

Université Mohamed El Bachir El Ibrahimi de Bordj Bou Arréridj
Faculté des Mathématiques et de l'informatique
Département de recherche opérationnelle



Mémoire

Présenté par

Houamed Fatima et Sekhara Amina

Pour l'obtention du diplôme de

Master

Filière : Mathématiques

Spécialité : Recherche Opérationnelle

Thème

**Prédiction de séries chronologiques en utilisant un modèle de Deep Learning
Etude de cas : Sonelgaz**

PRÉSIDENT	Saha Adel
ENCADREUR	Fillali Ferhat
EXAMINATEUR	Maache Salah

Promotion 2022/2023

تهدف هذه الدراسة إلى توضيح مراحل النمذجة التنبؤية باستخدام شبكتين عصبيتين ومقارنتها، حيث تم تطبيق عملي على سلسلتين زمنيتين من قيم استهلاك الطاقة لشركة سونلغاز الجزائرية برج بوعرييرج استناداً إلى أسلوب التدريب واستخدام برنامجين هما *Matlab2023*، *EVIEW S* في معالجة البيانات و تحقيق النتائج . أثبتت الخوارزميتان لنموذجي الشبكتين العصبيتين الاصطناعيتين المستخدمتين لوصف حركة نشاط استهلاك الطاقة في المؤسسة قدرتهما على تقديم تنبؤات جيدة مع القليل من الأخطاء المتوقعة، وأكدت هذه الدراسة أيضاً أهمية تحليل البيانات الإحصائية وعلاقتها بالتخطيط الاقتصادي وعمليات صنع القرار. أخيراً، نثبت القدرة التنبؤية العالية لنماذج الشبكة الاصطناعية مقارنة بالطرق الأخرى.

الكلمات المفتاحية : استهلاك الطاقة ،التنبؤ ، السلاسل الزمنية ، الشبكات العصبية
.LSTM،NARX،

Résumé

Cette étude vise à clarifier les étapes de la modélisation prédictive à l'aide de deux réseaux de neurones et à les comparer , où une application pratique a été faite sur deux séries temporelles de valeurs de consommation d'énergie pour l'entreprise algérienne Sonelgaz Bordj Bou Arreridj , basée sur la technique d'entraînement et l'utilisation de deux programmes EVIEWS et «Matlab R2023», dans le traitement des données et l'obtention des résultats.

Les deux algorithmes pour les deux modèles des deux réseaux de neurones artificiels utilisés pour décrire le mouvement de l'activité de consommation d'énergie d'entreprise ont prouvé leur capacité à fournir de bonnes prédictions avec peu d'erreurs attendues. Cette étude a également confirmé l'importance de l'analyse statistique de données et leur relation avec la planification économique et les processus décisionnels. Enfin, nous témoignons la grande capacité prédictive des modèles de réseaux artificiels par rapport aux autres méthodes.

Les mots clés : *Consommation d'énergie , Prédiction ,Série temporelle , Réseaux neurones , LSTM ,NARX .*

Abstract

This study aims to clarify the steps of predictive modeling using two neural networks and to compare them , where a practical application was made on two time series of energy consumption values for the Algerian company Sonelgaz Bordj Bou Arreridj , based on the training technique and the use of two programs EVIEWS and “Matlab R2023”, in data processing and obtaining results.

The two algorithms of the two models of the two artificial neural networks , that were relied upon to describe the movement of energy consumption activity in the organization proved their ability to provide good predictions with a small number of expected errors, and this study also emphasized the importance of statistical analysis of data and its relationship to economic planning and decision-making processes. Ultimately, we witness the higher predictive power of neural network models compared to other methods.

Keywords : *Energy consumption , Prediction , time series , Neural networks , LSTM ,NARX .*

REMERCIEMENT

Cette étude a été menée au Département de Recherche Opérationnelle du Département de Mathématiques, Faculté de Mathématiques et Informatique, Université El-Bachir el-Ibrahimi, Bordj Bou Arreridj, avec un stage à la Société Centrale Algérienne Sonelgaz, Bordj Bou Arreridj.

Nous remercions Dieu pour l'aide, la patience et le courage de faire ce travail.

Nos remerciements les plus sincères vont à Monsieur Fillali Ferhat qui nous a proposé ce sujet, ainsi que pour son aide, ses explications claires et son intérêt pour ce travail.

Aucune civilisation ne s'est élevée dans l'univers, mais elle est venue de la lumière d'un enseignant. Remerciements particuliers au Président Saha Adel, Merci cher professeur, pour votre travail acharné et votre dévouement à notre éducation, ainsi que pour tous les précieux conseils et informations que vous nous avez donnés et qui ont contribué à enrichir le sujet de notre étude. Nous adressons également nos sincères remerciements au Examineur Maach Salah pour les précieuses notes qu'il a fournies et ses efforts pour nous enseigner pendant le parcours universitaire.

Nous tenons à remercier la grande Mlle Dr. Bouthaina, notre superviseur à Sonelgaz, qui nous a aidés à obtenir les données et nous a également aidés dans le travail.

Nos remerciements vont également à l'ensemble de la division relations commerciales pour son aide, sa disponibilité et ses efforts.

Un grand merci et appréciation à la sœur, amie et chercheuse, le Dr Assai, pour ses grands efforts et son aide dans l'accomplissement de ce travail et pour ses encouragements, et ce travail a été achevé grâce à vos efforts, ma chère.

En fin, nous exprimons notre profonde reconnaissance envers tous ceux et celles qui ont participé à la réalisation de ce travail de près ou de loin.

Dédicaces

Je dédie ce travail
A mes très chers parents,
A mon cher mari,
A mon cher frère,
A mes chères sœurs,
À ma chère binôme Amina ,
A tous mes proches,
A tous ceux qui me sont chers.

Fatima

Dédicaces

Je dédie ce travail

A celle qui j'ai pas eu assez de son souffle et qui a été enterrée,
A l'âme qui serre ma joie
«ma mère Louiza Dieu repose son âme» ,

À mon cher père Boualem ,

À mes frères Djalal et Ilyas,

À ma chère sœur Siham et son mari,

À ma chère binôme Fatima ,

À tous mes proches

A tous ceux qui me sont chers.

