sqrt{h_{t-k}}} + \sum_h^s \xi_h \Psi(Z_t \lambda_h)$$

- **Un modèle GJR-GARCH-NN :**

$$h_t = \omega + \sum_{m=1}^p a_m \varepsilon_{t-m}^2 + \gamma \varepsilon_{t-1} I_{t-n} + \sum_{i=1}^p \varphi_i \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| + \sum_h^s \xi_h \Psi(Z_t \lambda_h)$$

$$\Psi(Z_t \lambda_h) = \left[1 + \exp\left(\lambda_{h,d,w} + \sum_{d=1}^v \left[\sum_{w=1}^m (\lambda_{h,d,w} Z_{t-d}^w) \right] \right) \right]^{-1}$$

Section 3 : Les Techniques de l'Intelligence Artificielle (Deep Learning)

$$\text{Où } z_{t,d} = [\varepsilon_{t-d} E(\varepsilon)] / \sqrt{E(\varepsilon^2) \cdot \frac{1}{2} \lambda_{h,d,w}} \text{ tend vers } [-1, +1]$$

Conclusion

Le Deep Learning révèle les gemmes cachées des données de séries temporelles. Comprendre les défis uniques et sélectionner des algorithmes appropriés permet des prédictions précises. Grâce à un prétraitement minutieux des données, à l'ingénierie des fonctionnalités, à la sélection de modèles et à l'évaluation des performances, nous exploitons la puissance du Deep Learning pour maîtriser le temps.

PARTIE II

Pratique

Introduction de la partie pratique

Partie II : Pratique

Introduction

Dans cette partie, nous mettons en œuvre les techniques et approches théoriques étudiées précédemment pour tester leur efficacité dans la prévision des séries temporelles, particulièrement dans le domaine des séries financières. L'objectif est d'appliquer les modèles d'intelligence artificielle, en les combinant avec des méthodes traditionnelles, afin de vérifier leur capacité à prédire la volatilité des marchés financiers, notamment les fluctuations des prix du pétrole.

La partie pratique est divisée en quatre sections nous commençons une section descriptive de la volatilité des cours pétroliers un simple historique des fluctuations et une présentation des différentes techniques utilisées dans le traitement de la volatilité pétrolière avec des études précédentes.

Une deuxième section présente une application d'une modélisation GARCH, traitement des données et les prévisions.

Une troisième section l'utilisation d'une technique de l'IA réseau de neurone récurrent NAR sur les données historiques des prix du pétrole discussion des résultats de la modélisation et de la prévision

La dernière section est une application d'un modèle hybride combinant les deux techniques précédente GARCH-NN, une présentation des étapes et la discussion des résultats.

Cette approche pratique vise non seulement à valider les résultats théoriques, mais aussi à apporter une contribution concrète à l'amélioration des méthodes de prévision dans des environnements complexes et volatils.

Section 1

Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

Sommaire

SECTION 1 : LES TECHNIQUES DE L'INTELLIGENCE ARTIFICIELLE UTILISEES DANS LA PREVISION DES PRIX DU PETROLE	53
INTRODUCTION.....	53
1 HISTORIQUE DES FLUCTUATIONS DES PRIX 2020-2022	54
2 LA VOLATILITE DES PRIX DU PETROLE	55
3 LA PREVISION DES PRIX DU PETROLE	56
4 L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE POUR PREDIRE LA VOLATILITE DES PRIX DU PETROLE	57
1- Modèles ARIMA-GARCH	58
2- Modèles SVM-ANN	59
3- Modèles VARX-MLP	60
4- Modèles LSTM-CNN	60
5- Les modèles GARCH-NN	61
6- Le modèle NN-GARCH	62
CONCLUSION	64

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

Section 1 : Les Techniques de l'Intelligence Artificielle Utilisées dans la Prévision des prix du pétrole

Introduction

Analyser et anticiper la volatilité des prix du pétrole est essentiel pour les acteurs des marchés énergétiques, tels que les investisseurs, les entreprises pétrolières, les gouvernements et les analystes financiers. En tant que ressource clé de l'économie mondiale, le pétrole joue un rôle central dans les secteurs industriels, les transports et la production d'énergie. Les variations de ses prix, ou sa volatilité, influencent directement les choix stratégiques, les orientations économiques et les perspectives financières. Cette instabilité peut avoir des répercussions majeures sur les économies nationales et mondiales, ainsi que sur les marchés financiers. Comprendre et prévoir la volatilité des prix du pétrole est essentiel pour plusieurs raisons :

- 1- Gestion des Risques :** Les prévisions de volatilité des prix du pétrole aident les entreprises pétrolières, les gouvernements et les investisseurs à élaborer des stratégies de gestion des risques, de couverture et d'optimisation de portefeuille.
- 2- Planification Stratégique :** Les prévisions de volatilité sont utilisées par les décideurs politiques et les planificateurs économiques pour élaborer des politiques énergétiques et évaluer les impacts économiques et planifier les investissements dans les infrastructures énergétiques.
- 3- Décisions d'Investissement :** Les prévisions de volatilité sont utilisées par les investisseurs financiers pour prendre des décisions d'investissement éclairées dans les marchés pétroliers et déterminer les opportunités de rendement et de diversification de portefeuille.

Dans cette section, on va faire une projection sur un historiques des mouvements des prix du pétrole au cours de ces quatre dernières années, les principaux facteurs qui ont un impact sur ces fluctuations et les modèles les plus utiliser pour étudier et prévenir la volatilité des prix du pétrole.

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

1 Historique des fluctuations des prix 2020-2022

Depuis 2020, le marché mondial du pétrole a été le théâtre de fluctuations significatives des prix, influencées par plusieurs facteurs économiques, politiques et géopolitiques. L'année 2020 a été marquée par une chute spectaculaire des prix du pétrole, principalement en raison de la pandémie de COVID-19 qui avait un impact important et a entraîné une baisse radicale de la demande mondiale de pétrole, et cela après les restrictions de déplacement et les mesures de confinement étaient mises en place à travers le monde.

Cette baisse de la demande a été aggravé par une guerre des prix entre les principaux producteurs de pétrole, notamment l'Arabie saoudite et la Russie, qui ont inondé le marché de pétrole alors que la demande chutait, entraînant une surabondance de l'offre et une chute brutale des prix. En avril 2020, les prix du pétrole ont même brièvement chuté en zone négatif pour la première fois de l'histoire, ce qui signifie que les producteurs payaient pour se débarrasser de leur pétrole excédentaire.

Au cours de l'année 2021, les prix du pétrole ont rétabli à mesure que l'économie mondiale commençait à se remettre de la pandémie. Cependant, la reprise a été inégale, et les prix ont continué à être influencés par des facteurs tels que les fluctuations de la demande, les décisions de l'OPEP (l'Organisation des pays exportateurs de pétrole et ses partenaires) concernant la production, ainsi que les développements géopolitiques dans des régions productrices de pétrole clés comme le Moyen-Orient.

En 2022, les prix du pétrole ont continué à être volatils, avec des fluctuations dues à des événements tels que les tensions géopolitiques accrues, les perturbations de l'offre, et les fluctuations de la demande mondiale alors que l'économie continuait de se remettre des conséquences de la pandémie.

En 2023, les prix du pétrole ont connu une forte volatilité, influencée par les tensions géopolitiques comme les restrictions sur les exportations russes, notamment le plafonnement des prix imposé par le G7 et l'Union européenne, ce qui y ont créé une incertitude sur le marché du pétrole mondial. Cela a conduit à des fluctuations importantes des prix des cours pétrolières, aggravé par les annonces de réductions potentielles de la production de pétrole par la Russie et par l'augmentation de la demande chinoise étant un acteur majeur du marché énergétique

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

mondial après la levée des restrictions liées aux COVID-19. Cette reprise a contribué à soutenir les prix internationaux. Cette année a connu des perturbations dues à des phénomènes climatiques comme les ouragans et la baisse de l'offre de pétrole brut aux États-Unis ce qui a ajouté de l'instabilité dans le marché pétrolier.

2 La volatilité des prix du pétrole

Depuis 2020, les prix du pétrole ont connu de fortes variations dues à plusieurs facteurs tels que les changements économiques, les décisions politiques et les crises sanitaires à l'échelle mondiale. Ces fluctuations reflètent les impacts combinés de divers événements qui ont perturbé l'équilibre entre l'offre et la demande.

La volatilité des prix du pétrole fait référence à la mesure dans laquelle les prix du pétrole brut fluctuent sur une période donnée. Cette volatilité peut être influencée par une multitude de facteurs, notamment :

- ❖ **Offre et demande** : Les déséquilibres entre l'offre et la demande de pétrole brut peuvent entraîner des fluctuations importantes des prix du pétrole.
- ❖ **Facteurs géopolitiques** : Les tensions dans les régions productrices de pétrole, telles que le Moyen-Orient, peuvent avoir un impact significatif sur la volatilité des prix du pétrole et peuvent perturber l'approvisionnement en pétrole et entraîner des fluctuations rapides des prix.
- ❖ **Événements économiques mondiaux** : Les développements économiques mondiaux, tels que les crises financières, les changements dans les politiques monétaires des grandes économies ou les indicateurs économiques clés comme la croissance du PIB, peuvent également influencer la volatilité des prix du pétrole.
- ❖ **Spéculation financière** : Les investisseurs financiers peuvent spéculer sur les mouvements futurs des prix du pétrole, ce qui peut amplifier la volatilité des prix du pétrole. Les facteurs tels que les annonces de politique monétaire, les rapports sur les stocks de pétrole ou les prévisions météorologiques peuvent inciter les spéculateurs à acheter ou à vendre des contrats pétroliers, contribuant ainsi à la volatilité.

En raison de ces facteurs et d'autres encore, les prix du pétrole peuvent varier considérablement sur de courtes périodes, ce qui peut présenter des défis pour les acteurs du marché qui cherchent à gérer les risques liés à la fluctuation des prix.

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

3 La prévision des prix du pétrole

La prévision des prix du pétrole est un défi complexe en raison de la multitude de facteurs qui influent sur ces prix, ainsi que de leur nature volatile. Voici quelques approches couramment utilisées pour tenter de prévoir les prix du pétrole :

- ❖ **Analyse fondamentale** : Cette approche consiste à examiner les facteurs fondamentaux qui influencent l'offre et la demande de pétrole, tels que les données sur la production, les inventaires, la consommation mondiale, les développements géopolitiques, les politiques des principaux producteurs de pétrole, etc. Les analystes utilisent ces données pour élaborer des modèles de prévision des prix du pétrole.
- ❖ **Analyse technique** : Les analystes techniques examinent les graphiques des prix historiques du pétrole pour identifier les tendances, les modèles et les indicateurs techniques qui pourraient indiquer des mouvements futurs des prix du pétrole. Les méthodes courantes comprennent l'analyse des moyennes mobiles, des niveaux de support et de résistance, etc.
- ❖ **Modélisation économétrique** : Cette approche consiste à utiliser des modèles économétriques pour estimer les relations entre les variables économiques et les prix du pétrole. Ces modèles peuvent inclure des variables telles que le PIB, les taux de change, les prix des autres matières premières, etc., pour prédire les prix futurs du pétrole.
- ❖ **Modélisation des séries temporelles** : Les analystes utilisent des techniques de modélisation des séries temporelles pour analyser les données historiques des prix du pétrole et identifier les tendances, les saisons et les cycles saisonniers qui pourraient aider à prédire les prix futurs du pétrole.
- ❖ **Modélisation des réseaux neuronaux artificiels (RNA)** : Les RNA peuvent être formés sur des ensembles de données historiques de prix du pétrole pour apprendre des motifs complexes et faire des prédictions sur les prix futurs du pétrole.

Il est important de noter que malgré les progrès dans ces techniques, la prévision des prix du pétrole reste difficile et sujette à l'incertitude. Les prévisions peuvent fournir des indications utiles, mais elles ne sont jamais garanties.

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

4 L'utilisation de l'intelligence artificielle pour prédire la volatilité des prix du pétrole

L'utilisation de l'intelligence artificielle (IA) pour prédire la volatilité des prix du pétrole est devenue de plus en plus courante ces dernières années. Les techniques d'IA, telles que l'apprentissage automatique (Machine Learning ML) et le traitement du langage naturel, peuvent être appliquées à de grandes quantités de données historiques et en temps réel pour identifier des modèles complexes et des relations non linéaires qui pourraient ne pas être détectés par les méthodes traditionnelles. Voici quelques façons dont l'IA est utilisée pour prédire la volatilité des prix du pétrole :

- **Modèles prédictifs** : Les algorithmes d'apprentissage automatique ou le Machine Learning (ML) peuvent être formés sur des ensembles de données historiques de prix du pétrole ainsi que sur d'autres variables pertinentes, telles que les données économiques, géopolitiques et météorologiques, pour prédire la volatilité future des prix du pétrole. Ces modèles peuvent être basés sur une variété d'approches, y compris les réseaux de neurones, les arbres de décision, les machines à vecteurs de support (SVM), etc.
- **Analyse des sentiments** : L'analyse du sentiment à partir de sources telles que les actualités, les médias sociaux et les rapports financiers peut fournir des indications sur les perceptions du marché et les attentes des investisseurs, ce qui peut influencer la volatilité des prix du pétrole. Les techniques de traitement du langage naturel peuvent être utilisées pour extraire et analyser ces informations afin d'anticiper les mouvements des prix.
- **Modèles hybrides** : Certains modèles combinent des techniques d'IA avec des méthodes traditionnelles telles que l'analyse fondamentale et technique pour améliorer la précision des prévisions de la volatilité.
- **Surveillance en temps réel** : Les systèmes d'IA peuvent être utilisés pour surveiller en temps réel les événements économiques, politiques et géopolitiques ainsi que les actualités mondiales susceptibles d'avoir un impact sur la volatilité des prix du pétrole. Cela permet aux investisseurs et aux gestionnaires de risques de réagir rapidement aux changements du marché.
- **Adaptabilité** : Les modèles d'IA peuvent être adaptés et mis à jour régulièrement à mesure que de nouvelles données deviennent disponibles et que les conditions du

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

marché évoluent, ce qui permet d'améliorer continuellement leur précision et leur capacité à prédire la volatilité des prix du pétrole.

Dans ce travail on s'intéresse aux modèles hybrides pour l'étude et la prévision des prix du pétrole. Les modèles hybrides combinent généralement plusieurs approches pour prédire les prix du pétrole de manière plus précise. Voici quelques exemples de modèles hybrides récemment utilisés dans l'étude et la prévision des prix du pétrole :

1- Modèles ARIMA-GARCH : Cette combinaison associe deux types de modèles traditionnelles ARIMA (AutoRegressive Integrated Moving Average) qui est utilisé pour modéliser les tendances et les composantes saisonnières des données historiques des prix du pétrole, et le modèle GARCH (Generalized Autoregressive Conditional Heteroskedasticity) qui est utilisé pour modéliser la volatilité conditionnelle. En combinant ces deux modèles, on peut obtenir des prévisions plus précises qui capturent à la fois les tendances et la volatilité des prix du pétrole.

Il existe plusieurs études qui ont utilisé le modèle hybride ARIMA-GARCH pour prédire les prix du pétrole. Voici quelques exemples d'études pertinentes :

- "Forecasting Crude Oil Price Volatility" (2008) par L. Jammazi et A. Aloui : Cette étude examine l'utilisation du modèle hybride ARIMA-GARCH pour prédire la volatilité des prix du pétrole. Les auteurs constatent que ce modèle peut capturer avec précision les fluctuations de la volatilité et fournir des prévisions utiles pour les gestionnaires de risques et les investisseurs.
- "Forecasting the volatility of crude oil futures using intraday data" (2014) par R. Baumeister et G. Kilian : Cette étude utilise le modèle hybride ARIMA-GARCH pour prédire la volatilité des contrats à terme sur le pétrole brut en utilisant des données intra journalières. Les résultats montrent que le modèle hybride fournit des prévisions précises de la volatilité à court terme.
- "Forecasting oil price volatility" (2015) par R. Kocaarslan et al. : Cette étude compare les performances de différents modèles de prévision de la volatilité des prix du pétrole, y compris le modèle ARIMA-GARCH. Les résultats indiquent que le modèle hybride ARIMA-GARCH surpasse d'autres modèles en termes de précision des prévisions de volatilité.

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

- "Forecasting crude oil market volatility: Further evidence using GARCH-class models" (2018) par Y. Lin et S. Lu : Cette étude évalue l'efficacité des modèles GARCH pour prédire la volatilité des prix du pétrole et propose des améliorations pour mieux modéliser les variations de la volatilité.

Ces études démontrent l'efficacité du modèle hybride ARIMA-GARCH dans la prévision de la volatilité des prix du pétrole, ce qui peut être précieux pour les acteurs du marché qui cherchent à gérer les risques et à prendre des décisions éclairées dans le secteur énergétique.

2- Modèles SVM-ANN : Cette combinaison associe les machines à vecteurs de support (SVM) et les réseaux de neurones artificiels (ANN). Les SVM sont utilisées pour modéliser les relations linéaires et non linéaires entre les prix du pétrole et d'autres variables explicatives, tandis que les ANN sont utilisées pour capturer les motifs complexes et non linéaires des données. En combinant ces deux techniques, on peut obtenir une meilleure capacité de prédiction. Voici quelques études qui ont exploré l'utilisation des modèles SVM-ANN (Support Vector Machine - Artificial Neural Network) pour prédire les prix du pétrole :

- "Support vector machines, neural networks, and forecasting of the weekly prices of crude oil and petroleum products" (2007) par A. Reboredo et J. Rivera-Castro : Cette étude examine l'utilisation des SVM et des réseaux neuronaux pour prédire les prix hebdomadaires du pétrole brut et des produits pétroliers. Les résultats montrent que les modèles SVM-ANN peuvent fournir des prévisions précises, en particulier pour les horizons de prévision à court terme.
- "Prediction of crude oil price using artificial neural networks and commodity futures prices" (2014) par A. T. Atsalakis et A. Valavanis : Cette étude propose une approche hybride combinant les réseaux de neurones artificiels (ANN) et les prix des contrats à terme sur les matières premières pour prédire les prix du pétrole brut. Les résultats suggèrent que cette approche peut améliorer la précision des prévisions à court terme.
- "Forecasting crude oil price using artificial neural networks : A review of literature and comparison of algorithms" (2017) par T. Atsalakis et A. Valavanis : Cette revue de littérature analyse diverses études utilisant des réseaux de neurones artificiels pour prédire les prix du pétrole. Les auteurs comparent également différentes architectures de

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

réseaux neuronaux et techniques d'apprentissage pour améliorer les performances des prévisions.

- "Forecasting crude oil price volatility: A comparison of GARCH, SVR, and the hybrid models" (2019) par M. H. Chan et al. : Cette étude compare les performances des modèles GARCH, SVR (Support Vector Regression) et des modèles hybrides GARCH-SVR pour prévoir la volatilité des prix du pétrole brut. Les résultats montrent que les modèles hybrides peuvent offrir des prévisions plus précises que les modèles individuels.

Ces études démontrent l'efficacité des modèles SVM-ANN dans la prévision des prix du pétrole, en particulier lorsqu'ils sont combinés avec d'autres techniques ou données pertinentes telles que les prix des contrats à terme ou d'autres indicateurs économiques.

3- Modèles VARX-MLP : Cette combinaison associe les modèles vectoriels autorégressifs avec des exogènes (VARX) et les perceptrons multicouches (MLP). Les modèles VARX sont utilisés pour modéliser les relations dynamiques entre les prix du pétrole et d'autres variables explicatives, tandis que les MLP sont utilisés pour capturer les motifs complexes et non linéaires des données. En combinant ces deux approches, on peut obtenir des prévisions plus précises qui prennent en compte à la fois les relations dynamiques et les motifs non linéaires.

4- Modèles LSTM-CNN : Cette combinaison associe les réseaux de mémoire à court et long terme (LSTM) et les réseaux de neurones convolutionnels (CNN). Les LSTM sont utilisés pour capturer les dépendances à long terme dans les séries temporelles des prix du pétrole, tandis que les CNN sont utilisés pour extraire les caractéristiques spatiales des données. En combinant ces deux architectures, on peut obtenir des modèles plus robustes pour prédire les prix du pétrole. Voici quelques études qui ont exploré l'utilisation de ces modèles pour prédire les prix du pétrole :

- "Short-Term Crude Oil Price Forecasting with LSTM Neural Networks" (2017) par G. A. Diniz et al. : Cette étude examine l'utilisation de modèles LSTM pour prédire les prix du pétrole brut à court terme. Les auteurs comparent les performances des LSTM avec d'autres méthodes de prévision et constatent que les LSTM peuvent fournir des prévisions précises.
- "Deep Learning for Crude Oil Price Forecasting: A Comparative Study" (2018) par Z. Zhou et al. : Cette étude compare les performances de différents modèles de deep

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

learning, y compris les LSTM-CNN, pour prédire les prix du pétrole brut. Les résultats indiquent que les modèles LSTM-CNN peuvent offrir des prévisions compétitives par rapport à d'autres méthodes.

- "Crude oil price forecasting using hybrid CNN-LSTM neural network architectures" (2018) par S. Subudhi et al. : Cette étude propose une architecture hybride combinant des couches CNN et LSTM pour prédire les prix du pétrole brut. Les résultats montrent que cette approche peut capturer efficacement les motifs temporels et spatiaux des séries temporelles des prix du pétrole.
- "Forecasting the daily oil price using hybrid CNN–LSTM models with feature fusion technique" (2020) par X. Hu et al. : Cette étude propose une approche hybride qui combine des modèles CNN et LSTM avec une technique de fusion de caractéristiques pour prédire les prix du pétrole brut quotidiens. Les résultats montrent que cette méthode peut améliorer la précision des prévisions par rapport aux modèles individuels.

Ces études démontrent l'efficacité des modèles LSTM-CNN dans la prévision des prix du pétrole brut, en exploitant à la fois les relations séquentielles à long terme et les motifs spatiaux dans les données.