Amina

Table des matières

Introduction Générale	8
1 Généralités sur les séries chronologiques	10
Introduction	10
1.1 Définitions	11
1.1.1 Série chronologique	11
1.1.2 Périodicité	11
1.2 Classification des séries chronologiques	11
1.2.1 Domaine d'application	11
1.2.2 Intérêt des séries temporelles	12
1.2.3 Composantes des séries chronologiques (time series components)	13
1.3 Modèles de séries chronologiques	14
1.3.1 Modèle additif	15
1.3.2 Modèle multiplicatif	15
1.3.3 Modèle mixte	16
1.4 La stationnarité de la série chronologique	16
1.4.1 Série chronologique stationnaire	16
1.4.2 Série chronologique non stationnaire	17
1.5 Comment connaître la stationnarité de la série temporelle	17
1.5.1 la méthode graphique	18
1.5.2 Fonction d'auto corrélation	18
1.5.3 Test de racine unitaire (Test DF et ADF)	19
1.6 D'une série temporelle non stationnaire en une série temporelle stationnaire	20
Conclusion	22
2 Prédiction de séries chronologiques	23
Introduction	23
2.1 Généralités sur la prédiction	24
2.1.1 L'importance de la prédiction	24
2.1.2 Durée de prédiction	24
2.1.3 Méthodes et techniques de prédiction	24
2.2 Quelques méthodes statistiques de prédiction	25
2.3 Prédiction basée sur l'intelligence artificielle	27
2.3.1 Apprentissage automatique	27
2.3.2 Processus d'apprentissage	27
2.4 Différent type d'apprentissage	27
2.4.1 Apprentissage supervisé	28
2.4.2 Apprentissage non supervisé	29
2.4.3 Apprentissage par renforcement	29
2.5 Outils d'apprentissage	30
2.5.1 Régression linéaire	30

2.5.2	Régression logistique	30
2.5.3	Les machines à vecteurs supports (SVM)	31
2.5.4	Les K plus proches voisins (KNN)	31
2.5.5	Les réseaux neurones	32
2.6	fonction d'activation	34
2.7	Types de réseaux de neurones	35
2.7.1	Peceptron	35
2.7.2	Modèle Hopfield	37
2.7.3	Modèle Boltzmann	37
2.7.4	Modèle Back Propagation	37
2.7.5	Modèle de Rétropopagation	38
2.8	Apprentissage neuronal	38
2.9	Apprentissage profond	39
2.9.1	Domaines d'applications	39
2.10	Quelques réseaux profonds	39
2.10.1	Réseaux de neurones récurrents <i>RNN</i>	39
2.10.2	Réseaux de neurones profonds <i>DNN</i>	40
2.10.3	Réseaux Neurones Convolutifs <i>CNN</i>	40
2.10.4	Le LSTM (Long Short-Term Memory)	41
	Conclusion	41
3	Acquisition et pré-traitement de données	42
	Introduction	42
3.1	Présentation de lieu de stage	43
3.1.1	Lieux	43
3.1.2	Présentation de l'entreprise	43
3.1.3	Service distribution	43
3.1.4	L'intermédiaire entre les clients et sonelgaz :	45
3.1.5	Type d'énergie	45
3.2	Objectif (Prédiction à moyen terme)	46
3.3	Présentation des différentes séries	46
3.4	Pré-traitement des deux séries	48
3.4.1	Pré-traitement de série chronologique Bordj Bou Arreridj	48
3.4.2	Pré-traitement de série chronologique Ras El Oued	49
3.5	Conversion d'un série non stationnaire en une série stationnaire	50
	Conclusion	53
4	Application	54
	Introduction	54
4.1	Prédiction de la consommation d'énergie à l'aide d'un réseau de neurones non linéaire auto-régressif	55
4.1.1	Auto régressif non linéaire	55
4.1.2	Matériel et méthodes	55
4.2	Architecture des réseaux neurones	56
4.2.1	Conception des deux réseaux	56
4.2.2	Discuter des deux réseaux	57
4.3	Test de performance du réseau neurones artificiels	58
4.3.1	Évolution des carrés moyens des erreurs à travers les étapes de formation	58
4.3.2	Corréler la sortie du réseau neuronal avec les valeurs réelles	59
4.3.3	Représentations graphiques de la sortie d'un réseau de neurones	62
4.4	Les résultats de modèle NARX	63

4.5	Prédire la consommation d'énergie à l'aide du réseau de neurones profonds <i>LSTM</i>	65
4.6	Résultats de l'algorithme LSTM pour les données de consommation d'énergie pour chacune des deux communes	66
4.6.1	Réseau 1	67
4.6.2	Réseau 2	69
4.7	Comparaison entre les deux réseaux	70
	Conclusion	71
	CONCLUSION GÉNÉRAL	72

INTRODUCTION GÉNÉRALE

La prédiction des charges électriques est nécessaire pour le système électrique à diverses fins telles que la gestion des charges, la planification de l'expansion des stations, le fonctionnement intelligent, ainsi que la précision de la détermination du prix de l'énergie électrique. L'importance de la prévision augmente avec l'utilisation croissante des énergies renouvelables [1] [2].

Prédire la consommation d'énergie, c'est obtenir des informations futures basées sur des données antérieures qui nous aident à prendre les mesures appropriées pour atteindre un équilibre entre la production et la consommation [3] [4]. En plus d'éviter les turbulences de charges à prévision basse, et les problèmes de gaspillage de production à prévision haute.

Les technologies modernes et les réseaux intelligents ont rendu la prévision des valeurs de consommation d'énergie extrêmement importante. La prévision de la consommation d'énergie est classée en trois types : prédiction à court terme, prédiction à moyen terme et prédiction à long terme.

Cette procédure de prédiction utilise de nombreuses méthodes de prédiction traditionnelles et intelligentes, notamment l'analyse de séries chronologiques, les réseaux de neurones artificiels, la logique floue et les algorithmes génétiques.

Les valeurs de consommation d'énergie dans de nombreux endroits, y compris le cas de Bordj Bou Arreridj, sont affectées par les conditions météorologiques, en particulier les températures, les vagues épidémiques, l'expansion de l'urbanisation, l'augmentation de la demande et de nombreuses autres influences. L'énergie électrique sensible aux conditions météorologiques est divisée en trois composantes.

Le premier composant n'est pas sensible aux intempéries et comprend l'éclairage et l'électroménager, le deuxième composant est sensible aux températures élevées (composant d'été) et comprend les appareils de climatisation, etc., et le troisième composant est sensible aux basses températures (composante d'hiver). et comprend les appareils de chauffage et de chauffage de l'eau... etc.

Dans cette recherche, nous avons étudié la consommation énergétique des communes de Bordj Bou Arreridj et Ras El Oued, qui sont situées dans la wilaya de Bordj Bou Arreridj. Où nous avons fait une attente des valeurs de consommation futures. Et nous nous sommes appuyés sur deux méthodes d'intelligence artificielle, à savoir le modèle autorégressif non linéaire externe *NARX*, qui est l'un des modèles de réseaux de neurones artificiels et le meilleur à utiliser pour étudier et prédire les séries chronologiques parmi d'autres réseaux.

La deuxième méthode sur laquelle nous nous sommes appuyés dans cette recherche est le réseau neuronal profond *LSTM* (Long short-term memory), qui est un type avancé de réseau neuronal récurrent. c'est l'un des réseaux les plus populaires dans le domaine de l'étude et de la prédiction des séries temporelles.

Chapitre 1

Généralités sur les séries chronologiques

INTRODUCTION

Les séries chronologiques sont essentielles à la fois dans la science et dans la vie réelle. L'analyse de séries temporelles est un enjeu scientifique important et trouve des applications dans divers domaines tels que la finance, la production d'électricité, la climatologie, etc... Ce chapitre commencera par les séries chronologiques, leurs modèles, leur analyse ainsi que les objectifs de leur analyse.

Ensuite, il traite de la stationnarité de la série temporelle et de la manière de connaître la stationnarité de la série temporelle. Enfin, la méthode de conversion des séries temporelles non stationnaires en séries temporelles stationnaires.

1.1 Définitions

1.1.1 Série chronologique

On appelle série chronologique (chronique ou temporelles), $\{x_t, t=1 \dots T\}$ tout suite fini des observations numérique d'un phénomène au cours du plusieurs années à intervalle de temps régulier et successives appelés saisons. Et on dit d'une série temporelle qu'elle est continue (*ex : la température, tableau de fréquence cardiaque, dessin de cerveau...*) si ses composantes sont apparu de manière continue dans le temps, et discrète (*ex : le revenu national annuel, la quantité mensuelle de précipitations...*) dans le cas contraire. [5]

1.1.2 Périodicité

La périodicité d'une série chronologique est la durée qui sépare deux observations. La durée des observations peut être annuelles, trimestrielles, mensuelles, hebdomadaire, journalière.. [6]

1.2 Classification des séries chronologiques

On peut classifier les séries chronologiques selon des critères variés : domaines d'application, séries réelles, complexe, séries stationnaires ou non stationnaires.

1.2.1 Domaine d'application

Nous trouvons des séries chronologiques dans notre environnement sous de nombreuses formes dont les suivants :

— Séries chronologiques météorologiques

Les séries chronologiques sont utilisées pour prédire les changements météorologiques et climatiques, et la méthode de Box-Jenkins est l'une des méthodes les plus largement utilisées à cet égard.

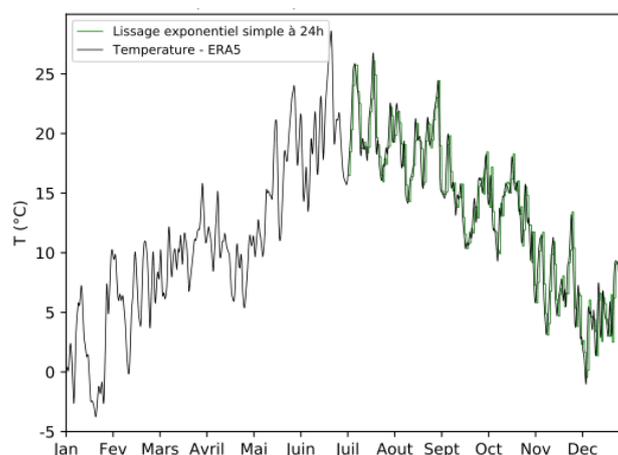


FIGURE 1.1: Temoératures journalières à Paris - 2017 [7]

— séries chronologiques économiques

(Economic Time Series) :par exemple cours des métières premières et des actions,valeurs

des importations et des exportations, production et consommation industrielle, dépôts auprès des banques.....

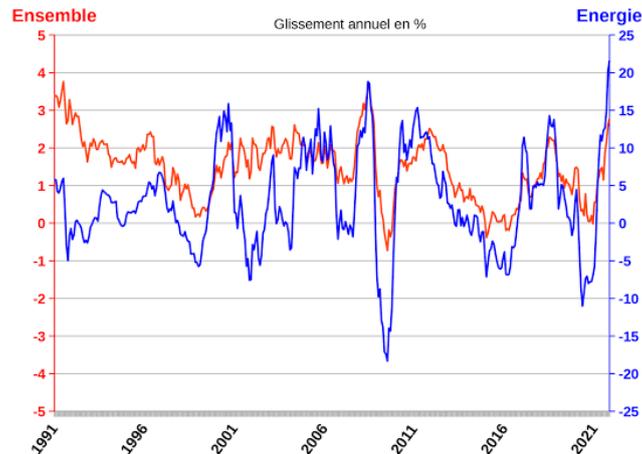


FIGURE 1.2: Indice des prix à la consommation : énergie et ensemble des produits [8]

— séries chronologiques de la Population

Données démographiques sur des années successives pour une zone il comprend également la répartition de la population entre les hommes et les femmes, et la structure par âge, ainsi que la concentration de la population dans les zones géographiques....

— médecine, biologie

suivi des évolutions des pathologies, analyse d'électron-encéphalogrammes et électrocardiogrammes

— Série chronologiques exprimant un événement

(Point Process) : cette série est souvent représentée sous la forme d'une ligne droite, par exemple, les accidents de la route ou les collisions ferroviaires...ect

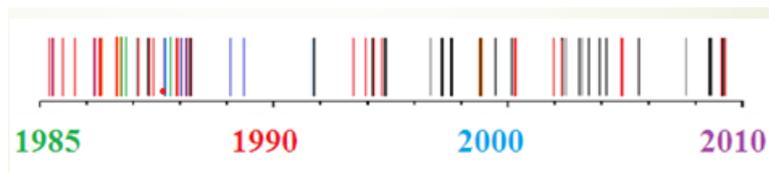


FIGURE 1.3: Échantillons aléatoires d'accidents de la route dans la ville pendant (1985/2010)

1.2.2 Intérêt des séries temporelles

L'intérêt pour les séries chronologiques est généralement considéré sous trois angles : descriptif, explicatif et prédictif.

1. Description :

- L'analyse temporelle permet de connaître la structure de la série de données étudiée.
- Elle peut être utilisée pour comparer une série à d'autres séries.

2. Explication :

- Les variations d'une série peuvent être expliquées par une autre série (exposition météorologique, pollution atmosphérique, etc.).
- il est possible de modéliser une intervention externe grâce à l'analyse de séries temporelles.
- Ces analyses permettent de réaliser des scénarios pour la période contemporaine : en agissant sur une variable explicative, il est possible d'observer le comportement de la variable expliquée.

3. Prédiction :

- La prédiction a priori permet la planification .
- La prédiction a posteriori permet d'estimer l'impact d'une perturbation sur la variable expliquée
- Des scénarios pour le futur, enfin, peuvent être réalisés.

1.2.3 Composantes des séries chronologiques (time series components)

Nous avons dit précédemment que l'un des objectifs de l'étude des séries temporelles est de décrire le phénomène étudié et d'identifier les changements qui s'y produisent au cours de temps en raison des différents facteurs auxquels le phénomène est exposé. En fait, ces changements se produisent en raison de quatre types différents de facteurs (influences), à savoir :

— **tendance (trend)**

La tendance (ou trend) y_t représente l'évolution à long terme de la série étudiée et traduit le comportement "moyen" de la série. C'est une fonction monotone qui peut contenir par exemple des composantes exponentielles.

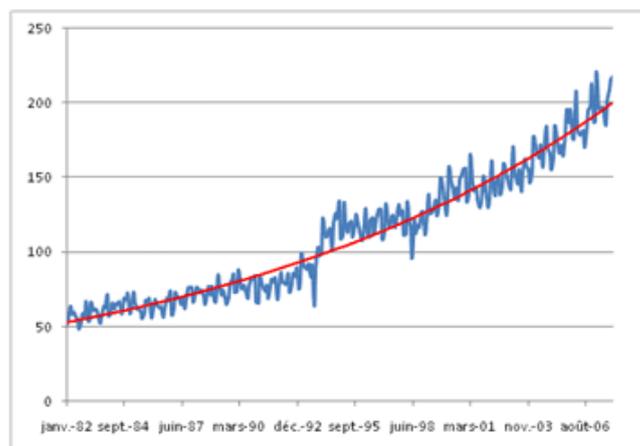


FIGURE 1.4: Tendence série chronologique [9]

— **Composante cyclique**

Ce sont les changements qui se produisent dans les valeurs de la série chronologique de manière régulière ou irrégulière sur une période de plus d'un an.