5- Les modèles GARCH-NN : Le modèle hybride GARCH-NN combine deux techniques : GARCH (Generalized Autoregressive Conditional Heteroskedasticity) pour modéliser la volatilité conditionnelle et les réseaux de neurones (NN) pour capturer les schémas non linéaires et complexes des données. Voici comment ce modèle fonctionne généralement :

- **GARCH :** La première partie du modèle est basée sur le modèle GARCH, qui est une méthode largement utilisée pour modéliser la volatilité des prix financiers. GARCH permet de capturer les variations de la volatilité dans le temps, ce qui est important dans la prédiction des prix du pétrole. En prenant compte la volatilité passée pour estimer la volatilité future.
- **Réseaux de neurones (NN) :** La deuxième partie du modèle utilise des réseaux de neurones pour capturer les schémas non linéaires et complexes des données historiques des prix du pétrole qui pourraient être difficiles à modéliser avec des approches traditionnelles.
- **Combinaison :** Dans ce modèle hybride, les sorties du modèle GARCH, qui représentent les estimations de la volatilité conditionnelle, sont utilisées comme

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

variables explicatives dans le réseau de neurones. Le réseau de neurones est ensuite formé sur les données historiques des prix du pétrole ainsi que sur les estimations de la volatilité conditionnelle fournies par GARCH. Cela permet au modèle de prendre en compte à la fois la structure temporelle des prix du pétrole et les informations sur la volatilité pour faire des prédictions plus précises.

Il y a moins d'études spécifiques sur les modèles GARCH-NN (Generalized Autoregressive Conditional Heteroskedasticity - Neural Network) pour prédire les prix du pétrole par rapport à d'autres combinaisons de modèles, mais voici quelques recherches pertinentes :

- "Oil Price Forecasting Using Hybrid Wavelet-GARCH Neural Networks Modeling" (2012) par A. Bouri et al. : Cette étude explore une approche hybride qui combine l'analyse en ondelettes, les modèles GARCH et les réseaux de neurones pour prédire les prix du pétrole brut. Les auteurs constatent que cette approche améliore la précision des prévisions par rapport aux modèles individuels.
- "Forecasting the volatility of crude oil futures using neural networks" (2018) par Y. Cui et al. : Cette étude utilise des réseaux de neurones pour prédire la volatilité des contrats à terme sur le pétrole brut. Bien que cette étude ne combine pas explicitement les modèles GARCH avec les réseaux de neurones, elle illustre l'utilisation des méthodes de prévision basées sur les réseaux de neurones pour les prix du pétrole.
- "Modelling the Volatility of Crude Oil Spot and Futures Prices Using Artificial Intelligence Approaches: A Comparative Analysis" (2020) par S. M. Hosseini-Motlagh et al. : Cette étude compare les performances des modèles GARCH, ARIMA et des réseaux de neurones pour prédire la volatilité des prix du pétrole. Bien que cette recherche ne combine pas explicitement les modèles GARCH avec les réseaux de neurones, elle offre des informations sur les performances des approches basées sur l'intelligence artificielle pour modéliser la volatilité du pétrole.

Ces études montrent que les approches hybrides qui combinent les modèles GARCH avec les réseaux de neurones ou d'autres techniques du Machine Learning peuvent offrir des prévisions plus précises des prix du pétrole ou de leur volatilité.

6- Le modèle NN-GARCH : Les NN-GARCH, en revanche, intègrent des réseaux de neurones pour modéliser la dynamique de la volatilité conditionnelle de manière plus

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

flexible et non linéaire. Dans un modèle NN-GARCH les entrées du réseau de neurones peuvent inclure des retards de la série temporelle ainsi que des indicateurs de volatilité passée. Le réseau de neurones peut être conçu pour capturer des modèles complexes et des dépendances non linéaires entre les variables. En ajustant les poids du réseau de neurones lors de l'estimation du modèle, les NN-GARCH peuvent s'adapter de manière plus précise aux données que les modèles GARCH traditionnels.

Les NN-GARCH ont été largement utilisés dans la modélisation financière pour la prévision de la volatilité des prix des actifs, la gestion des risques et d'autres applications où la modélisation précise de la volatilité conditionnelle est cruciale. Cependant, leur mise en œuvre peut nécessiter des ressources computationnelles plus importantes que les modèles GARCH traditionnels en raison de la complexité des réseaux de neurones.

Voici quelques exemples d'études qui ont utilisé les NN-GARCH pour la prévision des prix du pétrole :

- "Forecasting Crude Oil Prices Using Artificial Neural Networks : A Comparative Study of Linear and Non-linear Models" Cette étude compare les performances des modèles NN-GARCH avec d'autres modèles de prévision des prix du pétrole, tels que les modèles GARCH et ARIMA. Elle explore également l'efficacité des approches linéaires par rapport aux approches non linéaires dans la modélisation des séries temporelles financières.
- "Short-term crude oil price forecasting with machine learning techniques" : Cette étude utilise les NN-GARCH ainsi que d'autres techniques du Machine Learning pour prédire les prix du pétrole à court terme. Elle examine comment l'incorporation de données supplémentaires, telles que les données macroéconomiques et les nouvelles financières, peut améliorer la précision des prévisions.
- "Forecasting Crude Oil Price Using Artificial Neural Networks : A Time-Varying Approach" : Cette étude propose une méthode pour modéliser la volatilité conditionnelle du prix du pétrole en utilisant des NN-GARCH avec des paramètres variant dans le temps. Elle démontre comment cette approche peut capturer les variations de la volatilité à différentes périodes, ce qui peut améliorer la précision des prévisions.
- "Forecasting Oil Price Movements with Crack Spread Options Implied Volatility » : Cette étude explore l'utilisation de la volatilité implicite des écarts de prix (crack spreads) pour améliorer les prévisions des prix du pétrole en utilisant des modèles NN-GARCH. Elle

Section 1 : Les Techniques de l'Intelligence Artificielle utilisées dans la prévision des prix du pétrole

montre comment l'incorporation de cette information peut aider à mieux capturer les tendances et la volatilité des prix du pétrole.

Ces exemples illustrent différentes approches et applications des NN-GARCH dans la prévision des prix du pétrole, démontrant l'importance croissante de ces modèles dans la recherche en économie énergétique et financière.

Ces exemples ne représentent qu'une petite fraction des modèles hybrides utilisés dans l'étude et la prévision des prix du pétrole. En pratique, les chercheurs et les analystes peuvent expérimenter différentes combinaisons de techniques pour trouver celle qui offre les meilleures performances prédictives pour un ensemble de données.

Conclusion

En résumé, l'étude et la prévision de la volatilité des prix du pétrole jouent un rôle essentiel dans la gestion des risques, la planification stratégique et les décisions d'investissement dans les marchés énergétiques mondiaux. Les avancées dans les méthodes analytiques et les technologies de modélisation permettent aux acteurs du marché de mieux comprendre et anticiper les fluctuations des prix du pétrole, contribuant ainsi à une prise de décision plus informée et à une gestion plus efficace des risques.

Bien que l'IA puisse fournir des outils puissants pour prédire la volatilité des prix du pétrole, il est important de noter que ces prédictions sont toujours soumises à l'incertitude inhérente aux marchés financiers et que les modèles ne sont pas garantis. Une approche prudente et une compréhension approfondie des marchés pétroliers restent essentielles pour interpréter et utiliser efficacement les prévisions générées par l'IA.

Section 2

La Modélisation de la volatilité des prix du pétrole par le modèle GARCH

Sommaire

SECTION 2 : LA MODELISATION DE LA VOLATILITE DES PRIX DU PETROLE PAR LE MODELE GARCH **ERROR! BOOKMARK NOT DEFINED.**

INTRODUCTION..... **ERROR! BOOKMARK NOT DEFINED.**

1 ANALYSE DESCRIPTIVES ET EXPLORATION DES DONNEES..... **ERROR! BOOKMARK NOT DEFINED.**

1.1 Outils Utilisés pour l'Implémentation et la Préparation des DonnéesError! Bookmark not defined.

1.1.1 Langage de programmation MATLAB R2024a :..... **Error! Bookmark not defined.**

1.1.2 Toolboxes Principales **Error! Bookmark not defined.**

1.1.2.1 Econometrics Toolbox : **Error! Bookmark not defined.**

1.1.2.2 Statistics and Machine Learning Toolbox:..... **Error! Bookmark not defined.**

1.1.2.3 Financial Toolbox : **Error! Bookmark not defined.**

1.1.2.4 System Identification Toolbox : **Error! Bookmark not defined.**

1.1.3 L'environnement de travail Econometric Modeler **Error! Bookmark not defined.**

1.1.3.1 Fonctionnalités de l'Econometric Modeler **Error! Bookmark not defined.**

1.2 Etude graphique Error! Bookmark not defined.

1.3 Etude de la stationnarité des prix du pétrole..... Error! Bookmark not defined.

1.3.1 Stationnarité **Error! Bookmark not defined.**

1.4 Décomposition de la série des prix du pétrole : Error! Bookmark not defined.

2 MODELISATION DE LA SERIE DES PRIX DU PETROLE**ERROR! BOOKMARK NOT DEFINED.**

2.1 La Modélisation avec un GARCH Error! Bookmark not defined.

2.1.1 Identification du modèle GARCH..... **Error! Bookmark not defined.**

2.1.2 La spécification d'un modèle GARCH..... **Error! Bookmark not defined.**

2.1.3 L'estimation des paramètres du modèle **Error! Bookmark not defined.**

2.1.4 Validation du modèle GARCH (1.1) **Error! Bookmark not defined.**

2.1.4.1 Vérification des Résidus Standardisés **Error! Bookmark not defined.**

2.1.4.2 Autocorrélation des Résidus Standardisés et le Carré des résidus standardisés..**Error! Bookmark not defined.**

2.1.4.3 Test ARCH sur les Résidus Standardisés **Error! Bookmark not defined.**

3 LA PREVISION DE LA VOLATILITE DES PRIX DU PETROLE PAR UN MODELE GARCH (1.1)..... **ERROR! BOOKMARK NOT DEFINED.**

CONCLUSION **ERROR! BOOKMARK NOT DEFINED.**

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Introduction

Le modèle **GARCH (Generalized Autoregressive Conditional Heteroskedasticity)** est une technique statistique couramment employé pour analyser et prévoir la volatilité observée dans la plupart des séries temporelles. Ce modèle capture les variations conditionnelles de la volatilité en prenant en compte l'hétéroscédasticité (la variabilité de la variance au fil du temps). Dans le contexte des prix du pétrole, le modèle GARCH est particulièrement pertinent, car il permet de modéliser les périodes de forte volatilité souvent observées en raison des chocs géopolitiques, des changements dans l'offre et la demande, ainsi que des événements économiques mondiaux. Par conséquent, le GARCH aide à fournir des prévisions plus précises.

Cette section est dédiée à modéliser les données des prix du pétrole par un modèle GARCH pour avoir des prévisions et des résultats qu'on puisse les comparer avec les autres modèles et pouvoir utiliser les sorties du modèle GARCH dans l'hybridation objet du travail.

La suite de cette section est dédiée pour expliquer toutes les phases de la modélisation de la volatilité. Aussi, On va présenter les données des prix du pétrole ainsi que leurs propriétés statistiques, avant de procéder à l'étape de la modélisation et de la prévision.

1 Analyse descriptives et exploration des données

On commencera par une analyse descriptive et exploratoire des données des prix du pétrole extrait du site officiel de l'OPEP pour identifier les caractéristiques fondamentales des données, telles que les tendances générales, les cycles saisonniers et les anomalies.

L'étude porte sur la modélisation d'une série temporelle réel, qui représente les prix du pétrole {source : le site officiel de l'OPEP des données des prix du panier de l'OPEP}.

Nous disposons d'une série quotidienne composée de 1 062 observations, exprimées en dollars américains par baril, tirées des bulletins statistiques de l'OPEP et couvrant la période indiquée 02/01/2020 au 29/01/2024. La notation choisie pour la série est **P** pour désigner le prix du pétrole, **t** indique le temps, on note **P_t** : le prix du pétrole en \$ / baril. Les données des prix du pétrole, les graphes et les calculs sont effectués par MATLAB.2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

1.1 Outils Utilisés pour l'Implémentation et la Préparation des Données

Pour bien mener cette étude, plusieurs outils ont été utilisés pour l'implémentation et la préparation des données :

1.1.1 Langage de programmation MATLAB R2024a :

MATLAB signifie **MATrix LABoratory**, c'est un langage de programmation développé par **MathWorks**, il est largement utilisé dans les domaines de l'ingénierie, des sciences, de la finance, et de la recherche académique. La version utilisée dans notre travail est la version R2024a est l'une des versions les plus récentes on l'avait obtenu sur le site officiel de **MathWorks**, comme version d'essai spéciale pour les étudiants et les chercheurs. La nouvelle version introduit plusieurs améliorations et nouvelles fonctionnalités conçues à faciliter le calcul matriciel, l'algèbre linéaire, les statistiques, et l'analyse de données, avec une syntaxe simple et intuitive. MATLAB R2024a dispose de nombreuses bibliothèques (appelées "**Toolboxes**") spécialisées pour différentes applications, comme le traitement du signal, l'apprentissage automatique, l'optimisation, le traitement d'images, etc.

Dans la version MATLAB R2024a, et plus spécifiquement dans l'environnement **Econometric Modeler**, plusieurs bibliothèques et Toolboxes sont utilisées pour effectuer des analyses économétriques et des séries temporelles. Voici les principales Toolboxes et fonctions de MATLAB R2024a utilisées pour ces analyses :

1.1.2 Toolboxes Principales

1.1.2.1 Econometrics Toolbox :

C'est la principale boîte à outils utilisée pour l'analyse économétrique dans MATLAB R2024a. Elle inclut des fonctions pour la modélisation de séries temporelles, les modèles de régression, les tests d'hypothèses, les modèles ARIMA, GARCH, VAR, et bien plus encore.

1.1.2.2 Statistics and Machine Learning Toolbox:

Cette Toolbox est utilisée pour des analyses statistiques générales et des algorithmes d'apprentissage automatique, qui peuvent également être appliqués à des données économiques et de séries temporelles.

1.1.2.3 Financial Toolbox :

Cette boîte est spécialisée pour les applications financières, cette Toolbox comprend des fonctions pour la gestion de portefeuille, les dérivés financiers, et la modélisation des risques.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

1.1.2.4 System Identification Toolbox :

Cette Toolbox est utilisée pour l'identification des systèmes et la modélisation des séries temporelles, notamment pour les modèles dynamiques et les prévisions.

1.1.3 L'environnement de travail Econometric Modeler

Econometric Modeler est un environnement de travail intégré dans MATLAB R2024a, spécialement conçu pour l'analyse et la modélisation économétrique des séries temporelles. Il fait partie des fonctionnalités offertes par la boîte à outils Econometrics Toolbox.

1.1.3.1 Fonctionnalités de l'Econometric Modeler

- a- Econometric Modeler fournit une interface graphique interactive qui permet de charger, visualiser, et analyser des séries temporelles sans écrire de code MATLAB.
- b- Prétraitement des Données est un outil permet de nettoyer les données, de manipuler les séries temporelles, de détecter et de traiter les valeurs manquantes, et de transformer les données pour les rendre stationnaires.
- c- Econometric Modeler offre des outils pour effectuer des analyses statistiques descriptives, générer des corrélogrammes, et explorer la saisonnalité, la tendance et les résidus des séries temporelles.
- d- Modélisation il prend en charge divers modèles économétriques, tels que les modèles ARIMA, GARCH, VAR, et les modèles de séries chronologiques multivariées. Les utilisateurs peuvent ajuster ces modèles aux données, estimer les paramètres, et comparer les performances des modèles.
- e- Une fois les modèles ajustés, Econometric Modeler permet d'effectuer des tests diagnostiques, comme le test de Breusch-Godfrey pour l'autocorrélation des résidus, ou le test d'hétéroscédasticité, pour s'assurer de la validité du modèle.
- f- Générer des prévisions à partir des modèles ajustés et visualiser ces prévisions dans l'interface graphique.

1.2 Etude graphique

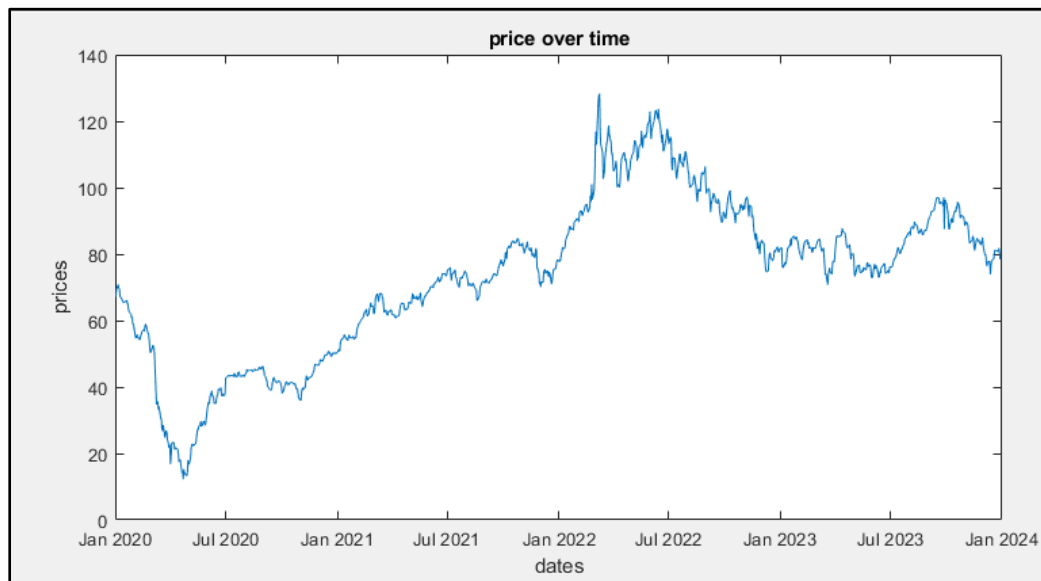
L'étude graphique de la série temporelle des prix du pétrole est une étape importante qui précède la modélisation et peut fournir des informations précieuses sur les tendances, les saisons et les cycles dans les données. Cette approche s'appuie sur une analyse visuelle du graphique représentant les données historiques de l'évolution des prix du pétrole, ainsi que sur l'examen

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

de la stationnarité, une étape cruciale dans la phase de modélisation de la série des prix du pétrole étudiée.

Dans ce travail on va diviser les données en deux séries une pour la modélisation et la prévision contient 1032 observations des prix du pétrole couvrant la période 02/01/2020 à 31/12/2023. La deuxième série est conservé à la vérification des résultats de prévision couvrant le période de 30 jour restant. La figure **04** suivante illustre l'évolution des prix du pétrole au cours de la période d'étude de 02/01/2020 à 29/01/2024.

FIGURE 4: Evolution des prix du pétrole entre 02/01/2020 et 29/01/2024



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

La lecture visuelle du graphe ci -dessus montre des fluctuations des prix du pétrole brut dans le temps. Et delà, on peut noter deux phases d'évolution ou bien de volatilité dans cette figure : Du 06 janvier 2020 au 22 Avril 2020 et de 20 décembre 2021 au 09 mars 2022.

La première phase (06 janvier 2020 au 22 Avril 2020) est marquée par une baisse brutale des prix du pétrole. Au cours de cette période, les prix du pétrole ont connu une période agitée en raison de différents facteurs, notamment la pandémie du COVID-19 et les tensions géopolitiques. En Janvier 2020 les prix du pétrole étaient relativement stables en début d'année, les prix se situant autour de 65-70 dollars le baril. Février 2020 les inquiétudes concernant la propagation du COVID-19 en dehors de la Chine ont commencé à peser sur les marchés mondiaux, y compris sur les prix du pétrole. Les prix ont légèrement baissé et en Mars 2020 la pandémie de COVID-19 s'est rapidement propagée à travers le monde, entraînant une forte baisse de la demande de pétrole alors que de nombreux pays imposaient des mesures de

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

confinement. L'Arabie saoudite et la Russie ont également entamé une guerre des prix en augmentant leur production, ce qui a fait chuter les prix du pétrole à des niveaux historiquement bas. Le Brent Crude est tombé en dessous de 30 dollars le baril. Les prix du pétrole ont continué de fluctuer de manière volatile en avril 2020. Les inquiétudes persistantes concernant l'excès d'offre et la faible demande, combinées à des problèmes logistiques liés au stockage, ont maintenu les prix à des niveaux bas. Cependant, les efforts concertés de l'OPEP pour réduire la production et stabiliser les marchés ont contribué à une légère remontée des prix vers la fin du mois.

Entre le 20 décembre 2021 et le 9 mars 2022, qui est la deuxième phase remarquable dans la figure, les prix du pétrole ont généralement affiché une tendance à la hausse, soutenue par une reprise économique mondiale et des tensions géopolitiques, bien que la volatilité ait persisté en raison de divers facteurs influençant l'offre et la demande. Voici un aperçu de cette période :

Début 2022 : Les prix du pétrole ont généralement augmenté au début de l'année 2022. Cette tendance a été soutenue par une reprise économique mondiale après les perturbations causées par la pandémie de COVID-19, ainsi que par des signes de reprise de la demande de pétrole dans de nombreux pays.

Tensions géopolitiques : Les tensions géopolitiques dans certaines régions productrices de pétrole, telles que le Moyen-Orient, ont également contribué à soutenir les prix du pétrole. Les préoccupations concernant les risques de perturbations de l'approvisionnement en raison de conflits ou de tensions entre les pays producteurs ont maintenu une certaine pression à la hausse sur les prix.

Variabilité des prix : Malgré cette tendance à la hausse globale, les prix du pétrole ont été sujets à une certaine volatilité pendant cette période en raison de facteurs tels que les rapports sur les stocks de pétrole brut, les données économiques publiées par les principaux pays consommateurs de pétrole, et les événements géopolitiques imprévus.