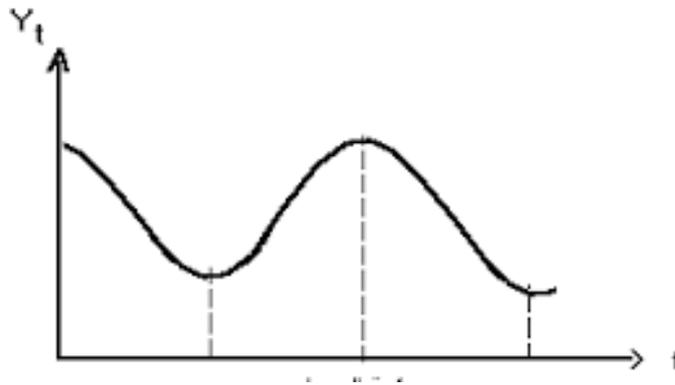


FIGURE 1.5: Série temporelle cyclique

— saisonnière

Ils expriment les fluctuations et les changements qui se produisent sur des périodes de temps similaires. Nous appelons chaque période une saison. Ces fluctuations cycliques se produisent dans les saisons de vacances, fêtes...etc. C'est-à-dire qu'il s'établit autour d'une certaine valeur d'une année sur l'autre, par exemple : les déplacements moyens le premier janvier de chaque année et pendant la saison du Hajj...etc Les changements saisonniers

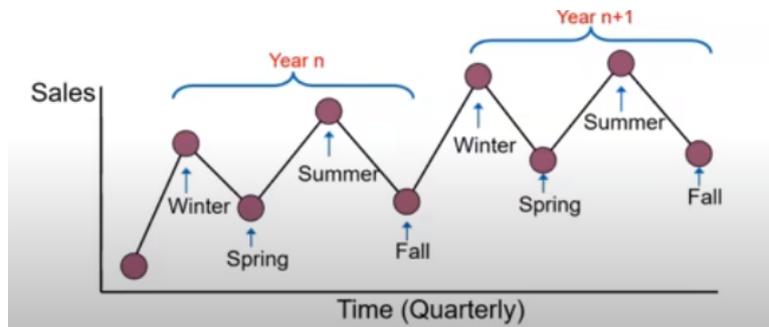


FIGURE 1.6: Série temporelle saisonnière

peuvent être déterminés soit en traçant la série chronologique si le modèle saisonnier est évident, soit par la fonction d'autocorrélation ACF , de sorte que plus la zone de fluctuations saisonnières est grande, plus les oscillations répétées dans le diagramme de la fonction ACF sont prononcées.

— Composante irrégulière

Bruit (ϵ_t) : Le bruit statistique correspond à des fluctuations de nature irrégulières, aléatoires et inexplicables.

Cette décomposition peut être additive $y_t = T_t + S_t + C_t + \epsilon_t$ ou multiplicative $Y_t = T_t * S_t * C_t * \epsilon_t$. Il est également possible de combiner ces deux décompositions $y_t = T_t * S_t + C_t + \epsilon_t$

1.3 Modèles de séries chronologiques

L'observation des séries chronologiques permet de distinguer trois types de séries :

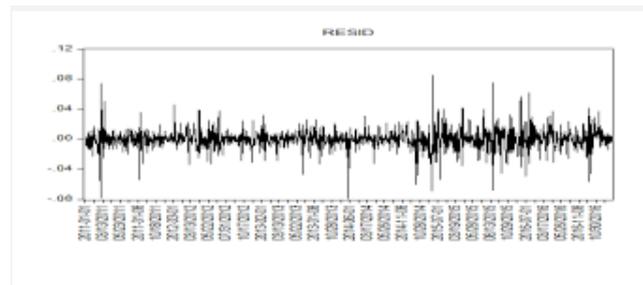


FIGURE 1.7: Série temporelle irrégulière

1.3.1 Modèle additif

Nous considérons dans cette section une série $X = (x_t)_t \in Z$ admettant une décomposition additive

$$X_t = Z_t + S_t + Q_t, \{t = 1 \dots T\} \quad (1.1)$$

où z_t est la composante tendancielle, s_t la composante saisonnière et Q_t représente l'erreur ou l'écart au modèle.

- la tendance z_t exprime un mouvement à moyen terme de la série. Elle est le plus souvent modélisée par une fonction polynomiale du temps.
- la composante saisonnière exprime un phénomène qui se reproduit de manière analogue sur chaque intervalle de temps successif.
- les erreurs sont des variables aléatoires centrées.

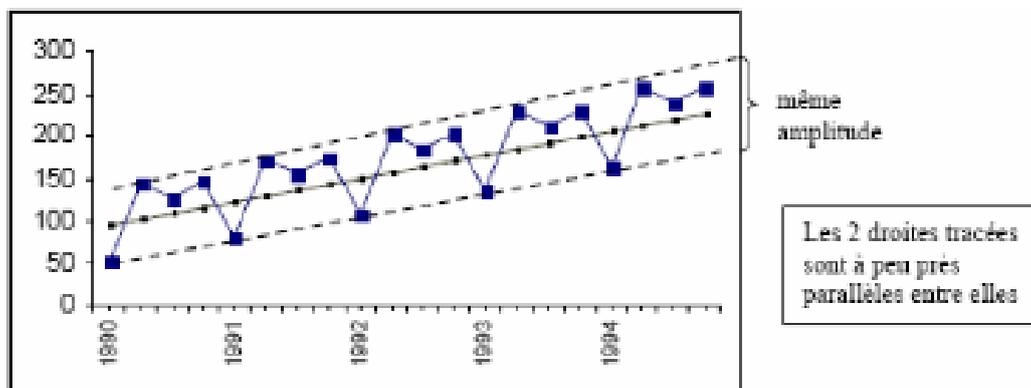


FIGURE 1.8: Modèle additif [10]

1.3.2 Modèle multiplicatif

Nous considérons dans cette section une série $X = (X_t)_t \in Z$ admettant une décomposition multiplicative :

$$X_t = Z_t(1 + S_t)(1 + Q_t), \{t = 1 \dots T\} \quad (1.2)$$

où z_t est la composante tendancielle, s_t la composante saisonnière et Q_t représente l'erreur ou l'écart au modèle.