On peut observer que la série des prix du pétrole semble non stationnaire. Selon la figure N°04, Les données brutes de la série temporelle des prix du pétrole montrent des fluctuations significatives sont schéma périodique évident à première vue, indiquant une forte volatilité avec des périodes marquées de hausse et de baisse. En examinant la tendance, on observe qu'elle capture les changements à long terme On observe un léger mouvement de croissance ou de

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

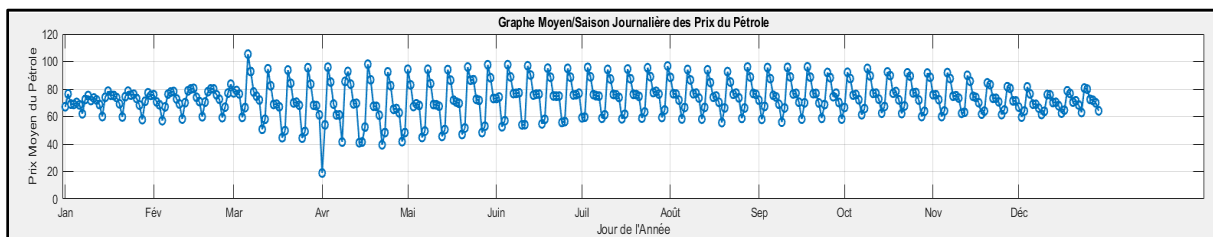
décroissance à moyen terme, indiquant la présence d'une composante déterministe dans la série, appelée "tendance" ou "trend", qui reflète l'évolution générale de la série.

On observe également des fluctuations imprévisibles, appelées variations irrégulières ou mouvements résiduels « stochastiques ». Ces variations sont souvent causées par des événements exceptionnels, tels que la pandémie de COVID-19.

Nous remarquons aussi un phénomène accidentel, il s'agit du pic anormalement élevé au mois Mars pendant l'année 2022. Une annonce d'une réduction de l'offre de pétrole de la part de l'OPEP ou de certains grands producteurs ait contribué à la hausse des prix.

Le graphique de la série montre également l'absence de périodicité dans les valeurs observées, ce qui indique l'absence de tout phénomène dit « saisonnier ». Cette observation peut être confirmée à l'aide du graphique de saisonnalité Figure 05 présenté ci-dessous.

FIGURE 5:Le graphe saisonnier des prix du pétrole entre 03/01/2020 et 29/12/2023.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Après avoir étudié et examiner le graphe de la série des prix du pétrole on a déterminé deux différentes composantes principales : T_t une tendance et fluctuations irrégulières ε_t ; le phénomène accidentel étant intégré au terme de fluctuations irrégulières. Il existe deux représentations possibles du processus permettent d'analyser la structure sous-jacente des données et d'identifier les effets de chaque composante :

1- Représentation Additive : Cette approche est appropriée lorsque l'amplitude des variations saisonnières reste constante sur toute la série. Dans une décomposition additive, La variable prix du pétrole s'écrit comme la somme des deux

composantes : $Y_t = T_t + R_t$.

- T_t : Tendence (Trend) à long terme.
- R_t : Résidu (Residual) ou bruit, qui capte les variations aléatoires.

2- Représentation Multiplicative : La décomposition multiplicative est utilisée lorsque l'amplitude des variations saisonnières change proportionnellement au

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

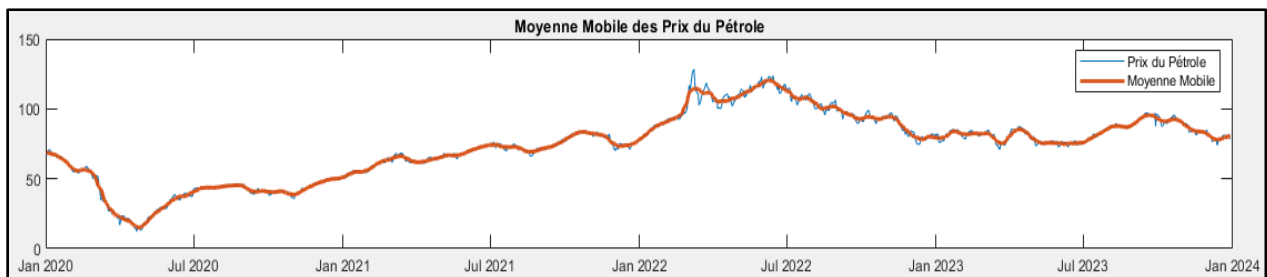
niveau de la série. Dans une décomposition multiplicative, la variable prix du pétrole s'écrit comme le produit de la tendance et de la composante aléatoire :

$$Y_t = T_t * R_t . \text{ L'ajustement multiplicatif intervient dans les modèles (G)ARCH.}$$

Pour déterminer la représentation la plus adaptée à la série des prix du pétrole, nous appliquerons le test paramétrique de Bays-Ballot, en procédant selon les étapes suivantes :

- Calculer les Moyennes μ_t et les variances σ_t pour chaque mois de la série. Les moyennes et les variances calculés peuvent être comparés pour évaluer la régularité des fluctuations saisonnières. La représentation graphique de la moyenne et la variance mensuelles calculé montre une instabilité remarquable au cours du temps. Les graphes de la moyenne et la variance mensuelles des prix du pétrole montrent une moyenne mobile en hausse avec une variance mobile plus au moins élevée cela indique une tendance haussière instable et durable mais avec une volatilité élevée, ce qui suggère des incertitudes ou des fluctuations importantes sur le marché.

FIGURE 6:Le graphe Moyenne mensuelles des prix du pétrole entre 03/01/2020 et 29/01/2024.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

FIGURE 7:Le graphe de la variance mensuelles des prix du pétrole entre 03/01/2020 et 29/01/2024.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

- Pour ces paires d'observations μ_t et σ_t on estimera les paramètres α et β de l'équation $\sigma_t = \alpha + \beta\mu_t + \varepsilon_t$ en utilisant la méthode des moindres carrés ordinaire.

Tableau 1: Régression Linéaire de la variance mensuelles sur les moyennes mensuelles des prix du pétrole.

Linear regression model:				
$y \sim 1 + x1$				
Estimated Coefficients:				
	Estimate	SE	tStat	pValue
(Intercept)	2.0358	0.79019	2.5764	0.013259
x1	0.009271	0.010244	0.90504	0.37016
Number of observations: 48, Error degrees of freedom: 46				
Root Mean Squared Error: 1.67				
R-squared: 0.0175, Adjusted R-Squared: -0.00386				
F-statistic vs. constant model: 0.819, p-value = 0.37				

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

L'estimation du modèle est donnée par l'équation suivante :

$$\sigma_t = 2.0358 + 0.009271\mu_t$$

La statistique t de student pour le paramètre β est égale à 0.90 avec un P-value de 0.370 qui est supérieur à 0.05, donc on accepte l'hypothèse nulle (disant que β est égal à 0, le coefficient est n'est pas significativement différent de 0 ainsi, la représentation du modèle adopté est de type additif.

Comme on a noté précédemment, pour une bonne estimation et vérification des méthodes de prédiction, les données sont divisées en deux classes. La première classe, regroupant les données des prix du pétrole sur une période de 04 années est réservée pour l'estimation du modèle et l'autre classe est utilisée pour vérifier le modèle estimé. Cette classe regroupe les données qui restent des données des prix du pétrole d'une période 30 jours le tableau N° 02 illustre la distribution des données.

Tableau 2: Division des données Utilisées

Données	Periode
Données d'estimation	02 janvier 2020 – 31 décembre 2023
Données de vérification	31 décembre 2023 – 29 janvier 2024

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

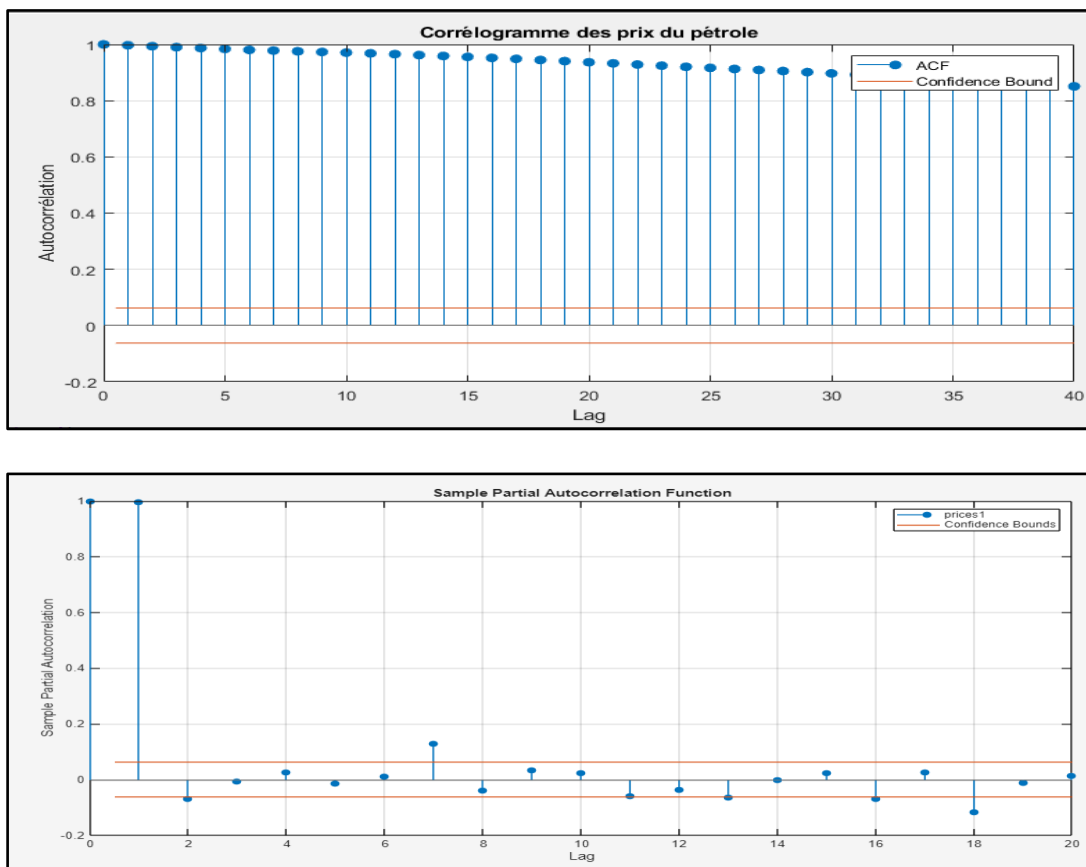
Pour cette étude, la méthode de Box-Jenkins est utilisée en premier lieu suivi des Réseaux de Neurones Artificiels.

1.3 Etude de la stationnarité des prix du pétrole

1.3.1 Stationnarité

L'analyse de la stationnarité d'une série temporelle, comme celle des prix du pétrole, est une étape essentielle dans les analyses économiques et financières. La stationnarité d'une série signifie que ses caractéristiques statistiques, telles que la moyenne, la variance et l'autocorrélation, restent constantes au fil du temps. Pour vérifier cela, nous allons réaliser un test de racine unitaire sur la série des prix du pétrole, en utilisant le test de Dickey-Fuller Augmenté. Ce test permet non seulement de détecter une éventuelle tendance (racine unitaire), mais aussi de déterminer la méthode la plus appropriée pour rendre la série stationnaire. Pour étudier la stationnarité des prix du pétrole, nous examinerons d'abord le corrélogramme (1), puis confirmerons les résultats à l'aide du test de Dickey-Fuller Augmenté.

Corrélogramme 1: Les autocorrélation simple et partielle de la série des prix du pétrole.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

la fonction d'autocorrélation (ACF) illustre la corrélation entre les prix du pétrole actuelles et leurs valeurs passées. Les barres qui dépassent la ligne Rouge, représentant l'intervalle de confiance, indiquent des autocorrélations significatives, suggérant que les prix du pétrole récentes sont fortement influencées par les prix des périodes précédentes.

la fonction d'autocorrélation partielle (PACF) montre une forte autocorrélation partielle au premier lag, indiquant une dépendance entre les valeurs du prix du pétrole successives, on remarque également une autocorrélation partielle significative, qui indique à une influence persistante importante pour la modélisation des données des prix du pétrole. Après le premier lag, la PACF diminue, montrant que l'autocorrélation à des lags plus éloignés est principalement due aux effets des lags plus proches.

On observe également que les valeurs de la fonction d'autocorrélation (ACF) sont proches les unes des autres et décroissent lentement vers zéro, ce qui indique une non-stationnarité due à une tendance. Afin de confirmer la non-stationnarité suggérée par le corrélogramme N°01 précédent, nous procéderons à l'exécution du test de racine unitaire de Dickey-Fuller Augmenté.

Tableau 3: Les test de Dickey-Fuller Augmentés des prix du pétrole

Augmented Dickey-Fuller Test(prices1)									
Null Hypothesis: prices1 contains a unit root									
Results									
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Lags	Model	Test Statistic	Significance Level
1	<input type="checkbox"/>	false	0.5820	-0.1910	-1.9416	0	AR	t1	0.0500

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Le tableau 3 présente les résultats du test de DFA sur la série des prix du pétrole effectué par Econometric Modeler Toolbox. Le test aboutit à l'acceptation de l'hypothèse nulle, indiquant la présence d'une racine unitaire. En effet, la p-value obtenue lors de l'application du test de Dickey-Fuller Augmenté (DFA) est supérieure à 0.05, ce qui signifie que la série des prix du pétrole n'est pas stationnaire et présente une racine unitaire. Par conséquent, il est nécessaire de décomposer la série des prix du pétrole afin de la rendre stationnaire.

1.4 Décomposition de la série des prix du pétrole :

La décomposition de la série temporelle est une méthode utilisée pour séparer une série en plusieurs composantes sous-jacentes, généralement la tendance (trend) qui représente la

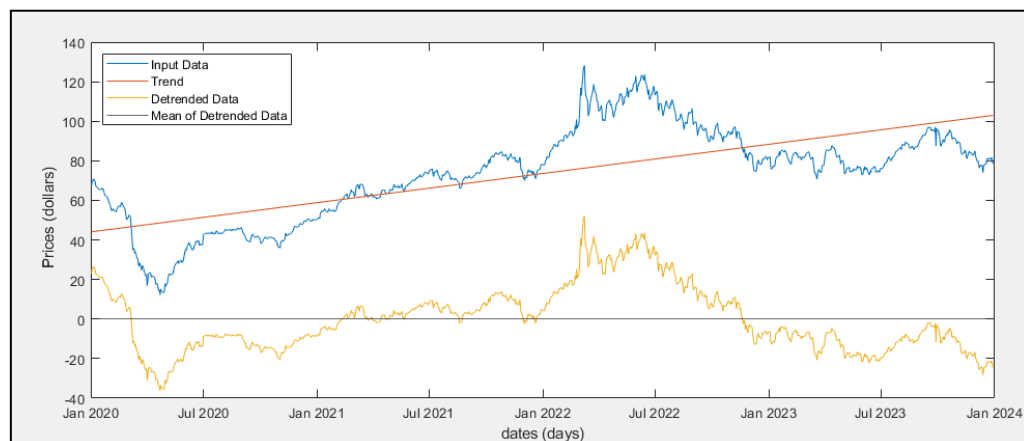
Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

direction générale des données à long terme, la saisonnalité (seasonality) qui capture les variations régulières qui se répètent à des intervalles régulières, et le résidu (residual ou noise) Les variations restantes après avoir retiré la tendance et la saisonnalité. Elles sont souvent considérées comme du bruit aléatoire. Cette technique est très utile pour analyser les différentes influences sur la série et pour préparer les données pour une modélisation plus avancée.

Dans le cas de notre travail, les données montre des phases distinctes de croissance et de décroissance, aussi les données montrent un taux de croissance qui ralentit dans ce cas une tendance linéaire peut capturer cette décélération et peut modéliser de manière simple ces phases, en capturant l'augmentation et la diminution.

L'estimation de la tendance est une méthode statistique utilisée pour modéliser une relation où le prix du pétrole change à un rythme variable au fil du temps.

FIGURE 8:L'évolution des prix du pétrole et la ligne trend / moyenne.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

La stationnarisation de la série des prix du pétrole se déroule en plusieurs étapes : tout d'abord, on élimine la tendance en estimant la fonction de tendance et en analysant les résidus obtenus, considérés comme une série stationnaire. Si la série des résidus n'est pas stationnaire, la deuxième étape consiste à appliquer un filtre de différence première, puis à modéliser la nouvelle série différenciée. Cette approche permet de traiter cette non-stationnarité en calculant les résidus par rapport à la tendance estimée à l'aide de la méthode des moindres carrés ordinaires. Les résultats obtenus sont présentés dans le tableau 4 suivant :

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Tableau 4: Régression linéaire de la tendance.

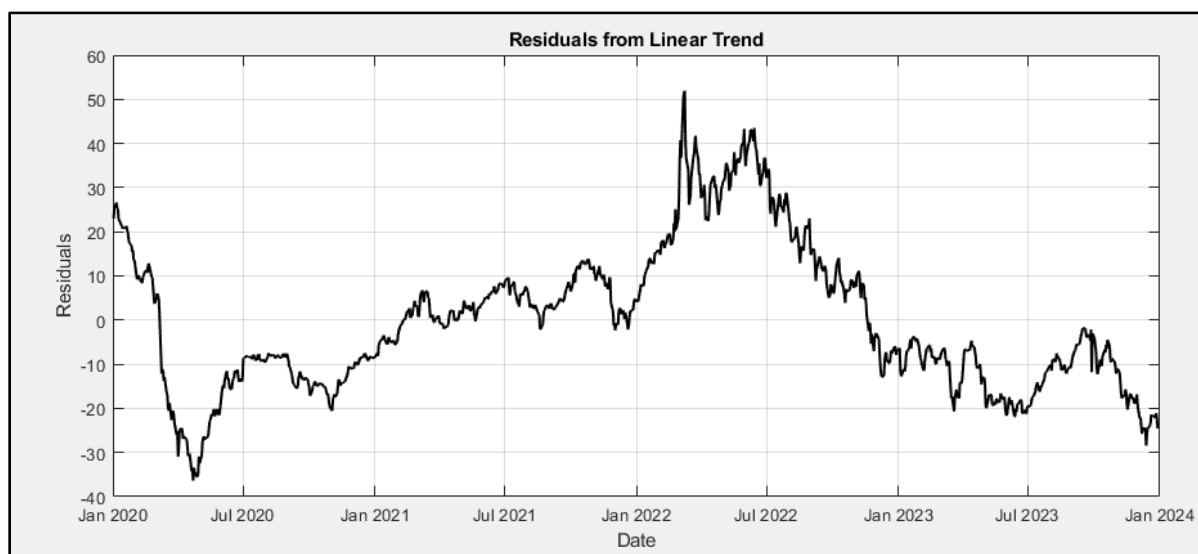
```
Linear regression model:  
y ~ 1 + x1  
  
Estimated Coefficients:  
  
              Estimate      SE      tStat      pValue  
-----  
(Intercept)  35.636      1.2035     29.61     6.6761e-140  
x1           0.51564     0.015572    33.114    2.5065e-164  
  
Number of observations: 1032, Error degrees of freedom: 1030  
Root Mean Squared Error: 11.8  
R-squared: 0.516, Adjusted R-Squared: 0.515  
F-statistic vs. constant model: 1.1e+03, p-value = 2.51e-164
```

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

L'estimation donne l'équation suivante : $P_t = 35.636 + 0.515@trend$

Les résidus obtenus de cette estimation sont considérés comme une série stationnaire, et nous procédons à leur analyse en suivant la méthodologie de Box-Jenkins. L'étude commence par une analyse graphique de la série des résidus. La figure 09 illustre l'évolution des résidus issus de l'estimation de la fonction de tendance de la série P_t .

FIGURE 9: Evolution des résidus obtenus de l'estimation de la fonction de la tendance de la série P_t .

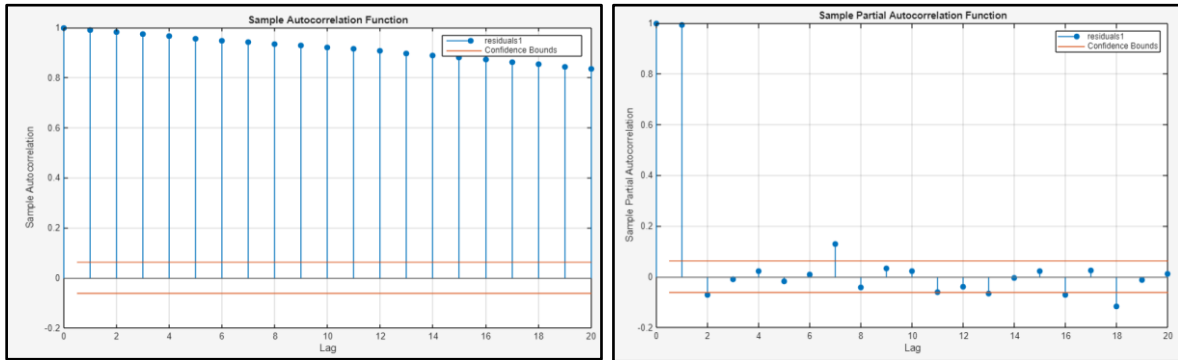


Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Le graphe montre des fluctuations importantes ce qui confirme l'instabilité de la série des prix du pétrole même après l'élimination de la tendance (trend), donc on conclut que la série est toujours non stationnaire. L'analyse visuelle du corrélogramme 2 des résidus peut confirmer cette non stationnarité.

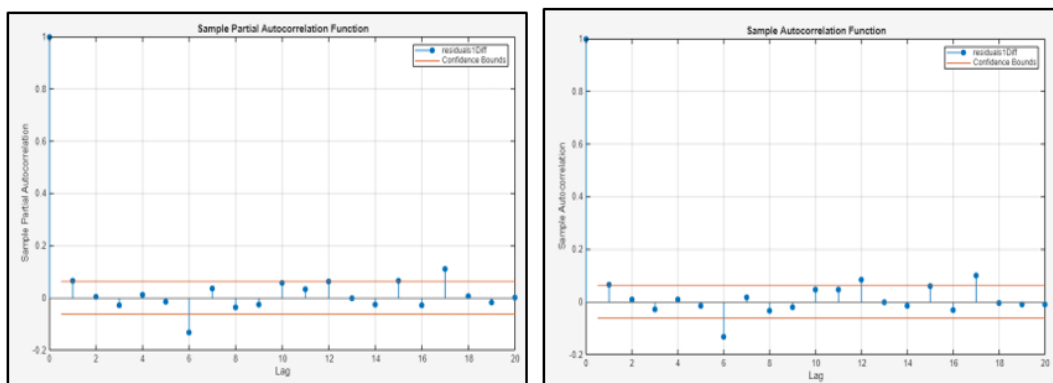
Corrélogramme 2: Les autocorrélations simple et partielles des résidus de la régression trend.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

la fonction d'autocorrélation (ACF) montre toujours une corrélation entre les prix du pétrole actuelles et leurs valeurs passées. On voit des barres qui dépassent l'intervalle de confiance, On remarque également que les valeurs des résidus sont proches les unes des autres et décroissent très lentement vers zéro, ce qui indique un signe de non-stationnarité. Dans ce cas, on passe au filtre de différence première, cette transformation est souvent utilisée pour éliminer la non stationnarité, il s'agit de créer une nouvelle série différenciée de la série des prix du pétrole ΔP_t et on refait l'analyse des fonctions de corrélation de la nouvelle série.

Corrélogramme 3: Les autocorrélations simples et partielles de la nouvelle série des prix du pétrole.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Il apparait clairement que la série différencié ne présente ni tendance ni dérive. Après la différenciation, la série devient stationnaire, ce qui est confirmé par l'application du test de racine unitaire DFA.

Tableau 5: Test de DFA de la série des prix différencier.

Augmented Dickey-Fuller Test(ResidualsDiff)									
Null Hypothesis: ResidualsDiff contains a unit root									
Results									
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Lags	Model	Test Statistic	Significance Level
1	<input type="checkbox"/>	true	1.0000e-03	-30.0339	-1.9416	0	AR	t1	0.0500
2	<input type="checkbox"/>	true	1.0000e-03	-30.1097	-2.8648	0	ARD	t1	0.0500
3	<input type="checkbox"/>	true	1.0000e-03	-30.0950	-3.4145	0	TS	t1	0.0500

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Dans les trois modèles (AR et ARD et TS) le test repose sur le rejet de l'hypothèse Null d'existence de racine unitaire. On remarque que les probabilités obtenues dans les trois modèles sont inférieures à 0,05, ce qui indique que la série des prix du pétrole est devenue stationnaire après la première différenciation. Ainsi, elle ne présente ni racine unitaire, ni tendance, ni dérive. Dans les étapes suivantes de l'analyse, l'étude se concentrera sur cette nouvelle série stationnaire des prix du pétrole ΔP_t .

2 Modélisation de la série des prix du pétrole

Dans notre étude on s'intéresse à la prévision de la volatilité plutôt que la tendance ou le cycle moyen de la série, un modèle GARCH est utilisé directement pour capturer la nature changeante de la variance. La série des prix est déjà stationnaire il peut être suffisant de modéliser directement la volatilité avec un modèle GARCH sans se soucier de la dynamique de la moyenne. Les modèles ARMA se concentre sur les relations linéaires et l'effet des valeurs passées sur les valeurs futures. Cependant, il suppose généralement que les résidus (les erreurs) sont homoscedastiques, c'est-à-dire qu'ils ont une variance constante ce qui n'est pas le cas de notre série qui présente une instabilité de la variance.

2.1 La Modélisation avec un GARCH

Trois étapes essentielles pour faire une modélisation GARCH :

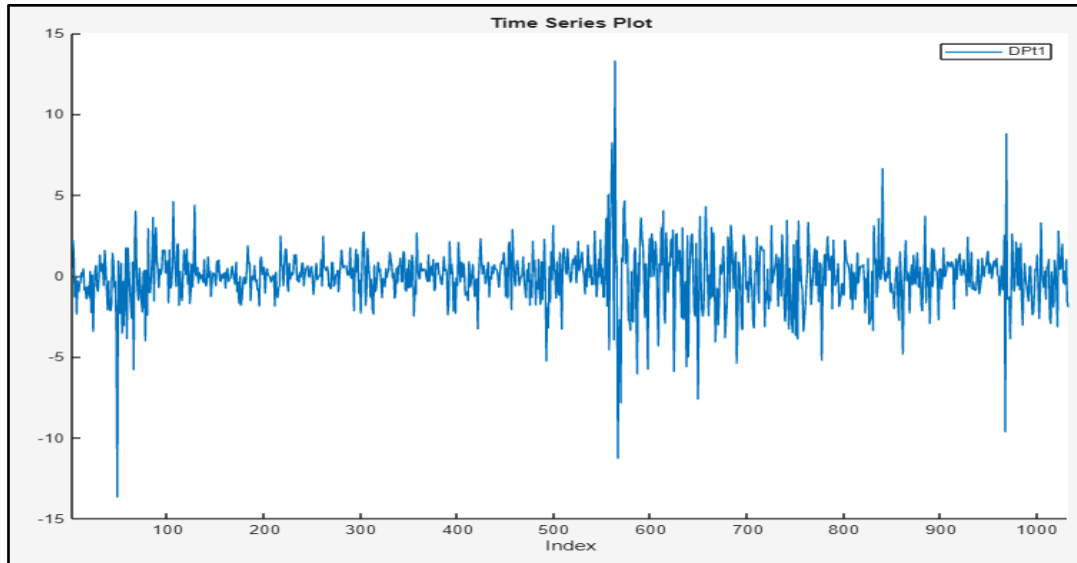
Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

2.1.1 Identification du modèle GARCH

L'identification est une étape très importante qu'elle se repose sur l'utilisation des graphiques de la série, des résidus et le test statistique Engle's ARCH pour vérifier la présence de volatilité conditionnelle.

- Visualisation de la série temporelle ΔP_t

FIGURE 10: La série des prix du pétrole différencier ΔP_t .



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

La série présente des périodes de grande volatilité suivie de période de faible volatilité ce qui suggère la présence d'hétéroscédasticité conditionnelle. Dans ce cas on examine la série des prix du pétrole si elle présente une variance hétéroscédastique en appliquant le test de Engle'ARCH. Le test d'Engle (ou test ARCH) est une méthode statistique spécifique pour détecter la présence d'hétéroscédasticité. Il vérifie si les résidus au carré de la série sont auto-corrélés.

Tableau 6: Test de Engle's ARCH.

Engle's ARCH Test for Heteroscedasticity(DPT1)						
Null Hypothesis: DPT1 exhibits no ARCH effects						
Results						
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Significance Level
1	<input checked="" type="checkbox"/>	true	1.6363e-05	18.5718	3.8415	0.0500

Hétéroscédasticité conditionnelle détectée (hypothèse nulle rejetée).
p-Value: 0.0010

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Le test donne un p-value inférieur à 0.05 ce qui signifie le rejet de l'hypothèse nulle et conclure que la série présente un effet d'hétéroscédasticité.

Après la confirmation que la variance des données des prix du pétrole dépend du temps, et la série ΔP_t présente un effet ARCH alors un modèle GARCH est bien adapté pour modéliser et prévoir cette série.

2.1.2 La spécification d'un modèle GARCH

L'étape de spécification est une étape cruciale qui consiste à déterminer la structure du modèle que on va utiliser pour analyser la volatilité conditionnelle des données des prix du pétrole. Cette spécification inclut le choix du nombre de termes autorégressifs (p) et de termes de moyenne mobile (q) dans le modèle. Voici les étapes théoriques pour spécifier un modèle GARCH.

On remarque le corrélogramme 3 on peut retenir les cinq modèles présenter dans le tableau 07 suivant :

Tableau 7:Modèle GARCH retenus.

Modèles	AIC	BIC
ARCH (1)	2.6489e+03	2.6588e+03
ARCH (2)	2.6186e+03	2.6334e+03
GARCH (1.1)	2.4569e+03	2.4717e+03
GARCH (2.1)	2.4589e+03	2.4787e+03
GARCH (2.2)	2.4530e+03	2.4777e+03

Source : Réalisée par l'auteur

Le tableau donne les résultats des deux critères de choix entre cinq modèles GARCH retenu par l'observation visuelle des corrélogramme de la série des prix du pétrole le choix du modèle est effectué par la sélection du modèle qui représente le minimum des deux critères donc selon les résultats le modèle qui représente le mieux notre série d'étude est le modèle GARCH (1.1).

Les modèles en rouge c'est des modelés qui donnent des coefficients significativement différents de zéro donc ces modèles ne peuvent pas représenter la série des prix du pétrole ou autrement dit ces modèles ne capturent pas correctement le comportement des données.

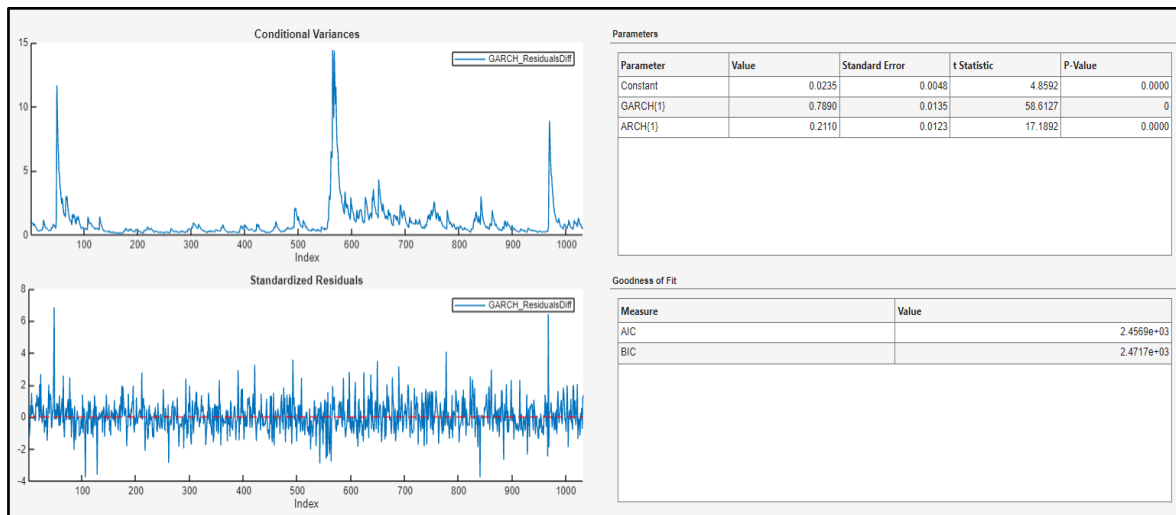
Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

2.1.3 L'estimation des paramètres du modèle

Le choix de la méthode d'estimation se diffère selon le type de processus étudié. L'étape de l'estimation des paramètres d'un modèle GARCH est cruciale pour s'assurer que le modèle représente correctement la dynamique de la volatilité dans une série temporelle. Parmi plusieurs méthodes d'estimation on trouve que là plus couramment utilisée pour des paramètres d'un modèle GARCH est l'estimation par Maximum de Vraisemblance (MLE - Maximum Likelihood Estimation).

Les résultats de l'estimation par MATLAB R2024a. Econometric Modeler Toolbox sont réunis dans le tableau 08 avec une représentation graphique de la variance conditionnelle et les résidus standardiser :

Tableau 8: Estimation du modèle GARCH (1.1).



Source : Réalisée par l'auteur dans le programme Matlab2024a.

Le modèle estimé est donc le suivant : $h_t = 0.0235 + 0.2110\varepsilon_{t-1}^2 + 0.7890h_{t-1}$

2.1.4 Validation du modèle GARCH (1.1)

Après l'estimation des paramètres du modèle retenu on passe à la validation en analysant et vérifiant les résidus obtenus de cette estimation et cela signifie qu'après avoir estimé le modèle GARCH (1.1) on vérifie que les résidus sont un bruit blanc et qui ne montrent pas de structure dans leur variance résiduelle. La validation du modèle de la moyenne conditionnelle s'effectue par l'analyse du corrélogramme des résidus standardisés et l'examen du

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

corrélogramme des carré des résidus standardisés pour confirmer et valider le modèle de la variance conditionnelle. La validation du modèle GARCH passe par des étapes principales :

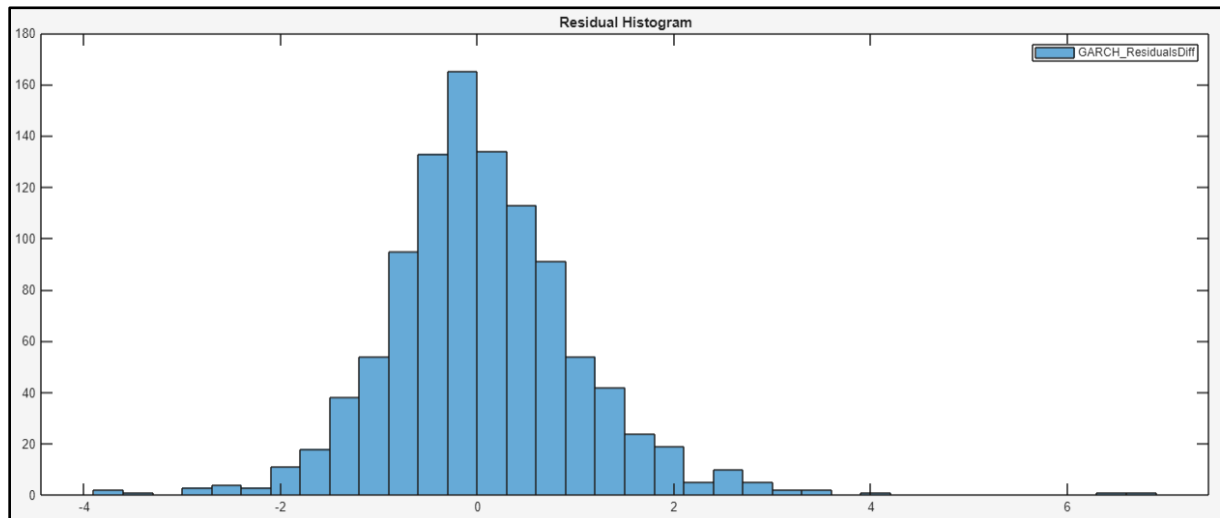
2.1.4.1 Vérification des Résidus Standardisés

Les résidus obtenus de l'estimation du modèle GARCH sont appelés résidus standardisés, définis comme : $z_t = \frac{\varepsilon_t}{\sigma_t}$

z_t sont les résidus standardisés qui devraient théoriquement être identiquement distribués et indépendants (iid), et suivre une distribution normale standard $N(0,1)$. Pour vérifier on trace l'histogramme et on calcule le test statistique de Jarque-Bera.

On relève de l'histogramme 1 que les résidus standardisés à peu près symétrique autour de son centre ce qui suggère que les données suivent une distribution normale indiquent que le modèle GARCH capture une partie importante des dynamiques de la série et suivent un bruit blanc non gaussien. Avec un test de Jarque-Bera qui donne un p-value inférieure de 0.05 qui confirme la distribution des résidus standardisés du modèle GARCH (1.1).

Histogramme 1:Résidus standardisés de l'estimation GARCH (1.1).



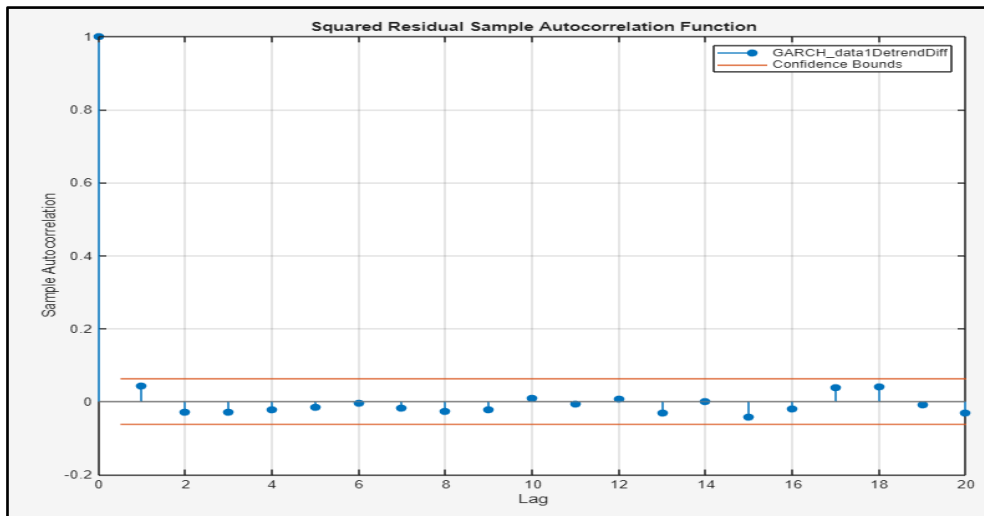
Source : Réalisée par l'auteur dans le programme Matlab2024a.

2.1.4.2 Autocorrélation des Résidus Standardisés et le Carré des résidus standardisés

On utilise le test de Ljung-Box sur les résidus standardisés et sur leurs carrés pour vérifier l'absence d'autocorrélation. Et on fait un examen du corrélogramme des résidus.

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Corrélogramme 4: Résidus standardisés et le carré des résidus standardisés.



Source : Réalisée par l’auteur dans le programme Matlab2024a.

Tableau 9: Test Ljung-Box des résidus standardisés.

Ljung-Box Q-Test for Autocorrelation(GARCH_data1DetrendDiff)								
Null Hypothesis: The first m autocorrelations of the residuals of GARCH_data1DetrendDiff are jointly 0								
Results								
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Lags	DOF	Significance Level
1	<input type="checkbox"/>	false	0.1593	26.1998	31.4104	20	20	0.0500

Source : Réalisée par l’auteur dans le programme Matlab2024a.

Le Corrélogramme des carrés des résidus standardisés et les résidus standardisés montre des fonctions d’autocorrélation dans l’intervalle de confiance qui signifie la non-corrélation des résidus confirmé par le teste statistique de Ljung-Box qui donne une valeur égale à 26.1998 avec un p-value égale à 0.1593 donc on peut valider le modèle de la variance conditionnelle.

2.1.4.3 Test ARCH sur les Résidus Standardisés

Ce test évalue la présence d'hétéroscédasticité résiduelle non modélisée dans les résidus standardisés. Si le modèle GARCH_(1,1) est correctement capturé la volatilité conditionnelle, il ne devrait pas y avoir de l’effet ARCH significatif restant dans les résidus. On applique le teste de Engle’s ARCH sur les résidus standardisés obtenus de l’estimation GARCH_(1,1) le tableau **10** montre les résultats du test :

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

Tableau 10: Test Engle's ARCH sur les résidus standardisés.

Engle's ARCH Test for Heteroscedasticity(GARCH_ResidualsDiff)							
Null Hypothesis: The residuals of GARCH_ResidualsDiff exhibit no ARCH effects							
Results							
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Lags	Significance Level
1	<input type="checkbox"/>	false	0.1570	2.0030	3.8415	1	0.0500

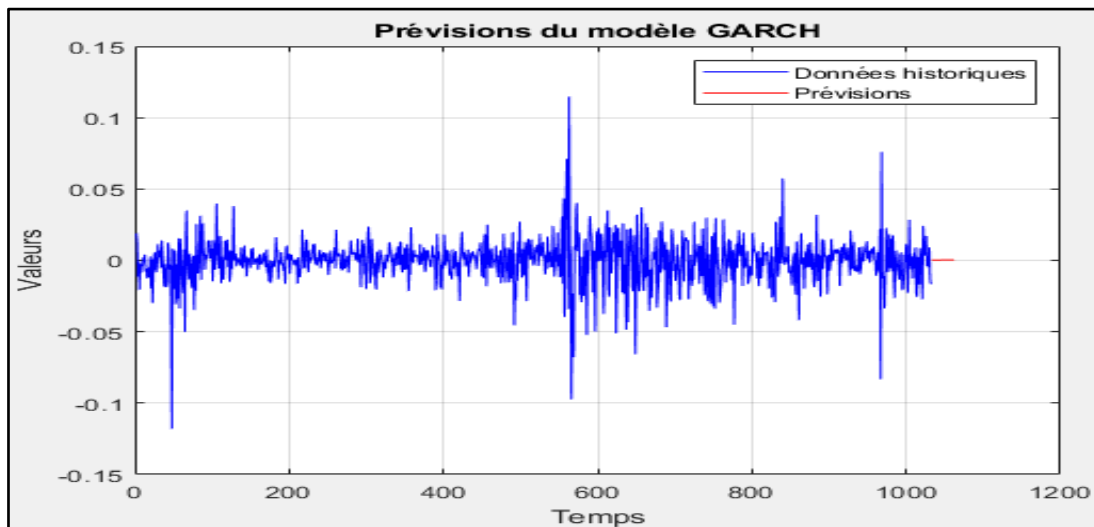
Source : Réalisée par l'auteur dans le programme Matlab2024a.

Le teste donne un p-valus égale à 0.1570 qui est supérieure à 0,05 on accepte l'hypothèse de l'homoscédasticité cela indique que l'effet ARCH est correctement modélisé et qu'il n'y a pas d'effet ARCH qui reste dans les résidus d'où l'adéquation du modèle estimé GARCH (1.1).

3 La Prévion de la volatilité des prix du pétrole par un modèle GARCH (1.1)

Dans cette étape on utilise le modèle estimé pour prévoir la volatilité future des prix du pétrole. Les prévisions se fait pour une période de 30 jours (court terme) la figure N°11 montre les prévisions de la volatilité conditionnelle.