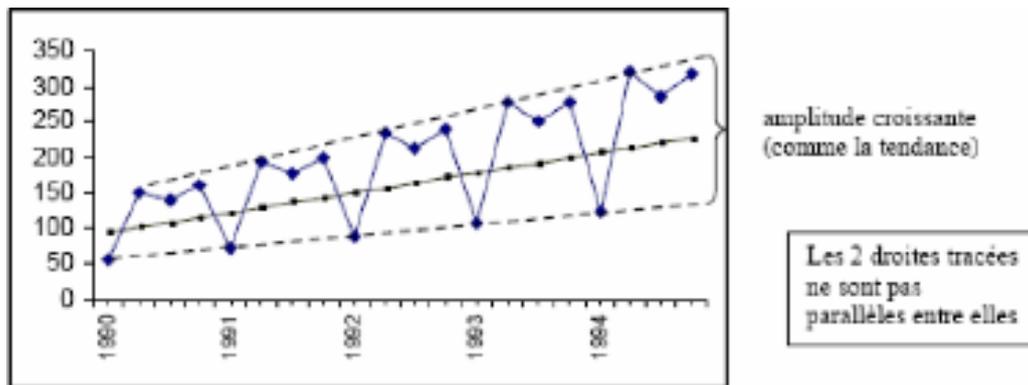


FIGURE 1.9: Modèle multiplicatif [10]

1.3.3 Modèle mixte

Il s'agit là de modèles où addition et multiplication sont utilisées. On peut supposer par exemple que la composante saisonnière agit de façon multiplicative alors que les fluctuations irrégulières sont additives :

$$X_t = Z_t S_t + \varrho_t, \{t = 1 \dots T\} \tag{1.3}$$

1.4 La stationnarité de la série chronologique

La stationnarité d'une série chronologique est une propriété importante utilisée dans l'analyse des séries temporelles. Une série chronologique est dite stationnaire si les propriétés statistiques de la série, telles que la moyenne, la variance et la covariance, sont constantes dans le temps.

1.4.1 Série chronologique stationnaire

On dit qu'une série temporelle y est stable si sa moyenne et sa variance sont constantes dans le temps et que la variance entre deux valeurs de la série ne dépend que de la durée séparant les deux valeurs, et non des instants réels où les variables sont observées [11]

Nous traduisons mathématiquement ce que nous avons dit comme suit :

- Moyenne arithmétique constante dans le temps

$$E(y_t) = E(y_{t+1}) = \alpha, \forall t \tag{1.4}$$

- variance constant dans le temps

$$V(y_t) = V(y_{t+k}) = \gamma^2 \tag{1.5}$$

- Aussi, la valeur de la covariance entre deux périodes de temps ne dépend que du décalage temporel entre les deux périodes et non du temps réel :

$$Cov(y_t, y_{t+k}) = [(y_t - \alpha) - (y_{t+k} - \alpha)] = cov(y_{t+k}, y_{t+k+s}) \tag{1.6}$$

Ou si les valeurs de l'observation pour la série chronologique y_t sont les suivant $y_1, y_2, y_3 \dots y_n$ et γ_y^2, \bar{y} sont des estimations de γ^2 et α respectivement, alors :

$$\bar{y} = \frac{1}{N} \sum_{t=1}^n y_t \tag{1.7}$$

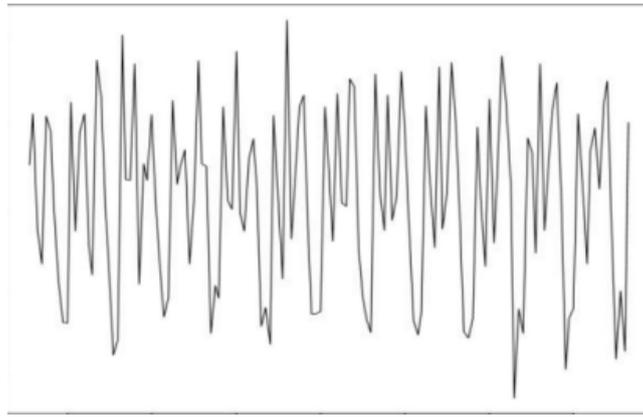


FIGURE 1.10: Série chronologique stationnaire [12]

$$\gamma_y^2 = \frac{1}{N-1} \sum_{t=1}^N (y_t - \bar{y})^2 \quad (1.8)$$

$$\text{Cov}(y_t, y_{t+k}) = \frac{1}{N-1} \sum_{t=1}^N (y_t - \alpha)(y_{t+k} - \alpha) \quad (1.9)$$

1.4.2 Série chronologique non stationnaire

Une série chronologique non stationnaire est l'opposé d'une série stationnaire, en ce sens que la moyenne des séries non stationnaire change continuellement avec le temps, soit par augmentation, soit par diminution, et qu'elle est également affectée par des variables saisonnières.

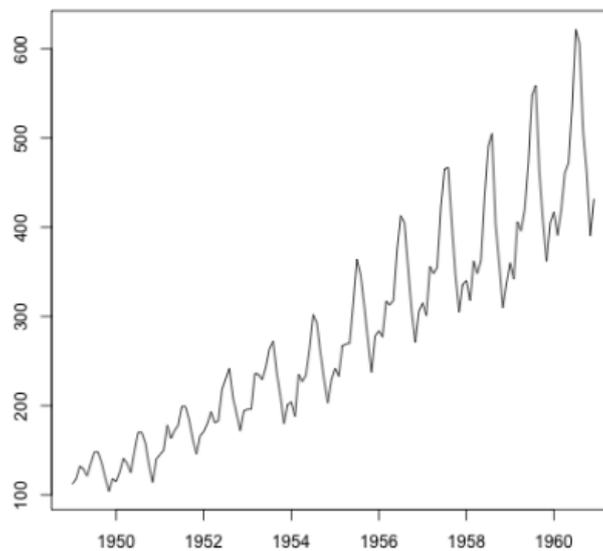


FIGURE 1.11: Série chronologique non stationnaire

1.5 Comment connaître la stationnarité de la série temporelle

La plupart des séries chronologiques économiques sont des séries instables (non statiques) car elles sont affectées d'une manière ou d'une autre par les variables qui les entourent, telles que les conditions économiques et les conditions politiques, ce qui entraîne une fluctuation de la série

(variation moyenne et instable).

Il est bien connu, et c'est une règle très importante, que les données non statiques ne peuvent pas être prédites (régression) car si cela est fait, les résultats obtenus seront faux (fausse régression) *Donc, savoir si la série est stable ou non est très important, et pour le savoir, nous avons effectué les méthodes suivantes :*

1.5.1 la méthode graphique

C'est la méthode la plus simple, et cela dépend du dessin de la série chronologique, puis de la détermination de sa stabilité ou non, et on peut s'appuyer sur Eviews pour cela .

1.5.2 Fonction d'auto corrélation

La fonction d'auto-corrélation (**ACF**) est une mesure de la force de la corrélation entre les données qui composent la série. Il s'agit de deux formes de la fonction d'auto corrélation d'échantillon (**SACF**) et la seconde est la fonction d'auto-corrélation partielle (**PACF**) [13]. Elle est définie mathématiquement comme suit :

$$R_k = \frac{\sum_{t=b}^{n-k} (x_t - \bar{x}) * (x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad (1.10)$$

avec R_t est la valeur de corrélation $k=1,2,3,..$

$$\bar{x} = \frac{\sum_{t=1}^n x_t}{(n - b - 1)} \quad (1.11)$$

avec \bar{x} est un moyenne arithmétique et n c'est la taille de la série ou nombre de vues,et b est le début de la prise des différences par exemple $y_2 - y_1$,donc $b=2$, et chaque r_k a un erreur type avec la relation suivante :

$$S_{R_k} = \frac{\sqrt{1 + 2 \sum_{j=1}^{k-1} R_j^2}}{\sqrt{n - b + 1}}, K = 0, 1, 2... \quad (1.12)$$

comme pour la statistique de test T_R ,est le quotient de :

$$t_{R_k} = \frac{R_k}{S_{R_k}} \quad (1.13)$$

REMARQUE :

Ces calculs ne sont pas effectués manuellement, mais sont effectués à l'aide d'un ordinateur, par exemple à l'aide du programme **Eviews**.

d'auto-corrélation partielle PAC

la fonction d'auto-corrélation partielle au décalage k est définie mathématiquement comme suit :

$$R_{kk} = \frac{R_k - \sum_{i=1}^{k-1} R_{k-1,i} R_{k-i}}{1 - \sum_{i=1}^{k-1} R_{k-1,i} R_i}, k = 2, 3, ... \text{ et si } k = 1 \text{ alors } R_{kk} = R_1 \quad (1.14)$$

et

$$R_{ki} = R_{k-i,i} - R_{kk} R_{k-1,k-1}, i = 1, 2, ..., k, k - 1 \quad (1.15)$$

son erreur standard est

$$S_{R_{kk}} = \frac{1}{\sqrt{n-b+1}} \quad (1.16)$$

et la statistique de test est

$$T_{R_{kk}} = \frac{R_{kk}}{S_{R_{kk}}} \quad (1.17)$$

1.5.3 Test de racine unitaire (Test DF et ADF)

Le test de racine unique est l'un des tests les plus importants dans l'étude des données de séries chronologiques. Une recherche qui nous permet de connaître la stabilité ou l'instabilité de la série, dans un état instable, elle est convertie en prenant les premières différences des données de les variables d'origine dans une série stable, ou d'autres méthodes de conversion peuvent être utilisées.

Si la série temporelle n'est pas stationnaire, mais qu'elle devient stationnaire à ses différences premières, on dit que c'est une série intégrale originale d'ordre $L \rightsquigarrow (1)$, on dit en général. Une série originale intégrale d'ordre d , où nous la notons $L \rightsquigarrow (d)$, et (d) est le nombre de fois que les différences de série sont prises.

Les méthodes de test unitaire les plus utilisées sont le test de Dickey-Fuller (DF) et le test de Dickey-Fuller étendu (ADF) [14].

Principe :

Les deux dépendent de l'hypothèse nulle, et nous disons que la série contient une racine unitaire (non stationnaire), et l'hypothèse alternative est que la série ne contient pas de racine unitaire (stationnaire) c'est comme suit :

$$H_0 = \rho = 0, \quad \text{non stationnaire}$$

$$H_1 = \rho > 0, \quad \text{stationnaire}$$

De l'expression processus aléatoire

$$y_t = \alpha y_{t-1} + \varepsilon_t, \quad \varepsilon_t \text{ est l'erreur aléatoire}$$

On ajoute à l'équation (y_{t-1}) Nous trouvons :

$$\Delta y_t = (\alpha - 1)y_{t-1} + \varepsilon_t$$

En substituant $(\alpha - 1)$ à (ρ) on obtient : $\Delta y_t = \rho y_{t-1} + \varepsilon_t$ (1.18) donc si $\rho = 0$ (série non stationnaire) alors, $\alpha = 1$ Cela signifie qu'il existe une racine unitaire.

Maintenant, pour tester la racine unitaire, on utilise **Dickey-Fuller**, qui détecte la stabilité de la série dans les trois cas suivants :

Il n'y a pas de déviation, égalisation sans constante et sans direction

$$\Delta y_t = \rho_1 y_{t-1} + \varepsilon_T \quad (1.19)$$

En présence de déviation de toute équation avec une constante sans direction

$$\Delta y_t = \rho_0 + \rho_1 y_{t-1} + \varepsilon_t \quad (1.20)$$

Une tendance autant qu'une constante

$$\Delta y_t = \rho_0 + \rho_1 y_{t-1} + \rho_2 t + \varepsilon_t \quad (1.21)$$

dans lequel ρ_0 est la déviation et ρ_1 est pente de l'équation et ρ_2 est la tendance avec le temps.

Dans le cas d'une série chronologique non stationnaire, l'erreur aléatoire ε_t n'est pas distribuée normalement et par conséquent, le test t ne peut pas être utilisé pour déterminer la signification statistique de ρ_1 . Par conséquent, **Dickey** et **Fuller** ont proposé des tableaux appelés statistiques **tau**.

Des critiques ont été faites au test de Dickey-Fuller, et l'une des principales critiques est qu'il s'agit d'un test qui contient une auto-corrélation dans les restes de régression. Par conséquent, ces deux scientifiques ont développé un meilleur test appelé **Dicky-Fuller étendu (ADF)**, et aujourd'hui c'est l'un des meilleurs tests, car la fonction du test comprend le comptage des différences de la variable dépendante, en bref, la conversion du modèle $AR(1)$ en un modèle $AR(p)$ de sorte que l'équation de la racine unitaire devienne la suivante :

$$\Delta y_t = \rho_1 + y_{t-1} + \rho_2 t + \sum_{i=1}^m a_i \Delta y_{t-1} + \varepsilon_t \quad , \quad a_i \text{ variable dépendante} \quad (1.22)$$

1.6 D'une série temporelle non stationnaire en une série temporelle stationnaire

La plupart des séries chronologiques de la vie économique sont des séries instabilité en raison de l'instabilité des conditions environnantes .
son instabilité due à l'une des raisons suivantes :

- Il y a une tendance général.
- La présentation de fluctuations saisonnières.
- Instabilité de la variance et moyenne arithmétique.

Notre objectif principal dans ce travail est la prédiction, et comme nous savons que la prédiction ne peut pas être faite sur une série instable, car cela présente des résultats très éloignés et erronés. Par conséquent, j'ai dû tester la stabilité de la série avant de mener l'étude, et dans le cas d'une série instable, nous avons recours à l'application de méthodes qui rendent cette série stable. Nous mentionnons les plus importants d'entre eux comme suit :

1. En cas d'instabilité de variance.

Dans ce cas, la meilleure façon d'installer la variance de la série est d'obtenir le *logarithme* naturel des données de la série, ou d'en obtenir la racine carrée, ou l'inverse des données, ou d'appliquer les transformations bien connues de Box et Cox dans le domaine de la conception et de l'analyse d'expériences. Cependant, la transformation des logarithmes est la meilleure et la première option dans le cas où il n'y a pas de silence pour la moyenne et la variance ensemble.

De nombreux phénomènes, notamment économiques et démographiques, ont une valeur

au temps T supérieure à leur valeur au temps $t-1$ d'un certain pourcentage, plus les erreurs aléatoires. On exprime la série comme suit :

$$y_t = y_{t-1} + \alpha y_{t-1}, \quad 0 < \alpha < 1 \quad (1.23)$$

Ce type de série se caractérise par la présence d'une tendance à la hausse à la fois de la moyenne et de la variance, en plus d'une stabilité approximative du taux de croissance du phénomène. Pour utiliser la transformation logarithmique, nous réécrivons le modèle sous la forme suivante :

$$y_t = (1 + \alpha)y_{t-1} \quad (1.24)$$

Donc

$$\ln(y_t) = \ln(1 + \alpha) + \ln(y_{t-1}) \quad (1.25)$$

Et cela conduit à :

$$\ln(y_t) - \ln(y_{t-1}) = \delta \quad (1.26)$$

Où δ est un montant fixe c'est-à-dire :

$$X_t = \Delta \ln(y_t) = \delta \quad (1.27)$$

Cela prouve que les différences premières des X_t logarithmes sont un processus statique.