FIGURE 11:Prévisions de la volatilité conditionnelle à l'horizon de 30jours.



Source : Réalisée par l'auteur dans le programme Matlab2024a.

La variance conditionnelle prévue obtenus du modèle GARCH (1.1) représente la volatilité anticipée pour la prochaine période de 30 jours. Les résultats obtenus du calcul des prévisions et la visualisation des résultats montre une faible variance conditionnelle suggère une volatilité prévue plus faible égale à 0.09, ce qui implique des variations de prix plus stables dans l'avenir entre (70\$ et 80\$). Une fois les variances conditionnelles prévues calculées, on peut obtenir la

Section 2 : La modélisation de la volatilité des prix du pétrole par le modèle GARCH

moyenne de ces valeurs. Cette moyenne donne une estimation globale de la volatilité prévue sur l'horizon de prévision qu'on a choisi 30 jours avec une moyenne légère la variance conditionnelle suggère une volatilité modérée dans les prévisions.

Conclusion

A la fin des prédictions on conclut que le modèle GARCH_(1,1) fournit des prédictions utiles de la volatilité future des prix du pétrole, Toutefois, il est important de noter que bien que le modèle capture efficacement la dynamique de la volatilité, il ne prévoit pas directement les prix futurs, mais plutôt l'incertitude associée à ces prix. Malgré sa capacité à modéliser la volatilité comme tous les modèles classiques et d'une nature linéaire il présente certaines limites. En supposant des chocs de volatilité avec un impact symétrique, ce qui ne peut pas être toujours le cas des prix du pétrole. Ainsi, pour une analyse complète, il pourrait être combiné avec d'autres modèles de prévision des prix ce qui est l'objectif de notre travail.

Section 3

La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Sommaire

SECTION 3 : LA MODELISATION DE LA VOLATILITE DES PRIX DU PETROLE PAR LE MODELE NAR.....	90
INTRODUCTION.....	90
1 LA STRUCTURE DU MODELE NAR.....	90
2 LA MODELISATION DES PRIX DU PETROLE PAR UN MODELE NAR	90
a. L'architecture du réseau neurone	92
b. Entraînement du modèle NAR.....	93
c. Test de la performance du réseau de neurone NAR	93
3 LA PREVISION DES PRIX DU PETROLE PAR LE NAR	99
d. Prévision muti-step	99
e. Prévision un pas Step-Ahead Prediction Network.....	100

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Introduction

L'étude de la volatilité des prix du pétrole à l'aide des réseaux de neurones NAR (Nonlinear Autoregressive) est une méthode puissante pour modéliser et prévoir des données complexes où des relations non linéaires peuvent être présentes. Les réseaux de neurones NAR sont adaptés pour modéliser des séries temporelles univariées, où les valeurs futures dépendent uniquement des valeurs passées. La technique NAR peut capturer des relations non linéaires dans les données, ce qui est souvent le cas pour les fluctuations des prix du pétrole.

Dans cette section on va faire une modélisation de la volatilité des prix du pétrole par un modèle NAR seulement pour voir les résultats des prévisions par rapport aux résultats obtenus de la modélisation GARCH puis on fait une hybridation des deux modèles pour améliorer la qualité des prévisions de la volatilité des prix du pétrole.

1 La structure du Modèle NAR

Un modèle NAR prédit la valeur future d'une série temporelle en utilisant une fonction non linéaire des valeurs passées. Le modèle peut être représenté comme suit :

$$y_t = f(y_{(t-1)}, y_{(t-2)}, \dots, y_{(t-d)}) + \varepsilon_t.$$

Où

y_t la valeur de la série temporelle à l'instant t .

f la fonction non linéaire modélisée par un réseau neurones.

d le nombre de retards ou "lags".

ε_t un terme d'erreur.

2 La modélisation des prix du pétrole par un modèle NAR

Dans la section précédente de la modélisation GARCH on a discuté la représentation graphique des données des prix du pétrole et on a vu que la série présente une forte volatilité. Comme on a conclu que la série contient la composante trend et présente une non stationnarité.

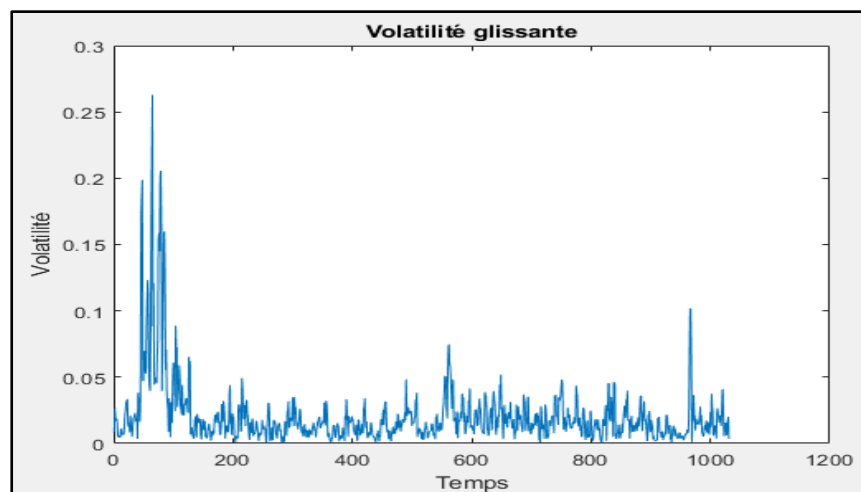
Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Pour faire une modélisation NAR on doit traiter les données des prix du pétrole pour extraire la volatilité glissante de la série des prix brute, on calcule la volatilité à partir de la série normalisée et différenciée ce qui permet d'obtenir des données stationnaires adaptées à la modélisation. Il est nécessaire d'éliminer la tendance et rendre la série stationnaire et cela à cause de la capacité des réseaux neurones de capturer mieux toutes les relations complexe existant dans les données et ils sont utilisé pour modéliser les non-linéarités. La normalisation des données consiste à mettre les données dans une certaine échelle, comme $[0,1]$ ou $[-1,1]$ sans changer la structure ou la nature des relations dans la série. Et la différenciation transforme la série en calculant les variations entre des points consécutifs.

Après cette étape on a calculé la volatilité des prix à partir de la série traitée appelé la volatilité glissante qui est défini comme la mesure de variabilité ou la dispersion d'une série temporelle calculée sur une fenêtre temporelle spécifique qui se déplace au fil du temps. Elle reflète la manière dont la volatilité change localement dans le temps. Elle est caractérisée par l'évolution dans le temps car la fenêtre glissante inclut de nouvelles observations et exclut les anciennes.

Pour cette volatilité on a choisi une fenêtre courte de 5 réagit plus rapidement aux changements brusques pour pouvoir évaluer la capacité du NAR à capturer la dynamique complexe de cette série la visualisation de la volatilité glissante est dans la figure N°12.

Figure 12:La volatilité Glissante des prix du pétrole.

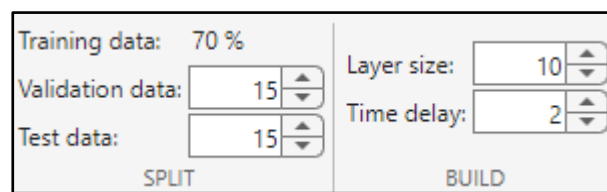


Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

On observe une forte volatilité au début atteignant des valeurs élevées ce qui indique à une période d'incertitude dans les prix « la pandémie COVID-19 ». Suivie d'une période intermédiaire où la volatilité diminue progressivement et reste relativement faible. Cette période semble être plus stable, avec des fluctuations modérées. Elle peut refléter une période d'équilibre sur le marché cette volatilité reste globalement faible dans la fin de la période d'étude on remarque des pics indiquant à des épisodes isolés de forte volatilité, qui pourraient être liés à des événements spécifiques. Un modèle NAR pourrait mieux s'adapter à ces comportements.

On commence notre modélisation par la séparation de la série en des groupes de données : Entraînement, validation, et test. On divise les données de la série en trois ensembles : Ensemble d'entraînement (70%) comprend à 720 observations, Ensemble de validation (15%) comprend à 155 observations, et ensemble de test (15%) comprend à 155 observations. Ce qui permet de évaluer la robustesse et la performance du modèle et de prévenir le surapprentissage. Cette division est effectuée automatiquement par Matlab R2024a.

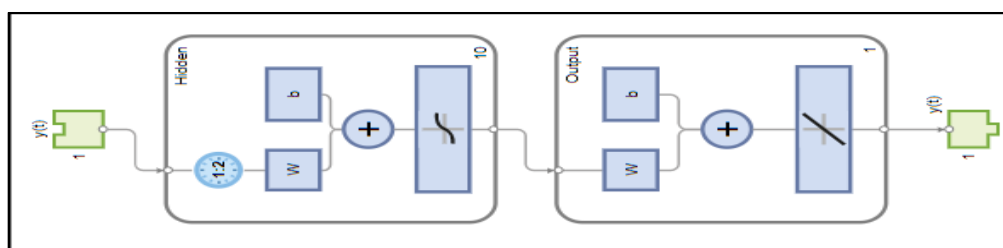


Pré-traitement et division des données par Matlab R2024a.

a. L'architecture du réseau neurone

L'architecture proposer par Matlab est présenté comme le montre le schéma dans la figure N°13 :

Figure 13:L'architecture du NAR (1 :2) (10) des prix du pétrole.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Selon la figure 13 on peut retenir que le réseau de neurones autorégressif se compose de trois couches principales une couche d'entrée (Input Layers) qui représente la série des prix du pétrole avec un décalage de 2 (time delay) afin que le réseau reconnaisse les caractéristiques des données pour l'utiliser dans le processus de prédiction. Une couche cachée (hidden layers) qui traite les données reçues de la couche d'entrée et extrait des caractéristiques complexes. Cette couche effectue des transformations non linéaires sur les données d'entrée leur nombre est déterminé par le type d'étude et le nombre d'entrées, qui sont souvent faibles, on appuiera donc dans ce modèle sur 10 (layer size) neurones cachés. La Couche de sortie (Output Layer) Cette couche fournit les résultats finaux du modèle, et puisque notre objectif d'étude et la prévision donc l'Output layer est construite d'un seul neurone.

b. Entraînement du modèle NAR

Entraînement du modèle est effectué sur l'ensemble d'entraînement à l'aide de l'algorithme **Levenberg-Marquardt**. Les résultats de l'entraînement est affiché dans le tableau N°11.

c. Test de la performance du réseau de neurone NAR

La performance d'un modèle NAR (Nonlinear Autoregressive) est évaluée à travers plusieurs métriques pour s'assurer qu'il est capable de bien prédire les valeurs futures d'une série temporelle. Les métriques de performance pris par Matlab R2024a sont :

La moyenne des carrés des erreurs MSE (la moyenne des carrés des erreurs entre les valeurs prédites et les valeurs réelles) : $MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$.

Le coefficient de détermination (R²) indique la proportion de la variance des données qui est expliquée par le modèle. Un R² proche de 1 indique que le modèle explique bien les données.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Tableau 11: Résultat d'entraînement et les statistiques d'ajustement du réseau de neurone.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

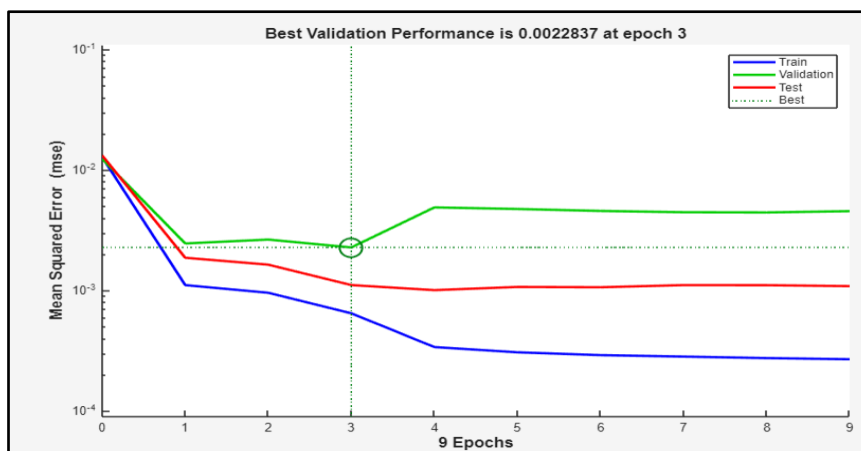
Layer size:	10		
Time delay:	2		
	Observations	MSE	R
Training	721	0.0006	0.9072
Validation	154	0.0023	0.9096
Test	154	0.0011	0.8682

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

L'entraînement du modèle NAR pour l'ensemble d'entraînement un MSE égale à 0.0006 l'erreur de prédiction sur cet ensemble est faible ce qui est un signe que le modèle a bien appris à prédire les données d'entraînement. Un R= 0.9072 proche de 1 signifie une forte corrélation entre les valeurs prédites et les valeurs réelles des données d'entraînement et montre que le modèle capture bien la dynamique des données. Et pour l'ensemble de validation on peut remarquer que le MSE égale à 0.0023 un peu plus élevée par rapport à l'ensemble d'entraînement ce qui est normal car les données de validation servent à évaluer la capacité généralisante du modèle sur des données inconnus. Avec une corrélation forte, similaire à l'entraînement ce qui indique que le modèle généralise bien sur les données de validation ce qui est essentiel pour éviter le surapprentissage. Et pour l'ensemble de test le MSE égale à 0.0011 une valeur faible par rapport à l'ensemble de validation, ce qui est un bon signe indique que le modèle prédit bien les données de test avec une corrélation légèrement plus faible que dans l'ensemble précédent indique que le modèle perd un peu de précision sur les données totalement inconnues. Donc on peut dire que le modèle est performant il n'y a pas un surapprentissage et généralise très bien sur des données nouvelles de manière efficace. La figure 14 montre l'évolution des MSE au cours de la phase d'apprentissage.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Figure 14: Performance du réseau de neurone.



Source : Réalisée par l’auteur dans le programme Matlab R2024a.

Le graphe montre la meilleure performance de validation du modèle a été atteinte à l’époque 3 (ou itération 3 sur l'ensemble de données d'entraînement) égale à 0.0022837 cette valeur représente MSE qui indique à quel point le modèle se comporte bien sur l'ensemble de validation qui est un ensemble de données séparé non utilisé pour l'entraînement direct.

On peut remarquer que le coefficient de corrélation dans l’ensemble de Test est un peu plus bas ($R=0.8682$). Cela pourrait être amélioré en ajustant des hyperparamètres comme le nombre de neurones de la couche cachée ou la taille des décalages (lags), ou en fournissant plus de données d'entraînement pour plus de précision.

Tableau 12: Résultat d’entraînement et les statistiques d’ajustement du réseau de neurone NAR (1:1)(10).

Layer size:	10		
Time delay:	1		
	Observations	MSE	R
Training	720	0.0009	0.9049
Validation	155	0.0006	0.8974
Test	155	0.0018	0.9431

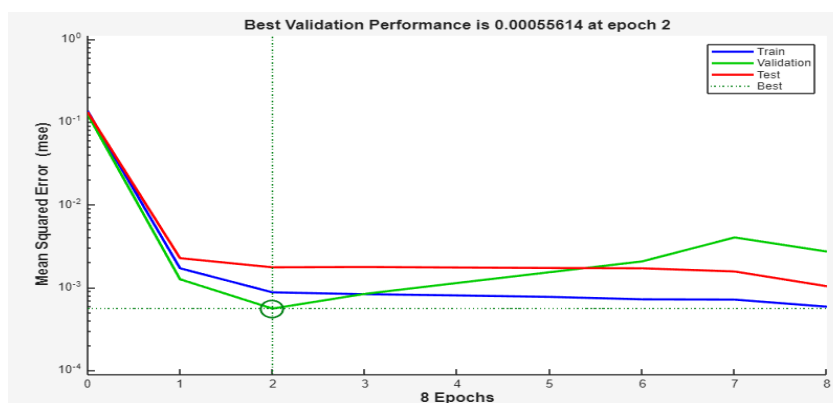
Source : Réalisée par l’auteur dans le programme Matlab R2024a.

Sur l’ensemble d’entraînement on a obtenu un $MSE = 0.0009$ qui est une valeur faible mais légèrement augmenté par rapport à MSE du modèle précédent ($0.0006 \rightarrow 0.0009$) ce qui est normal après avoir réduit le nombre de retards ($2 \rightarrow 1$) avec une corrélation $R = 0.9049$ élevée

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

et proche de 1, ce qui signifie que le modèle a toujours bien appris les relations sur l'ensemble d'entraînement mais a légèrement perdu en précision. Les résultats sur l'ensemble de validation montrent un faible MSE = 0.0006 considérablement diminué par rapport au modèle précédent (0.0023 → 0.0006), ce qui indique une amélioration notable de la capacité généralisante du nouveau modèle avec une corrélation toujours élevée proche de celle de l'ensemble d'entraînement $R = 0.8974$ on peut dire que le modèle généralise beaucoup mieux sur les données de validation. Et pour l'ensemble de Test l'erreur a augmenté légèrement (0.0011 → 0.0018) mais reste toujours faible et corrélation plus élevée qu'avec le précédent modèle (0.8682 → 0.9431) indique que le modèle est très performant sur des données inconnues. Donc le modèle a nettement amélioré sa précision sur l'ensemble de test, ce qui est un excellent signe pour des prévisions fiables et une meilleure généralisation. En générale on peut retenir du nouveau modèle ajusté qu'il y a une légère dégradation de la performance sur l'ensemble d'entraînement, ce qui est souvent acceptable lorsque cela améliore la généralisation. La figure 14 montre le meilleur MSE sur l'ensemble de validation pour le modèle NAR (1.10).

Figure 15: Performance du réseau de neurone NAR (1 :1)(10).



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

une meilleure performance de validation a été atteinte à l'époque 2 égale à 0.00055614 le modèle a amélioré sa précision ce qui est excellent pour des prévisions le modèle est performant mais cette performance a légèrement diminué sur l'ensemble d'entraînement on peut améliorer la performance en ajustant le nombre de neurones de la couche cachée et garder le même nombre de retard. Les résultats de l'entraînement sont rassemblés dans le tableau N°13.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Tableau 13: Résultat d'entraînement et les statistiques d'ajustement du réseau de neurone NAR (1:1)(14).

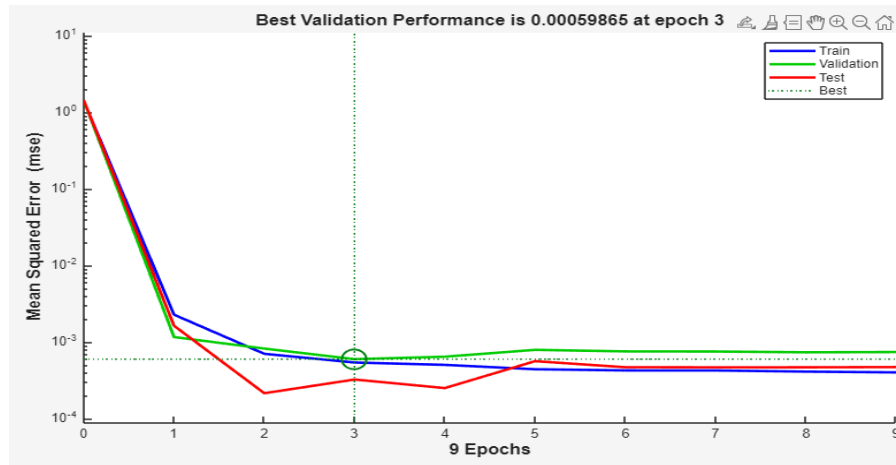
Layer size:	14		
Time delay:	1		
	Observations	MSE	R
Training	720	5.4135e-04	0.9455
Validation	155	5.9865e-04	0.9063
Test	155	3.2482e-04	0.9590

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

En analysant les résultats on peut remarquer une forte amélioration sur le modèle en générale, le modèle fonctionne très bien sur l'ensemble d'entraînement avec une valeur faible du MSE = 0.000541 et une excellente corrélation indiquant que le modèle a capturé la dynamique des données de manière très précise. Et aussi le modèle généralise bien sur les données de validation avec une valeur de MSE = 0.0005986 et un R= 0.9063. Même pour l'ensemble de Test l'erreur est la plus faible parmi les trois ensembles MSE = 0.000324 et un R= 0.9590 qui est plus élevée, ce qui est un excellent signe de robustesse du modèle NAR (1:1)(14). Le graphe de la meilleure performance du modèle NAR (1:1)(14) :

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

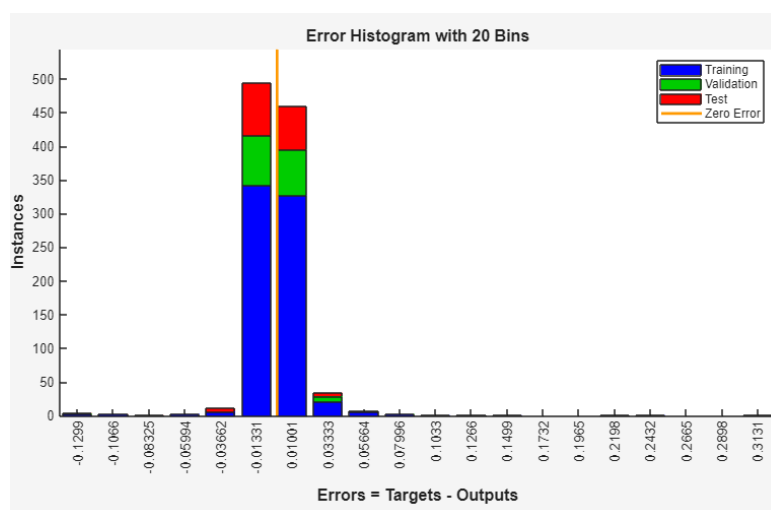
Figure 16: Performance du réseau de neurone NAR (1 :1)(14).