Note importante 1 Vous devez vous assurer que toutes les valeurs de la série y_t sont strictement positives avant d'utiliser cette transformation .

Note importante 2 La série des différences premières des logarithmes peut parfois rester instable, et ici les différences secondes des logarithmes doivent être prises pour stabiliser la série.

2. En cas de tendance général .

Il y a deux Méthode :

(1) *Méthode de Régression linéaire :*

Cette méthode estime la tendance, puise l'isole et traite le résidu comme une série chronologique stationnaire .

(2) *Méthode de la différence :*

Cette méthode consiste à soustraire les valeurs des observations les unes des autres pour certaines périodes lentes. Par exemple, les différences du premier degré prennent la forme suivante :

$$X_t = \nabla y_t = y_t - y_{t-1}$$

Et les différences du second degré sont écrites sur la figure :

$$\begin{aligned} X_t &= \nabla^2 y_t = \nabla y_t - \nabla y_{t-1} \\ &= y_t - 2y_{t-1} + y_{t-2} = (1 - \beta)^2 y_t \end{aligned}$$

Note importante 3 : La série des première et deuxième différences peut ne pas être stationnaires, et c'est ce qui pousse le chercheur à recourir à l'application D (plusieurs fois) des différences pour se débarrasser de la tendance générale et obtenir une série chronologique stationnaires.

CONCLUSION

Une série chronologique représente des données historiques mesurées de manière ordonnée. Ces dernières évoluent dans le temps, impliquant leur caractère dynamique, et pour les modéliser, ce chapitre cite plusieurs méthodes selon l'un des trois objectifs de l'analyse des séries temporelles. Parmi ces objectifs, nous avons choisi la prévision de séries temporelles. Étant donné que la prévision de séries chronologiques est choisie pour ce chapitre, la méthode de prévision la plus utilisée est les réseaux de neurones artificiels.

Chapitre 2

Prédiction de séries chronologiques

INTRUDUCTION

La prédiction de séries chronologiques est au cœur de nos recherches utilisant les réseaux de neurones, la technologie d'IA la plus puissante, imitant les réseaux de neurones biologiques trouvés dans le cerveau humain, apprenant par l'exemple et s'entraînant d'une manière similaire à l'apprentissage humain, les réseaux de neurones sont préparés et applications concrètes tels que les modèles discriminatifs, la cognition, la classification des données, etc. sont organisés à travers le processus d'apprentissage.

Dans ce chapitre, nous verrons l'importance de la prévision et de ses méthodes. Enfin, les prédictions basées sur l'IA.

2.1 Généralités sur la prédiction

Définition : Il s'agit de planifier et d'établir des hypothèses pour des événements futurs à l'aide de techniques et de méthodes spéciales sur différentes périodes, car il s'appuie sur l'utilisation de données antérieures sur le phénomène, qui servent d'entrées pour faire des estimations prédictives qui déterminent les tendances futures. Les décideurs s'y fient pour développer et améliorer leurs hypothèses sur les conditions futures.

Pour les institutions et les entreprises, la prédiction est la fenêtre vers l'avenir, en particulier pour les activités économiques qui dépendent de l'allocation des ressources disponibles à diverses activités. Par conséquent, l'objectif principal de la prédiction est de tirer le meilleur parti de l'information actuellement disponible.

Et l'investir dans des activités futures qui servent les objectifs particuliers de l'institution. En général, une prédiction parfaite de l'avenir n'est pas possible, mais les études nous donnent des tendances générales.

Les méthodes de prédiction diffèrent selon le phénomène étudié et la méthode sur laquelle le chercheur veut s'appuyer, et parmi elles se trouvent des méthodes statistiques qualitatives et quantitatives basées sur des modèles mixtes et des modèles aléatoires, qui sont largement acceptées par les chercheurs, les chercheurs et les décideurs, dont les plus célèbres sont Box-jenkins, et d'autres méthodes qui s'appuient sur l'intelligence artificielle, considérée comme nouvelle à l'époque, comme les réseaux de neurones [15]etc

2.1.1 L'importance de la prédiction

- Il aide l'administration à prendre des décisions correctes dans une large mesure, ainsi que l'autorité responsable à prendre des décisions et à s'assurer que ces décisions sont mises en œuvre.
- Il travaille sur l'estimation des besoins financiers lorsque l'utilisation efficace du capital est une question sensible.
- La prédiction offre un travail fluide et continu à toute organisation, en particulier une organisation nouvellement créée. Grâce à la prédiction, les bénéfices et les pertes attendus peuvent être estimés.
- La prédiction aide à réussir au travail car une prédiction précise des ventes stimule l'achat des matières premières nécessaires à travers lesquelles de nombreuses activités commerciales sont menées, et en fait le processus de prédiction précise est essentiel pour de nombreux budgets [16].

2.1.2 Durée de prédiction

C'est peut-être une prédiction :

1. Court terme , moins de deux ans .
2. Moyen terme de deux à cinq ans .
3. Long terme , plus de cinq ans .

2.1.3 Méthodes et techniques de prédiction

Il existe de nombreuses méthodes et types que les analystes financiers utilisent pour prévoir les revenus, les dépenses et les coûts en capital futurs des entreprises. Il existe également un large

éventail d'outils de prédiction budgétaire quantitatifs qui sont utilisés fréquemment ou pour prédire tout autre phénomène comme les changements météorologiques ou les événements naturels etc. Parmi les phénomènes, nous mentionnons à partir de ces méthodes les suivants :

- Tendances linéaire .
- Régression linéaire simple et multiple .
- Lissage exponentiel simple.
- Moyennes mobiles *MA*.
- Modèles *AR* .
- Modèles *ARDL*.
- Modèles *ARIMA*.
- Modèles *ARICH*.
- Modèles *SARIMA*.

Celles-ci font partie des méthodes statistiques, et en plus de ces méthodes, il existe d'autres techniques concurrentes et récemment diffusées dans le domaine de la prévision qui sont connues pour leur efficacité et leurs bons résultats en matière de prévision. Ce sont des technologies basées sur ce qu'on appelle **l'intelligence artificielle**.

2.2 Quelques méthodes statistiques de prédiction

Le développement technologique moderne dans tous les domaines de notre vie contemporaine d'une part, et l'entrée du monde dans l'ère informatique, d'autre part, tout cela a conduit à l'importance croissante de l'utilisation de méthodes d'analyse statistique dans tous les domaines de la connaissance et à tous les niveaux.