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Une meilleur performance de validation à été attient à l'époque 3 égale à 0.00059865 le modèle semble très performant à ce stade l'augmentation du nombre de neurones a permis au modèle de mieux capter les relations complexes dans les données, tout en maintenant une bonne capacité de généralisation sur les ensembles de validation et de test. À ce stade, on peut arrêter ici parce que l'objectif est d'obtenir un modèle stable et performant. On continu le diagnostique du modèle retenu pour l'utiliser dans la prévision de la volatilité des prix du pétrole.

Histogramme 2: histogramme des erreurs du modèle NAR (1 :1)(14).



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

L'histogramme montre la distribution des erreurs entre les valeurs prédites et les valeurs réelles. Les erreurs sont principalement concentrées autour de zéro avec une faible dispersion ce qui indique que le modèle fait de bonnes prédictions dans la majorité des cas et que le modèle a un biais faible et généralise bien sur les trois ensembles (entraînement, validation, test) ou autrement dit les erreurs semblent être symétriques autour de 0 ce qui suggère que le modèle ne présente pas de biais systématique. On peut aussi observer quelques erreurs éloignées de zéro sur les côtés ce sont des cas où le modèle a fait des prédictions moins précises mais leur fréquence est très faible ce qui limite leur impact global.

3 La prévision des prix du pétrole par le NAR

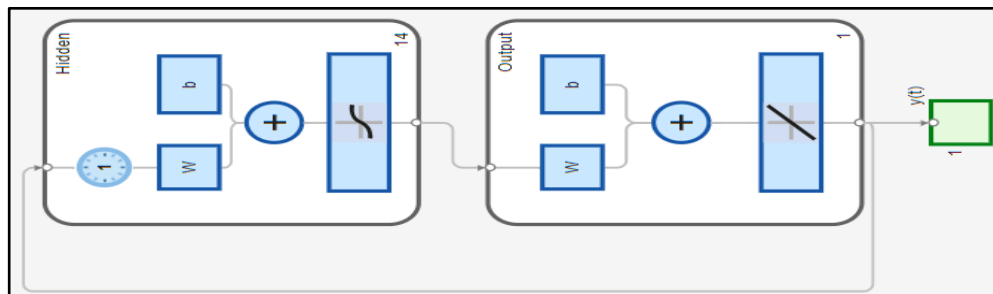
Le modèle NAR est un outil puissant pour la prévision des séries temporelles non linéaires, utilisant les forces et la puissance des réseaux neurones et les avantages des approches autoregressives. On a deux approches de prévision :

d. Prévision muti-step

Cette prévision s'effectue par un réseau en Boucle fermée (CloseLoop) c'est une approche utilisée pour effectuer des prévisions à plusieurs étapes en utilisant le réseau de neurone. Dans la prévision par un réseau en boucle fermée on n'utilise pas les données réelles comme entrées, les sorties prédites par le réseau pour une étape sont utilisées comme entrées pour la prédiction. Les prédictions du réseau sont réinjectées comme nouvelles entrées pour prédire les étapes suivantes. Cette méthode est essentielle pour les prévisions à long terme car elle permet au réseau de continuer à faire des prédictions au-delà de l'horizon d'entraînement, en se basant uniquement sur ses propres sorties, on a calculé la performance de ce réseau est égale à [closed Loop Performance = 0.0045]. La figure N° 16 présente l'architecture du réseau en boucle fermée du modèle.

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Figure 17: Le réseau en boucle fermé (Close Loop)

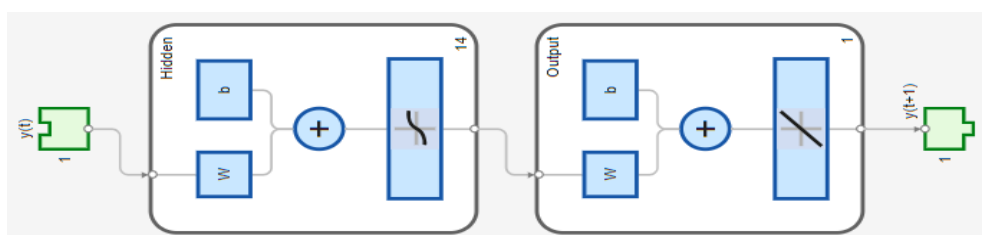


Source : Réalisée par l'auteur dans le programme Matlab R2024a.

e. Prédiction un pas Step-Ahead Prediction Network

Prédiction en Boucle Ouverte (Open Loop) c'est un réseau en boucle ouvert qui aide à avoir des prédictions un pas de temps plus tôt où le réseau utilise les observations passées pour prédire les valeurs futures. Cela fonctionne bien pendant l'entraînement parce que les données réelles sont disponibles à chaque étape. La performance de ce réseau est calculée par Matlab [stepAheadPerformance = 0.00091730]. La figure N°17 présente le réseau en boucle ouverte :

Figure 18: Le réseau en boucle ouverte (Open Loop)



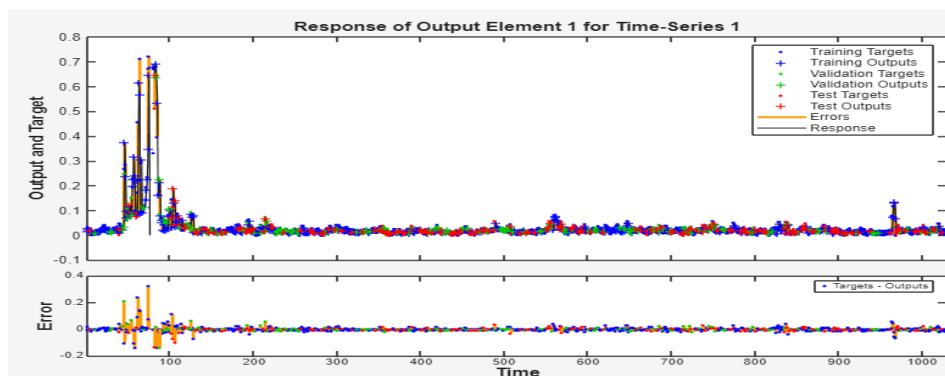
Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Les deux valeurs de MSE sont faibles ce qui indique un modèle performant et capable de faire des prédictions précises, aussi bien pour des prévisions directes step ahead que pour des prévisions closed loop. La performance en Close Loop égale à 0.0045 est environ 5 fois plus élevée que celle en "step ahead" ou du Open Loop = 0.00091730 ce résultat est raisonnable indique que des erreurs cumulées se produisent dans le réseau de boucle fermée. Cette différence est relativement modérée ce qui montre que le modèle reste stable même en Close

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

Loop. Le MSE= 0.00091730 est très faible ce qui reflète une excellente performance à court terme.

Figure 19:Response des éléments de sortie



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

La figure N°19 montre deux graphes un dans la partie supérieure représente Target et les outputs et une partie inférieure représente les erreurs « Error ».

Le graphe supérieur compare les valeurs de la série réelles « Targets » et les sorties du modèle « Outputs » pour les ensembles d'entraînement, de validation et de test. On peut remarquer que les valeurs de la série réelle et prédite sont presque identiques les sorties du modèle sont exactement identiques à celles de l'évaluateur d'efficacité pendant toutes les étapes (Training, Validation, Test). Au début de la série entre 0 et 150 il y a des pics remarquables peut-être des anomalies ou des fluctuations importantes. Mais les sorties du modèle restent proches des valeurs réelles. Tous les indicateurs présentés ci-dessous indiquent l'efficacité et l'efficacité du modèle dans le processus de l'apprentissage et le test.

Le graphe inférieur montre l'erreur résiduelle, calculée comme la différence entre les valeurs réelles et les sorties (Targets - Outputs) on peut remarquer que la majorité des erreurs sont proches de zéro ce qui indique une bonne performance du modèle et pour les erreurs les plus élevées sont concentrées dans la zone des pics initiaux indique la difficulté de modéliser des variations brusques.

On conclut que Le modèle est globalement performant et fonctionne bien pour la majorité des données avec des erreurs faibles et une bonne correspondance entre les Target et les sorties donc on peut l'utiliser pour la prévision. On a calculé la valeur de la volatilité des prix du pétrole

Section 3 : La Modélisation de la volatilité des prix du pétrole par le modèle NAR

pour un pas en avant 29/12/2023 $VP = 0.0064$ la volatilité réelle de ce jour $VR = 0.0035$ la performance de cette prévision est calculer par la différence des deux valeurs $Erreur = - 0.0029$

Cela signifie que le modèle à légèrement surestimer les valeurs réelles. La magnitude de l'erreur est très faible ce qui indique que le modèle est globalement précis et qu'il capture efficacement la dynamique non linéaire dans les données.

Section 4

La Modélisation des prix du pétrole par le modèle GARCH_{NN}

Sommaire

SECTION 4 : LA MODELISATION DE LA VOLATILITE DES PRIX DU PETROLE PAR LE MODELE GARCH-NN	105
INTRODUCTION.....	105
1 LA PREVISION PAR LE MODELE GARCH-NN	105
a. La modélisation GARCH.....	105
b. Extraction des résidus et des volatilités conditionnelles.....	105
c. Conception du Réseau de Neurones	106
2 VALIDATION DU MODELE GARCH-NN (1.1) (1 :1) (10).....	109
3 LA PREVISION DE LA VOLATILITE DES PRIX DU PETROLE.....	112
CONCLUSION	115

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

Section 4 : La Modélisation de la volatilité des prix du pétrole par le modèle GARCH-NN

Introduction

L'utilisation d'un modèle hybride GARCH-NN combine les avantages et la force des modèles traditionnels de séries temporelles GARCH et la flexibilité des réseaux de neurones artificiels (NN). Le modèle hybride GARCH-NN peut mieux capturer la dynamique complexe des données de la série étudiée et modéliser les résidus non capturés par le modèle GARCH, en améliorant la précision des prévisions (réduire les erreurs de prévision) et la performance globale du modèle en capturant des patterns qui évoluent dans le temps.

1 La prévision par le modèle GARCH-NN

Les étapes de cette modélisation dépendent des modèle pré-estimé dans les sections précédentes.

a. La modélisation GARCH

On commence par la modélisation des fluctuations des prix du pétrole par le modèle GARCH. Les données sont traitées et transformées pour la stationnarité et éliminé la trend et sont modélisé par un modèle GARCH (1.1) présenté dans l'équation suivante :

$$h_t = 0.0235 + 0.2110\varepsilon_{t-1}^2 + 0.7890h_{t-1}.$$

h_t Présente la variance conditionnelle à l'instant t.

b. Extraction des résidus et des volatilités conditionnelles

Une fois le modèle GARCH ajusté et estimé on extrait les résidus standardisés ε_t ainsi que les volatilités conditionnelles σ_t . Et on prépare les données des résidus standardisés et les volatilités conditionnelles pour les utilisés comme entrées pour le réseau de neurones. Il faut convertir et organiser les données des résidus en série temporelle. Dans cette étape, les données sont divisées en groupes tels que :

- Groupe d'entraînement : Il comprend 70% des observations 721 observations.

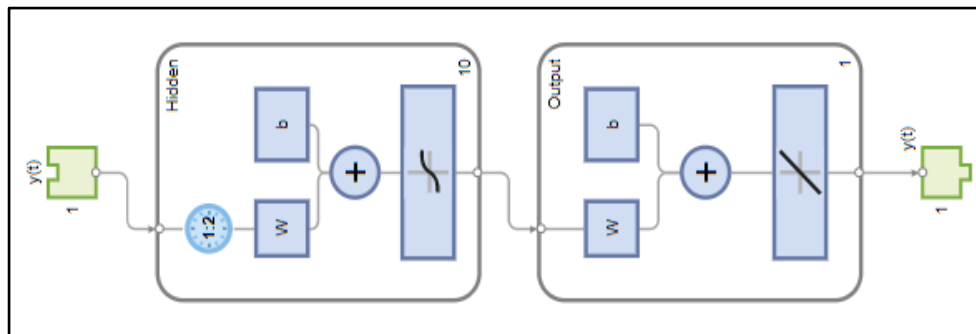
Section 4 : La Modélisation des prix du pétrole par le modèle *GARCH_NN*

- Groupe de vérification : Il comprend 15% des observations 154 observations.
- Groupe de test : Il comprend également 15% des observations 154 observations.

c. Conception du Réseau de Neurones

L'architecture du réseau est donnée automatiquement par MATLAB R2024a après la sélection de la série des résidus standardisés obtenus de la modélisation GARCH_(1.1) comme une variable d'entrée du réseau. Après on choisit la fonction d'apprentissage de la liste présentée par MATLAB R2024a. on a choisi l'apprentissage par l'algorithme de **Levenberg-Mmarquardt** (prend un temp plus rapide). L'architecture primère retenu par MATLAB est :

Figure 20: Architecture du NAR des résidus standardisés.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Selon la figure précédente, le réseau neurone a été conçu pour coder les entrées comme une seule couche cachée. Avec **10** neurones avec un degré de décalage introduit dans le programme Lag :**2**. Et une couche de sortie. Les résultats de l'entraînement sont consolidés dans le tableau **14** présenté ci-dessous :

Tableau 14: Résultats d'entraînement des résidus standardisés (1 :2) (10).

Layer size:	10		
Time delay:	2		
	Observations	MSE	R
Training	721	1.7830e-08	0.9234
Validation	154	1.4127e-07	0.8774
Test	154	5.6277e-08	0.5614

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

Le tableau 14 montre des coefficients de corrélations et les MSE des trois ensembles. Pour l'ensemble d'entraînement on peut remarquer un MSE très faible ce qui indique que le modèle ajuste bien les données d'entraînement avec une forte corrélation entre les valeurs réelles et prédites. Pour l'ensemble de validation le MSE est faible mais légèrement plus élevée par rapport à l'entraînement avec un R faible que l'ensemble précédent mais toujours élevé ce qui suggère que le modèle généralise correctement sur l'ensemble de validation. L'ensemble de Test affiche un MSE nettement plus élevée que celle sur l'entraînement et la validation indiquant une possibilité de sur-apprentissage et aussi un coefficient de corrélation très faible $R = 0.5614$ suggère que le modèle est moins performant pour capturer la relation entre les prédictions et les valeurs réelles dans l'ensemble de test. Cela peut entraîner des résultats non fiables. Une amélioration est suggérée sur le modèle pour éviter le sur-apprentissage en réduisant le nombre de retard de deux au 1 et on refait l'entraînement du modèle les résultats dans le tableau N° 15 suivants :

Tableau 15: Résultats d'entraînement des résidus standardisés (1 :1) (10).

Layer size:	10		
Time delay:	1		
	Observations	MSE	R
Training	720	3.0652e-08	0.9045
Validation	155	3.0848e-08	0.9388
Test	155	1.6233e-08	0.9634

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

On peut remarquer du tableau 15 que le modèle est amélioré après la diminution de nombre de retard. L'amélioration est dans les deux métriques de performances les erreurs quadratiques et les coefficients de corrélation pour chaque ensemble.

Interprétation des résultats du tableau :

- Ensemble d'entraînement (Training) un $MSE = 3.0652 \times 10^{-8}$ une valeur faible indique à un bon ajustement et suggère que le modèle capture bien les caractéristiques et la structure des données d'entraînement. Un $R = 0.9045$ une valeur élevée proche de 1 indique une forte corrélation positive signifie que les prédictions du modèle sont bien alignées avec

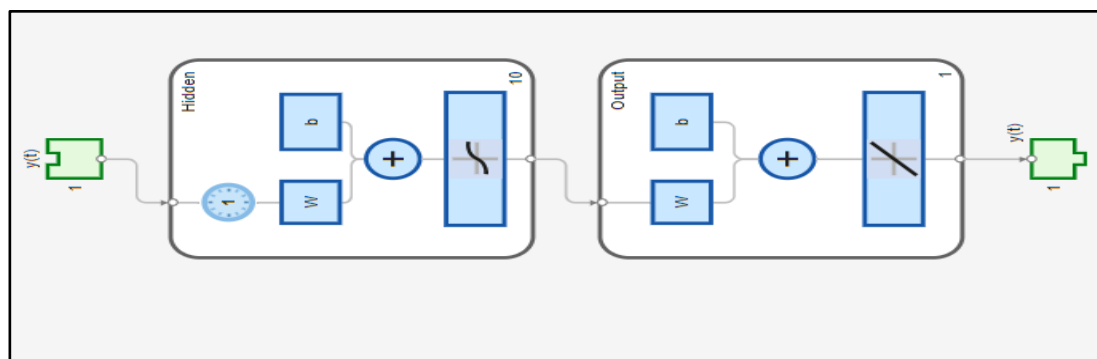
Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

les valeurs réelles sur l'ensemble d'entraînement et que le modèle a pu capturer la relation entre les valeurs réelles et prédites.

- Ensemble de Validation un $MSE = 3.0848 \cdot 10^{-8}$ la valeur est très faible ce qui indique que le modèle a une excellente précision sur l'ensemble de validation. Avec un $R = 0.9388$ reflète une très forte corrélation positive entre les prédictions du modèle et les valeurs réelles sur l'ensemble de validation. Donc le modèle capture bien la structure sous-jacente des données de validation.
- Ensemble de Test un $MSE = 1.6233 \cdot 10^{-8}$ des erreurs de prédiction extrêmement faibles sur l'ensemble de test. Suggère que le modèle fait un excellent travail pour prédire des données inconnues. Le même ensemble donne un $R = 0.9634$ une forte corrélation signifie que les variations des prédictions suivent étroitement celles des valeurs réelles.

Les MSE ont diminué dans tous les ensembles (entraînement, validation et test). Cela indique que le modèle est maintenant plus précis et parvient mieux à s'ajuster aux données, probablement grâce à une complexité réduite liée à l'utilisation d'un seul retard. Le coefficient de corrélation R a augmenté, en particulier pour l'ensemble de test, où il atteint $R = 0.9634$ cela montre une très forte corrélation entre les prédictions et les valeurs réelles. Ces résultats montrent que le modèle est performant et fiable pour prédire sur des données externes et que le modèle est bien entraîné. Avec cette performance, le modèle GARCH-NN $(1,1) (1:1)_{10}$ semble bien ajusté et généralisé. L'architecture du réseau est présentée dans la figure N° 21.

FIGURE 21: Architecture du réseau de neurone ajusté.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

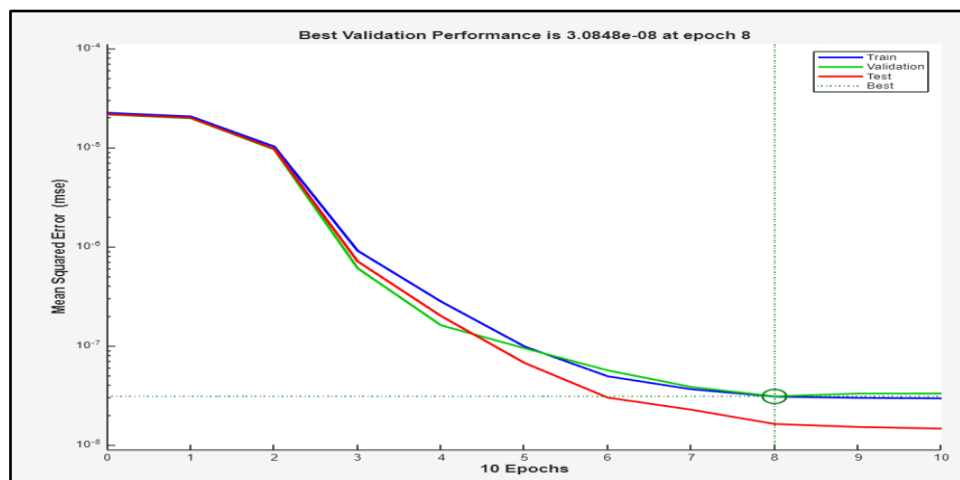
Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

Un examen visuel sur les valeurs réelles et prédite est nécessaire pour valider la performance du modèle retenus avant son exploitation dans la prévision.

2 Validation du modèle GARCH-NN (1.1) (1 :1) (10)

L'entraînement du modèle GARCH-NN (1.1) 1 :1) (10) atteint la meilleure performance sur les données de validation à la 8 -ème éviration suggère que le réseau a convergé rapidement ce qui indique que la structure du réseau est bien adaptée et que les données d'entraînement sont bien prétraitées avec une valeur de 3.08448×10^{-8} qui est extrêmement faible indiquant la capacité du réseau à modéliser avec une précision remarquable les relations dans les données de validation.

FIGURE 22:La performance du réseau de neurone des résidus standardisés.

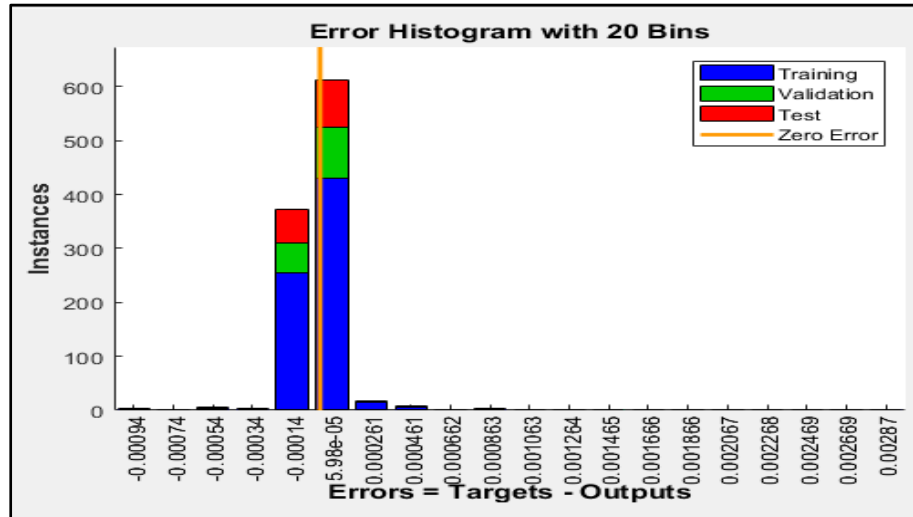


Source : Réalisée par l'auteur dans le programme Matlab R2024a.

L'analyse de l'histogramme des erreurs de prévision avec 20 Bins permet d'évaluer la distribution des écarts entre les valeurs prédites par le modèle et les valeurs réelles.

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

Histogramme 3: Erreur de prévision du modèle.



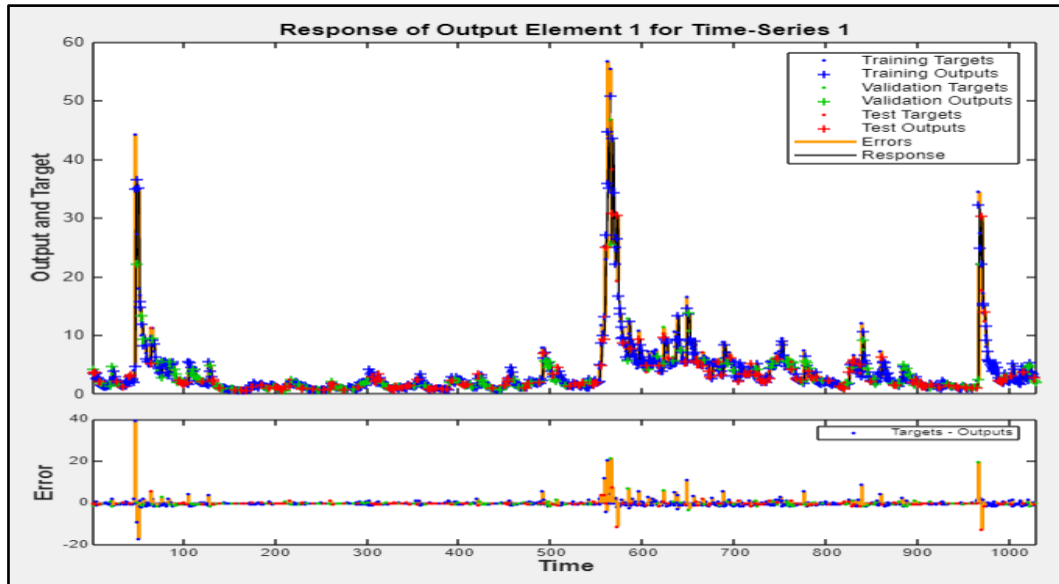
Source : Réalisée par l'auteur dans le programme Matlab R2024a.

On peut remarquer que la majorité des erreurs se trouvent dans les bins proches de zéro ce qui indique que le modèle est globalement précis, avec des prévisions très proches des valeurs réelles pour la plupart des données. Quelques erreurs éloignées dû à des anomalies dans les données que le modèle n'a pas pu bien capturer. Donc on peut conclure que Le modèle montre une performance solide avec des erreurs principalement concentrées autour de zéro, indiquant une bonne précision des prévisions. Quelques erreurs rares et éloignées pourraient nécessiter une analyse plus approfondie.

Le modèle est bien adapté aux données de la volatilité des prix du pétrole et lors de l'analyse de la figure N° 23 qui est généralement utilisé pour évaluer la performance d'un réseau de neurones appliqué à une série temporelle.

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

FIGURE 23: Response of output élément 1 for Time series.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

La figure N°23 représente un graphique qui compare les valeurs réelles de la série des résidus standardisés avec les valeurs prédites par le modèle NAR. La figure contient de graphes.

La partie supérieure de la figure représente les valeurs prédite (output) et les valeurs réelle (Target), on peut voir que outputs du modèle s'aligne correctement avec les données réelles cela montre que le réseau a bien appris les relations dans les données et bien ajusté à la dynamique de la série.

La partie inférieure de la figure représente les erreurs de prédiction. La courbe d'erreur est centrée autour de 0 avec des fluctuations limitées ce qui indique que le modèle est globalement performant et que les prédictions sont proches des valeurs réelles. Des pics remarquables dans la courbe d'erreur ce qui peut indiquer que le modèle ne parvient pas à capturer certains comportements spécifiques dans les données, comme des changements imprévisibles.

L'analyse et le diagnostic réalisés mettent en évidence la capacité du modèle combiné GARCH-NN à modéliser efficacement une série temporelle hautement volatile, comme celle

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

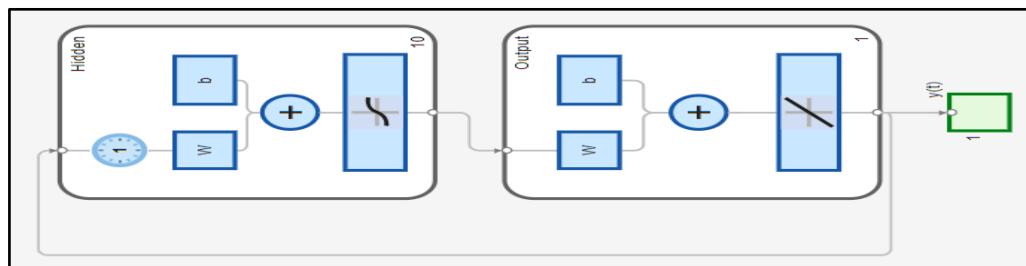
des prix du pétrole. La décomposition de la volatilité par le modèle GARCH, suivie de son exploitation par le réseau de neurones NAR, a démontré une excellente performance en termes de précision et de généralisation. Cette approche hybride s'est avérée robuste, permettant de capturer à la fois les dynamiques non linéaires et les comportements volatils caractéristiques de la série. Les résultats obtenus confirment que cette combinaison est adaptée pour prédire des séries complexes et volatiles, offrant ainsi un outil fiable pour l'analyse et la prise de décision.

3 La prévision de la volatilité des prix du pétrole

Après la validation et l'ajustements du modèle on a utilisé le réseau neurone avec l'architecture approuvée pour anticiper les fluctuations des prix du pétrole. On a vu dans la section précédente que Matlab donne deux approches pour effectuer des prévisions :

L'approche **Close Loop** (réseau en boucle fermée) cette approche est représenté par l'architecture de la figure N°24 et donne une valeur de performance [closedLoopPerformance = 1.6310×10^{-7}]. Une valeur faible indique à la bonne précision du modèle mais cette valeur doit être comparé avec celle du réseau Open Loop.

FIGURE 24: Réseau en boucle fermée.

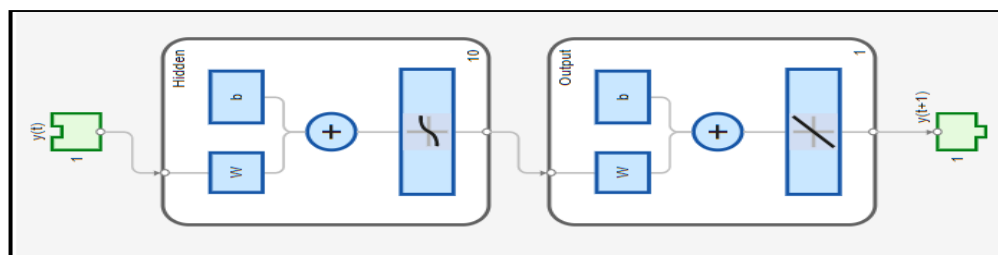


Source : Réalisée par l'auteur dans le programme Matlab R2024a.

La deuxième approche c'est le réseau **Open Loop** (réseau en boucle ouvert) qui calcule les prédictions d'un pas en avant. Cette approche donne une valeur de performance égale à [stepAheadPerformance = 3.0254×10^{-8}] une valeur extrêmement faible ce qui signifie que l'écart entre les valeurs réelles et prédites est très petit ce qui indique d'une grande précision dans la prédiction à un pas. La figure N°25 présente l'architecture Open Loop.

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

FIGURE 25:Réseau en boucle ouverte.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

En comparant les deux réseaux on peut conclure que la performance en boucle fermée $1.6310e-7$ est légèrement moins bonne que celle de la boucle ouverte $3.0254e-8$ ce qui est attendu. En boucle fermée le modèle doit s'appuyer uniquement sur ses propres sorties ce qui peut entraîner une légère accumulation d'erreurs. La boucle ouverte utilise les valeurs réelles ce qui réduit le risque des erreurs cumulées.

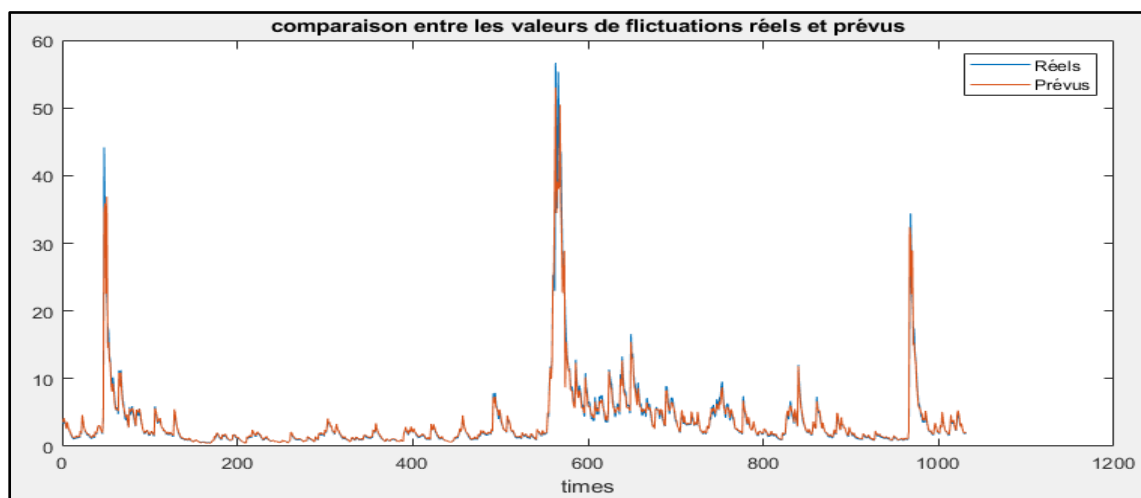
Le réseau qui présente une faible valeur de MSE (performance) est le réseau en boucle ouverte donc on va l'utiliser pour faire les prévisions. On a utilisé les valeurs passées de la suite des résidus obtenus du modèle GARCH (1.1) nommés **XS** cette série contiens les valeurs réelles des étapes antérieures. Une série **xis** représente les états initiaux des entrées du réseau ou états de retard. Dans un réseau NAR, l'état du réseau à un instant donné dépend des états précédents. Cette série présente ces états d'entrée pour le début de la prédiction. **ais** une série représente les états initiaux des couches cachées pour les délais initiaux du réseau .Tout comme **xis**, ces états sont nécessaires pour les neurones cachés afin de démarrer les prédictions à partir d'un certain point dans le temps permettant au réseau de continuer le calcul à partir de ces états. **ys** est la sortie prédite du réseau de neurone. Il s'agit des valeurs que le réseau prédit pour chaque pas de temps futur basé sur les entrées **xs**, ainsi que les états initiaux.

Après la prévision des résidus standardisés par un NAR en boucle ouverte on a obtenu la série **YS** ces valeurs prédites sont des estimations des résidus futurs basées sur les résidus passés. Ou autrement dit, c'est des valeurs qui représentent les résidus ajustés après avoir appliqué le modèle NAR.

La figure ci-dessous présente une comparaison entre les fluctuations des prix du pétrole réels et les fluctuations prévus par un modèle GARCH-NN :

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

FIGURE 26: Comparaison entre les valeurs réelles et les valeurs prévus.



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Il est clair que les valeurs prédites (Résidus ajustés) imitent les valeurs d'origine durant la période d'étude, cela est dû à la similitude des deux courbes. En comparant les résidus réels et prédite par NAR pour évaluer si les résidus ajustés sont proches des résidus réels en calculant les erreurs de prédiction obtenue par la différence entre les résidus - YS.

Après l'analyse de la série des erreurs de prévision on remarque que ces erreurs sont plus faibles que celles du modèle GARCH seul, cela indique que le NAR (Nonlinear Autoregressive) est capable de capturer les dynamiques ou les régularités dans les données qui pourraient ne pas être détectées par le modèle GARCH (1.1). Et pour quantifier l'amélioration apportée par le modèle NAR sur le modèle GARCH (1.1) on compare les MSE ou les RMSE des résidus GARCH et des résidus NAR ajustés. Les résultats de calculent présenter dans le tableau ci-dessous :

Tableau 16: Comparaison des MSE et RMSE des résidus réels et ajustés.

Modèles	Résidus GARCH Réels	Résidus NAR Ajustés
MSE	6.6227×10^{-7}	3.0254×10^{-8}
RMSE	8.180×10^{-4}	1.730×10^{-4}

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Selon le tableau N°16 on remarque une réduction significative des erreurs (MSE/RMSE) indique que le modèle NAR améliore les prédictions. Donc le modèle GARCH-NN est considéré comme un très bon modèle pour prédire la volatilité future des prix du pétrole. Ces résultats mettent en évidence une amélioration de la précision des prévisions par rapport aux

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

modèles précédentes grâce à l'intégration des aspects non linéaires. Cette approche hybride est particulièrement efficace pour le comportement volatil constaté dans les séries des cours du pétrole.

Conclusion

Et afin d'assurer la qualité des prédictions obtenues par l'application des trois modèles sur la série de la volatilité des prix du pétrole : GARCH_(1,1), NAR_{(1:1)(14)} GARCH-NN_{(1,1)(1:1)(10)} on utilise les métriques de performance de précision des prévisions présentés dans le tableau suivant pour comparer la performance des trois modèles :

Tableau 17: Comparaison entre les performances des trois modèles.

Modèles	MSE	RMSE
GARCH _(1,1)	6.6227 ^{e-7}	8.180 ^{e-4}
NAR _{(1:1)(14)}	9.1730 ^{e-4}	0.0303
GARCH-NN _{(1,1)(1,17,1)}	3.0254 ^{e-8}	1.730 ^{e-4}

Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Le modèle GARCH_(1,1) et GARCH-NN_{(1,1)(1:1)(10)} ont des MSE assez faibles mais le GARCH-NN présente des performances légèrement meilleures en termes de MSE et RMSE. Cela suggère que l'ajout du réseau de neurones améliore la précision de la prévision, bien qu'il soit un peu plus complexe. Le modèle NAR_{(1:1)(14)} présente un MSE et un RMSE significativement plus élevé ce qui suggère que bien qu'il puisse capturer la non-linéarité, il n'est pas aussi performant en termes de minimisation des erreurs moyennes par rapport aux modèles GARCH. Donc si l'objectif est de minimiser l'erreur et d'obtenir une prévision précise, le modèle GARCH-NN_{(1,1)(1:1)(10)} semble être le plus performant en termes de MSE et RMSE. Cependant, le modèle GARCH_(1,1) reste compétitif si la simplicité et la rapidité sont plus importantes. Le modèle NAR_{(1:1)(14)} pourrait être plus utile pour des séries non linéaires complexes, mais il ne donne pas les meilleurs résultats en termes de minimisation des erreurs.

Les résultats montrent que le modèle NAR_{(1:1)(14)} présente des erreurs plus élevées et un que le modèle GARCH_(1,1). Ce qui suggère qu'il ne parvient pas à surpasser le modèle GARCH en termes de précision des prévisions. Cependant, il est important de noter que les

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

modèles NAR peuvent exceller dans des situations où la non-linéarité est plus présente et où les relations complexes entre les variables doivent être modélisées. Mais en termes de précision mesurée par les critères MSE et RMSE, le GARCH semble plus performant.

On peut aussi conclure que le modèle GARCH-NN $(1.1) (1:1) (10)$ a montré une performance supérieure à la fois en termes des métriques par rapport aux modèles GARCH (1.1) et NAR $(1:1) (14)$ suggère que le modèle GARCH-NN pourrait effectivement remplacer l'utilisation exclusive des modèles GARCH ou NAR dans certaines situations.

Après une comparaison générale entre les trois modèles (GARCH, NAR et GARCH-NN), il est essentiel d'approfondir l'analyse en examinant les résidus issus de la modélisation du GARCH-NN. Cette étape permet d'évaluer si le modèle, bien qu'il affiche de meilleures performances globales en termes de MSE et RMSE, capture efficacement les dynamiques sous-jacentes des données, notamment dans les périodes de forte volatilité ou de comportements non linéaires. En comparant les résidus du GARCH-NN à ceux du modèle GARCH seul, nous pouvons déterminer si le GARCH-NN surpasse réellement le GARCH en termes de capacité à expliquer les variations des données, ou si les performances améliorées sont uniquement dues à des ajustements spécifiques, sans résoudre certaines limites fondamentales du modèle de base. Les modèles GARCH et GARCH-NN reposent sur des hypothèses spécifiques comme l'absence d'autocorrélation dans les résidus et la capture efficace de la volatilité conditionnelle. Si ces hypothèses ne sont pas respectées, cela indique que le modèle n'a pas pleinement capturé les dynamiques de la série temporelle. Des résidus bien modélisés indépendants, non auto corrélés et normalement distribués garantissent que les prévisions du modèle sont fiables. Si les résidus présentent encore des structures ou des patterns, cela peut signaler des lacunes dans le modèle.

Une analyse approfondit sur les résidus obtenus du modèle GARCH (1.1) est effectué dans la section de la modélisation de la volatilité des prix du pétrole par un modèle GARCH on a conclu que les résidus sont à Bruit Blanc et ne présente pas un effet d'hétéroscédasticité, le modèle a bien capturer le comportement de la volatilité conditionnelle dans les données.

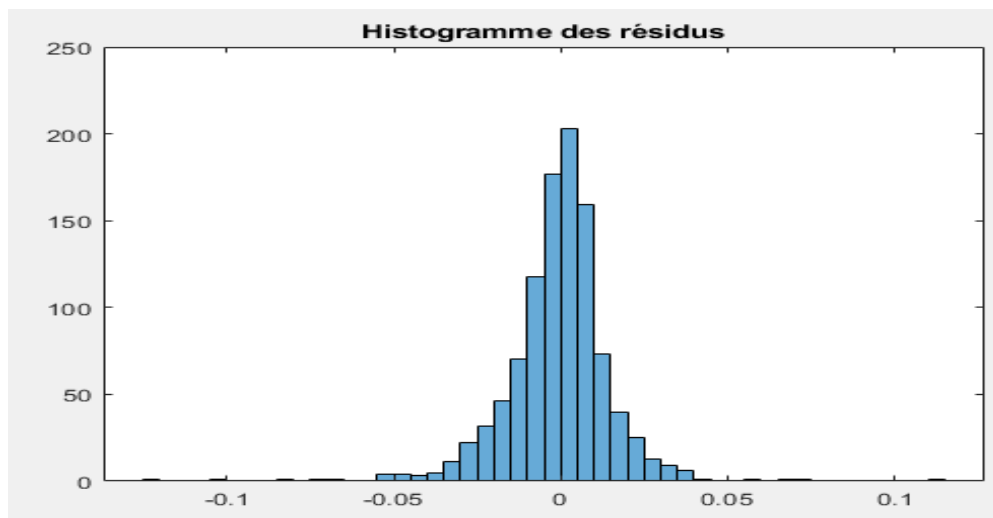
Pour un modèle GARCH-NN l'analyse des résidus montre les résultats suivants :

❖ **La distribution des résidus du modèle GARCH-NN :**

Un examen d'histogramme est nécessaire pour cette étape est vérifier la distribution des résidus

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

Histogramme 4: Résidus du modèle GARCH -NN (1.1) (1 :1) (10).



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

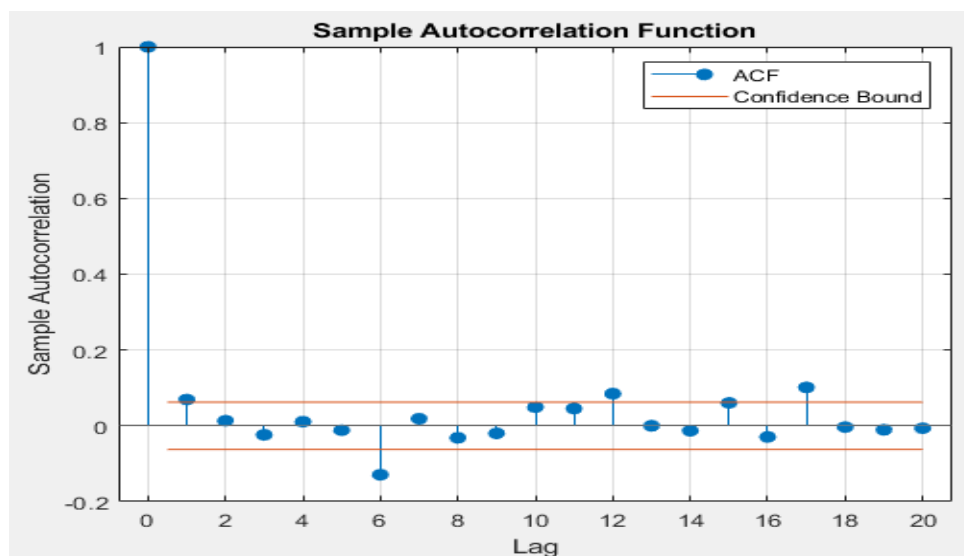
La distribution semble centrée autour de 0 ce qui est une bonne indication signifie que les erreurs de prévision ou les résidus n'ont pas de biais systématique. La forme générale de l'histogramme est symétrique aussi la concentration des erreurs autour du zéro indique que les prédictions du modèle sont globalement précises on peut remarquer que les résidus présentent des queues relativement légères ce qui pourrait indiquer qu'il n'y a pas de fortes erreurs de prévision. On peut remarquer que la distribution est légèrement inclinée vers la gauche, avec une queue gauche plus longue ou plus épaisse que la queue droite Skewness égale -0.6945 et Kurtosis = 12.6830 une valeur très élevée, ce qui suggère que la distribution présente des queues plus lourdes ou des valeurs aberrantes et plus pointue qu'une distribution normale. Ces observations visuelles sont affinées avec le test de Jarque-Berra qui donne un P-value égale à 0.001 une valeur inférieure à 0.05 ce qui suggère le rejet de l'hypothèse de normalité des résidus.

❖ Autocorrélation des résidus du modèle GARCH-NN :

On examine si les résidus sont corrélés entre eux sur différentes périodes pour cela on effectue une analyse visuelle sur le corrélogramme des résidus et on applique le test de Ljung-Box pour confirmer les commentaires retenus par le corrélogramme.

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

Corrélogramme 5: Les résidus du modèle GARCH-NN (1,1) (1 :1) (10).



Source : Réalisée par l'auteur dans le programme Matlab R2024a.

Un corrélogramme montre des autocorrélations à différents lags et l'intervalle de confiance représente la zone autour de zéro dans laquelle les autocorrélations ne sont pas significatives.

On remarque des autocorrélations significatives des erreurs du réseau des bars en dehors de l'intervalle de confiance ce qui peut indiquer que le modèle n'a pas correctement capturé toutes les dynamiques présentes dans les données des résidus. Des coefficients d'autocorrélation dépassent l'intervalle de confiance indique qu'il y a une dépendance significative à ce lag. Pour la volatilité les autocorrélations significatives dans le résidu ou la volatilité peuvent suggérer qu'il existe encore une dépendance non expliquée par le modèle.

Le test de Ljung-Box donne une probabilité inférieure à 0.05 ce qui confirme les commentaires retenus par l'examen visuelle. Donc les résidus présentent toujours une dépendance temporelle.

❖ L'homoscédasticité des résidus du modèle GARCH-NN :

Pour vérifier l'homoscédasticité des résidus un test de ARCH est effectué. Le résultat du test donne un P-value égale à 0.0021 qui est inférieure à 0.05 d'où le rejet de l'hypothèse nulle. On peut conclure que les résidus présentent toujours un effet ARCH.

Pour ces résultats on conclut que :

Section 4 : La Modélisation des prix du pétrole par le modèle GARCH_NN

- Le modèle GARCH a mieux capturé les dépendances conditionnelles dans la volatilité et a réussi à modéliser la structure présente dans les données, d'où des résidus indépendants et identiquement distribués « bruit blanc ».
- Le modèle GARCH-NN malgré sa complexité a échoué à expliquer pleinement certaines caractéristiques des données, laissant des structures non expliquées dans les résidus. Cela peut indiquer une mauvaise configuration du réseau neuronal dans le modèle hybride ce qui nécessite à réajuster le modèle pour mieux expliquer la structure des données ou une optimisation ou bien de tester différentes configurations pour s'assurer que le modèle n'introduit pas de biais non nécessaire.

Conclusion Générale

Conclusion Générale

Conclusion Générale

Diverses méthodes de prévision sont suggérées dans la littérature, ces méthodes s'avèrent inefficaces lorsqu'elles sont appliquées à des séries financières caractérisés par une forte volatilité, soit difficiles à mettre en œuvre. Cette étude vise à utiliser les techniques de l'IA pour tester leurs performances dans la prévision et de mettre en place une méthode plus efficace, qui combine deux approches traditionnelles et celle de l'intelligence artificielle, cette technique est simple à utiliser et pratique pour la prévision de la volatilité des séries financières.

L'inspiration de cette étude réside dans la conviction que le partage d'expériences et de connaissances entre différents domaines est essentiel. Les techniques de l'IA sont utilisées dans le domaine de prévision et ont réussi à produire des bons résultats mais aussi ont montrer quelques limites selon l'étude de (Nikolopoulos,2010), ce qui a conduit les experts en prévision à ne pas faire confiance dans ces techniques en proposant des modèles combine les forces de la régression et l'IA.

Les modèles hybrides ont été amplement étudiés dans la littérature comme une méthode pour améliorer la qualité des prédictions de la volatilité. En combinant les forces des approches traditionnelles et les avantages des réseaux de neurones. La nature hybride de ces modèles a montré des résultats prometteurs dans diverses applications financières.

Sur la base de l'analyse comparative entre les trois modèles (GARCH, NAR et GARCH-NN) utiliser dans la recherche les résultats ont montré que le modèle GARCH-NN $(1,1)(1:1)(10)$ a démontré une supériorité en termes de performances surpassant à la fois les modèles GARCH $(1,1)$ et NAR $(1:1)(14)$. Ces résultats suggèrent que l'intégration des réseaux de neurones au modèle GARCH offre une meilleure capacité de modélisation et de prédiction, renforçant son potentiel à remplacer l'utilisation isolée des modèles traditionnels GARCH ou NAR dans des contextes spécifiques où une précision accrue est essentielle.

Le modèle GARCH, conçu spécifiquement pour la modélisation et la capture de la volatilité dans les données financières et économiques, est reconnu pour sa robustesse et sa capacité à fournir des estimations fiables par rapport aux modèles traditionnels. Cette robustesse repose sur sa structure permettant de prendre en compte la dépendance temporelle de la volatilité, ce qui le rend particulièrement adapté à l'analyse de séries temporelles volatiles. Cependant, en combinant le modèle GARCH avec un réseau de neurones artificiels (NN), il a été possible d'améliorer significativement la précision des prédictions. Cette combinaison

Conclusion Générale

exploite la capacité des réseaux de neurones à capturer des non-linéarités complexes et des relations cachées dans les données, complétant ainsi les forces du modèle GARCH. Néanmoins, une étude comparative approfondie entre les modèles GARCH traditionnels et cette approche hybride GARCH-NN reste nécessaire pour évaluer systématiquement leurs performances dans divers contextes. Une telle étude pourrait inclure des métriques variées, des horizons temporels différents et des séries de données provenant de divers secteurs, afin d'identifier les cas où le modèle hybride apporte une réelle valeur ajoutée.

La capacité des techniques de l'intelligence Artificielle dans l'apprentissage et le traitement des données sans exigence de contraintes dépasse les capacités limitées des techniques traditionnelles. Dans ce contexte, on ne peut ni nier leurs avantages ni les éliminer, mais les utiliser comme un background et un champ d'apprentissage pour améliorer l'efficacité des techniques intelligentes et augmenter leurs performances et permettre aux experts de prévision de mettre leurs jugements pour ajuster les résultats et pour renforcer les points faibles des techniques de l'intelligence artificielle comme la qualité des données : précision, complétude, absence de biais, l'interprétabilité, et les questions d'éthique.

Après avoir consulté plusieurs études et recherches précédentes réalisées par pas mal de chercheurs et analystes et d'après les résultats de notre recherche qui confirme l'efficacité de ces techniques dans l'amélioration des prédictions de la volatilité des séries financières et que ces techniques peuvent être utilisées pour leurs résultats non seulement dans le but d'améliorer la précision mais aussi pour l'objectif de tracer des stratégies à long terme et de prendre des décisions cruciales et fiables dans le court terme en prenant compte les défis de ces techniques.

En fin de travail on peut citer quelques recommandations pour inclure les techniques de l'intelligence artificielle dans le domaine de la prévision les recommandations sont organisées en deux catégories :

❖ **Recommandation pour une poursuite de cette recherche :**

- Mener des analyses comparatives entre le modèle GARCH-NN et d'autres modèles hybrides, tels que LSTM-GARCH pour identifier les approches les plus performantes ;
- Utiliser des techniques avancées comme les algorithmes génétiques pour affiner les paramètres du modèle hybride et maximiser ses performances ;
- Analyse des limites du modèle GARCH-NN en identifiant les scénarios où le modèle GARCH-NN pourrait échouer ;

Conclusion Générale

- Intégrer des variables exogènes telles que des indicateurs macroéconomiques ou des données géopolitiques, pour améliorer la précision des prédictions.

❖ **Recommandations générales :**

- L'utilisation des modèles d'IA explicables et interprétables pour augmenter la confiance et l'adoption en modèle utilisé ;
- Développer des approches robustes capables de gérer des données bruitées, incomplètes ou biaisées pour améliorer la fiabilité des prédictions ;
- Encourager la combinaison de plusieurs modèles tels que des réseaux neuronaux profonds et des modèles économétriques, pour capturer des dynamiques complexes dans les séries temporelles.

Ces recommandations mettent en avant une combinaison d'approches robustes et interprétables, généralement regroupées sous le concept de "l'intelligence artificielle explicable" **XAI**. Ces approches privilégient particulièrement l'intégration de techniques d'apprentissage supervisé, de mécanismes renforçant la robustesse, et de méthodologies hybrides. Le concept de **XAI** est utilisé pour les prévisions de séries temporelles, il représente un domaine relativement récent et en pleine évolution. Les outils XAI comme les modèles hybrides les rendent accessibles, compréhensibles, et donc plus adaptés à des environnements exigeants.

Bibliographie

Bibliographie

- Aljinović, Josip Arnerić - Tea Poklepović - Zdravka. 2014. «GARCH based artificial neural networks in forecasting conditional variance of stock returns.» *CRORR* 5: 329-343.
- Aloui, Manel Hamdi - Chaker. 2015. «Forecasting Crude Oil Price Using Artificial Neural Networks: A Literature Survey.» *Economics Bulletin*, 35 (2): 1339-1359.
- AYACHI, Fahima CHAREF - Fethi. 2016. «A Comparison between Neural Networks and GARCH Models in Exchange Rate Forecasting.» *International Journal of Academic Research in Accounting, Finance and Management Sciences* 6 (1): 94-99.
- Bourbonnis, R. 2010. *Econométrie manuelle et exercice corrigées*. 8. Paris: Dunod.
- Camacho-Miñano, Francisco Pérez-Hernández - Alvaro Arévalo-de-Pablos- María-del-Mar. s.d. «A Hybrid Model Integrating Artificial Neural Network with multiple GARCH-type models and EWMA for Performing the Optimal Volatility Forecasting of Market Risk Factors.» *University of Madrid, Madrid (Spain)*.
- Cao, Xunfa Lu - Danfeng Que - Guangxi. 2016. «Volatility Forecast Based on the Hybrid Artificial Neural Network and GARCH-type Models.» *Procedia Computer Science* 91: 1044-149.
- CHAREFFethi, AYACHIFahima and. 2016. «A Comparison between Neural Networks and GARCH Models in Exchange Rate Forecasting.» *International Journal of Academic Research in Accounting, Finance and Management Sciences* 6: 94-99.
- Chkili, Manel Hamdi - Walid. 2019. «AN ARTIFICIAL NEURAL NETWORK AUGMENTED GARCH MODEL FOR ISLAMIC STOCK MARKET VOLATILITY: DO ASYMMETRY AND LONG MEMORY MATTER? .» *Economic Research Forum*.
- Damasevicius, Robertas. 2023. «Artificial Intelligence Techniques in Economic Analysis.» *Economic Analysis Letters* 2 (2): 52-59.
- Desrochers, Jean-François Laplante - Jean. s.d. «LE CARACTÈRE PRÉVISIONNEL DU MODÈLE GARCH (1,1) SELON DIFFÉRENTS HORIZONS DE PRÉVISIONS DE LA VOLATILITÉ.» *Annual Atlantic Schools of Business Conference*.
- Dritsakis, Dimitrios Kartsonakis Mademlis - Nikolaos. 2020. «Volatility Forecasting using Hybrid GARCH Neural Network Models: The Case of the Italian Stock Market.» *International Journal of Economics and Financial Issues* 11 (1): 49-60.
- Elias K. Karuiru, John Mwaniki Kihoro, Thomas Mageto, Anthony Gichuhi Waititu. 2022. «ANN-Time Varying GARCH Model: Simulations and Application in Modelling Temperature for Weather Derivatives.» *Open Journal of Statistics* 12: 433-441.
- Estrada, Mario Arturo Ruiz. 2023. «How Artificial Intelligence (AI) Can Help in the Economic Modelling?» *Econographication Laboratory*.

- Farouk, BENABBAS. 2012. «Méthodes Heuristiques pour la Prédiction des Séries Temporelles.» PHD Thesis, Annaba Algeria.
- Fichtner, Emil Kraft - Dogan Keles - Wolf. 2020. «Modeling of frequency containment reserve prices with econometrics and artificial intelligence.» *WILEY Journal of Forecasting*.
- Gharacheh, Mohammed Reza Amin Nasri - Ehsan Ahmadi. s.d. «A Hybrid Artificial Intelligence Approach to Monthly Forecasting of crude oil price times series.»
- Ghassane Benrhmach, khalil Namir, Abdelwahed Namir, Jamal Bouyaghroumni. 2020. «Nonlinear Autoregressive Neural Network and Extended Kalman Filters for Prediction of Financial Time Series.» *Hindawi Journal of Applied Mathematics* 2020: 6.
- Gignac, Charles. 2023. «Applications MATLAB Séries temporelles & Télédétection.» PHD Thesis, USA.
- Gracia, Juan José Montaña Moreno - Alfonso Palmer Pol - Pilar Muñoz. 2011. «Artificial neural networks applied to forecasting time series.» *Psicothema* 23 (2): 322-329.
- Haidar, Siddhivinayak Kulkarni - Imad. 2009. «Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices.» *International Journal of Computer Science and Information Security* 1 (2).
- Homayounfar, Amir Daneshvar - Maryam Ebrahimi - Fariba Salahi - Maryam Rahmaty - Mahdi. 2022. «Brent Crude Oil Price Forecast Utilizing Deep Neural Network Architectures .» *HINDAWI*.
- J. Rynkiewicz, M. Cottrell, M. Mangeas, J.F. Yao. 2001. «Modèles de réseaux de neurones pour l'analyse des séries temporelles ou la régression: Estimation, Identification, Méthode d'élagage SSM.» *SMM*.
- KIM, EUNHO KOO AND GEONWOO. 2022. «A Hybrid Prediction Model Integrating GARCH Models With a Distribution Manipulation Strategy Based on LSTM Networks for Stock Market Volatility.» *IEEE Access* 10: 12.
- Laid, GASMI. 2018/2019. «Time Series forecasting using neural networks and Genetic algorithms.» PHD Thesis, Sidi Bel Abess Algeria.
- Le, Tien-Thinh Le - Binh Thai Pham - Hai-Bang Ly - Ataollah Shirzadi - Lu Minh. 2020. «Development of 48-hour Precipitation Forecasting Model using Nonlinear Autoregressive Neural Network .» *Innovation for Sustainable Infrastructure*.
- Liu, Yi Ren - Jinglin Liu - Xu Tan - Zhou Zhao - Sheng Zhao - Tie-Yan. 2020. «A Study of Non-autoregressive Model for Sequence Generation.» *Microsoft*.
- Ly, Arthur Charpentier - Emmanuel Flachaire - Antoine. 2018. «Économétrie & Machine Learning.» *HAL* 3.
- Mahesh S. Khadka, Benjamin Popp, K. M. George, N. Park. s.d. «A NEW APPROACH FOR TIME SERIES FORECASTING BASED ON GENETIC ALGORITHM.»
- Mechgoug, Raihane. 2013. «La Prédiction des Séries Temporelles utilisant les Paradigmes de Soft Computing .» PHD Thesis, Biskra .
- MUSTAPHA, HABIBECHE. 2019/2020. «Une étude de la modélisation et comparative entre Réseau de neurones artificiels, Algorithmes génétiques et Box-Jenkins en prévision de série chronologique chaotique .» PHD Thesis, Sidi Bel Abess Algeria.

- Núñez-Velázquez, Eduardo Ramos-Pérez - Pablo J. Alonso-González - José Javier. 2020. «Forecasting volatility with a stacked model based on a hybridized Artificial Neural Network.» *Economics Department, Universidad de Alcalá, Plaza de la Victoria.*
- NOËL, Valentin. 2022. «Séries temporelles et réseaux de neurones récurrents .» *Culture Sciences de l'Ingénieur.*
- Prof. Dr.Abudl –Azeem Abudl- Khareem Ali, Prof. Dr. Fouzai Khalib Omer. 2013. «The Use Nets of Artificial Neural Networks to Predict of Macroeconomic Multidimensional model Multidimensional model.» *مجلة الاقتصادي الخليجي (24).*
- S. Lek, I. Dimopoulos, M. Derraz et Y. El Ghachtoul. 2023. «Modélisation de la relation pluie-débit à l'aide des réseaux de neurones artificiels.» *Revue des sciences de l'eau* 9 (3): 320-331.
- Saud M. Al-Fattah, Saudi Aramco. 2019. «Artificial Intelligence Approach for Modeling and Forecasting Oil-Price Volatility.» *SPE Reservoir Evaluation & Engineering.*
- So, Wing Ki Liu - Mike K. P. 2020. «A GARCH Model with Artificial Neural Networks .» *MDPI.*
- Souad, HAMDANI. 2017. «Etude Prévisionnelle de la volatilité des prix spot du pétrole.» Magister, Mostaganem Algeria.
- Spieß, Sendhil Mullainathan and Jann. 2017. «Machine Learning: An Applied Econometric Approach.» *Journal of Economic Perspectives* 31 (2): 87-106.
- Tealab, Ahmed. 2018. «Time series forecasting using artificial neural networks methodologies: A systematic review.» *Future Computing and Informatics Journal* 3 (2): 334-340.
- TIR, Rédha. 2004. «Essai de Prévision des Indices du Marché Boursier Tunisien Approche par le Réseau de Neurones.» *ResearchGate.*
- Torgler, Steve J. Bickley · Ho Fai Chan · Benno. 2022. «Artificial intelligence in the field of economics.» *Scientometrics* 127: 2055-2084.
- Truchis, Gilles de. s.d. «Économétrie des séries temporelles non-stationnaires Implémentation sous MATLAB.» Master EIPMC.
- Yildizoglu, Thomas Vallée - Murat. 2021. «Présentation des algorithmes génétiques et de leurs applications en économie.» *PESSAC* 1 (2).
- Zatout, M.Malik Kamel Bensafta - Dr. Ali. s.d. «La Volatilité du Taux de change du dinar algérien Modélisation et Prévisions.» *مجلة اقتصاديات شمال افريقيا (2).*
- zineb, Benkhedda. 2024. «The relationship between trading volume, stock returns and volatility during the Covid-19 period for the shanghai stock exchange.» *MECAS* 20 (1).
- Zohra, Bousseba Fatma. 2017. «Utilisation des modèles Arch-Garch pour modéliser le marché de l'énergie :Prix de pétrole et de gaz.» PHD Thesis, Annaba Algeria.
- «استخدام نماذج الشبكة العصبية الاصطناعية للتنبؤ بسعر الفائدة.» احمد سلطان محمد القره غولي, اسراء عامر الحمداني *Diyala Journal for pure Sciences* (2) 6.
- الكفاءة السعرية والتنبؤ بمؤشر السوق المالي باستخدام النماذج القياسية والشبكات العصبونية. 2020. الحميد, أيهم محمود. جامعة حماة: PHD thesis, Syria.

- مفردة و ANN و ARIMA دراسة احصائية مقارنة بين نماذج». s.d. الصادق, د. شادي اسماعيل التلباني - أ. محمد عبد الهادي هجينة ZCOM 42 (2): 297-321.
- Master, Adrar «. استخدام الشبكات العصبية الاصطناعية في التنبؤ بأسعار البترول». 2022. بن شريف مريم, بله وهيبة Algeria.
- طريقة مقترحة لتحليل النحدر غير الخطى المتعدد باستخدام الذكاء». 2021. حجاج, محمد عبدالسالم العجمي - عبد الوهاب السيد 52: 557-592. *المجلة العلمية للاقتصاد والتجارة* «. الصطناعي
- International* «. للتنبؤ بسعر الصرف ANN-ARIMA استخدام النموذج الهجين». 2020. حسن قهوي, عبد القادر ساهد *Multilingual Journal of Science and Technology (IMJST)* 5 (9): 1681-1685.
- «An Application of Non-Linear Autoregressive Neural Networks to Predict the Inflation Rate in Algeria During the Period 1990 to 2021». 2023. د. حسن قهوي, أ.د. عبد القادر ساهد, أ.د. محمد مكيديش 18. *مجلة وحدة البحث في تنمية الموارد البشرية*
- استخدام منجية بوكس جنكنز والشبكات العصبية الاصطناعية لمتمنبؤ باستيالك الطاقة الكيربائية». 2017/2017. سبيمة, عتروس PHD Thesis, Beskra Algeria. «-بمؤسسة سونمغاز - دراسة حالة: والية بسكرة
- مجلة «(2000-2020) للعدة A H- التنبؤ بأسعار البترول باستخدام النموذج الهجين». 2020. قهوي, عبد القادر ساهد / حسن 10: 6. *تور للدراسات اقتصادية*
- مجلة *الدراسات المالية* «. استخدام الشبكات العصبية الاصطناعية في التنبؤ بسعر البترول الخام برنت». 2019. مرابطي, د. سناء (4). *المحاسبية والدارية*
- والشبكات العصبية الاصطناعية على تحليل تجانس أخطاء السلسلة GARCH تطبيق نماذج». s.d. مصطفى, صفاء محمد على 18 (1): 397-406. *JSST* «. الزمنية
- والشبكات العصبية الاصطناعية على تحليل تجانس أخطاء السلسلة GARCH تطبيق نماذج». 2018. مصطفى, صفاء محمد على 18: 379-406. *JSST* «. الزمنية
- كلية الإدارة و الاقتصاد بغداد «. استخدام نماذج الشبكات العصبية الاصطناعية للتنبؤ والمقارنة». s.d. ناظم, م. م إيقان علاء «. استخدام الطرق الهجينة في التنبؤ لسعر الصرف للدولار الامريكي مقابل الدينار العراقي». 2013. نوال الجراح, ندى الحكاك (34). *مجلة كلية بغداد للعلوم الاقتصادية الجامعة*