Au niveau de l'économie nationale, ou au niveau des unités économiques, qu'il s'agisse du secteur public ou du secteur privé, la nécessité de collecter des données et d'en extraire des informations sur la base d'une étude systématique prenant en compte la modernité est l'une des questions fondamentales dans notre monde moderne. C'est ce que font les méthodes d'analyse statistique. Ces méthodes sont utilisées dans la planification, l'élaboration des politiques, la prise de décision et la prédiction .nous mentionnons quelques méthodes importants dans ce qui suit :

1. Tendances linéaire générale [17] :

Son équation

$$Y = \alpha + \beta X \quad (2.1)$$

ou α est la constante de l'équation et β est le pente de l'équation et X représente les valeurs de la variable indépendante (Temps) et Y représente le phénomène prédit, la tendance générale a quatre états selon la nature de la série qui sont :

tendance linéaire

$$H_t(k) = \alpha' + \beta'(t + k) \quad (2.2)$$

Quadratique linéaire

$$H_t(k) = \alpha' + \beta'(t + k) + \gamma'(t + k)^2 \quad (2.3)$$

tendance exponentiel

$$H_t(k) = \exp(\alpha' + \beta'(t + k)) \quad (2.4)$$

2. Moyennes mobiles *MA(q)* [17]

Cette méthode est utilisée avec des données saisonnières ou périodiques et peu d'observations, elle donne plus d'importance aux observations récentes dans la prédiction, comme suit :

$$Y_{t+1} = \frac{x_t + x_{t-1} + x_{t-2} + \dots + x_{t-n+1}}{n} \quad (2.5)$$

Ou Y_t est la valeur prédite pour la période suivant , et x_t est la valeur réelle pour la période t et n est le nombre de périodes du média en mouvement.

3. Modèle Auto-régressif $AR(p)$ [17]

Nous notons le modèle auto-régressif de p-degré par $AR(p)$ ou il est noté $ARIMA(p,0,0)$. Cette méthodologie est largement utilisée dans la prédiction des variables macroéconomiques , mathématiquement c'est :

$$Y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + s_t \tag{2.6}$$

ou δ est la limite absolue de la série et $\phi_i, i = 1, 2..p$ poids accompagnent les valeurs de la série y_t , et p est le score du modèle , s_t sont des erreurs aléatoires avec une distribution normale.

4. Modèles $ARIMA(p, d, q)$ et $ARMA(p, q)$ [17]

Ils sont symbolisés par $ARMA(p, q)$ ou $ARIMA(p, d, q)$, et les deux sont des modèles mixtes d'une même famille. Le premier a une sorte de coordination entre $AR(p)$ et $MA(q)$, et il est caractérisé comme un processus stationnaire, son milieu est constant dans le temps, sa formule mathématique :

$$X_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \delta - \Theta_1 \xi_{t-1} - \Theta_2 \xi_{t-2} - \dots - \Theta_q \xi_{t-q} \tag{2.7}$$

Quant aux modèles $ARIMA(p, d, q)$, sont dus au fait que la série temporelle est non stationnaire , donc on dit qu'ils sont intégrés, c'est-à-dire qu'on applique la méthode des différences du premier ordre pour obtenir un série stationnaire.

5. Box-jenkins

Il s'agit d'un modèle de régression automatique et de moyenne mobile. On lui a donné ce nom en lien avec les deux statisticiens qui l'ont développé. Il est l'un des modèles statistiques les plus connus dans le domaine de la prévision, notamment après son célèbre livre (Time Series Analysis Forecasting and Control) qui a été publié en 1976.[17] Il est devenu le plus largement utilisé dans l'analyse moderne des séries chronologiques et est basé sur les processus illustrés dans le diagramme suivant :

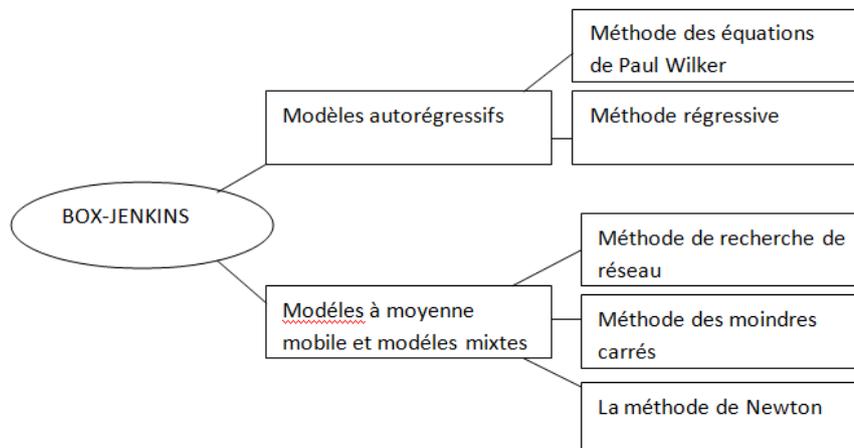


FIGURE 2.1: L'ensemble des opérations sur lesquelles repose la méthode Box Jenkins

2.3 Prédiction basée sur l'intelligence artificielle

L'intelligence artificielle prédictive est une méthode d'analyse de données capable de prédire et de prévoir les besoins ou les événements futurs d'une entreprise. Cela permet, entre autres, d'examiner les tendances futures ou de prévoir les risques et leurs solutions.

2.3.1 Apprentissage automatique

Il existe plusieurs définitions :

- "... explores way of estimating functional dependency from a given collection of data" .
- "...vast set of tools for modelling and understanding complex data"

Autrement dit :

L'apprentissage automatique (en anglais machine Learning), apprentissage artificiel ou apprentissage statistique, est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune.

2.3.2 Processus d'apprentissage

L'apprentissage est probablement la propriété la plus intéressante des réseaux de neurones. Cependant, il ne fait pas référence à tous les modèles, mais aux plus couramment utilisés. L'apprentissage est une étape de développement du réseau de neurones au cours de laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement souhaité. C'est-à-dire que la valeur du poids reliant un neurone d'une couche à une autre change". Considérons le poids w_{ij} reliant le neurone i à son entrée j . Au temps τ , le changement de poids

$$\Delta w_{ij}^{(\tau)} \quad (2.8)$$

peut être simplement exprimé comme :

$$\Delta w_{ij}^{(\tau)} = w_{ij}^{(\tau+1)} - w_{ij}^{(\tau)} \quad (2.9)$$

Par conséquent,

$$w_{ij}^{(\tau+1)} = w_{ij}^{(\tau)} + \Delta w_{ij}^{(\tau)} \quad (2.10)$$

avec $w_{ij}^{(\tau+1)}$ et $w_{ij}^{(\tau)}$ représentent respectivement les nouvelles valeurs et les anciennes aussi du poids w_{ij} .

2.4 Différent type d'apprentissage

Nous pouvons classer les types d'apprentissage en plusieurs catégories très distinctes : apprentissage supervisé, non supervisé et apprentissage par renforcement, comme le montre la figure ci-dessous :

L'apprentissage automatique se décompose en deux étapes : une étape d'entraînement c'est l'apprentissage sur un ensemble de données et une deuxième étape de vérification c'est le teste. Nous aurons donc trois phases : la représentation, l'évaluation et l'optimisation. La phase de représentation consiste à trouver le modèle mathématique le plus adapté. L'évaluation elle mesure l'écart entre le modèle et la réalité des données de tests. Enfin , l'optimisation vise à amenuiser cet écart.

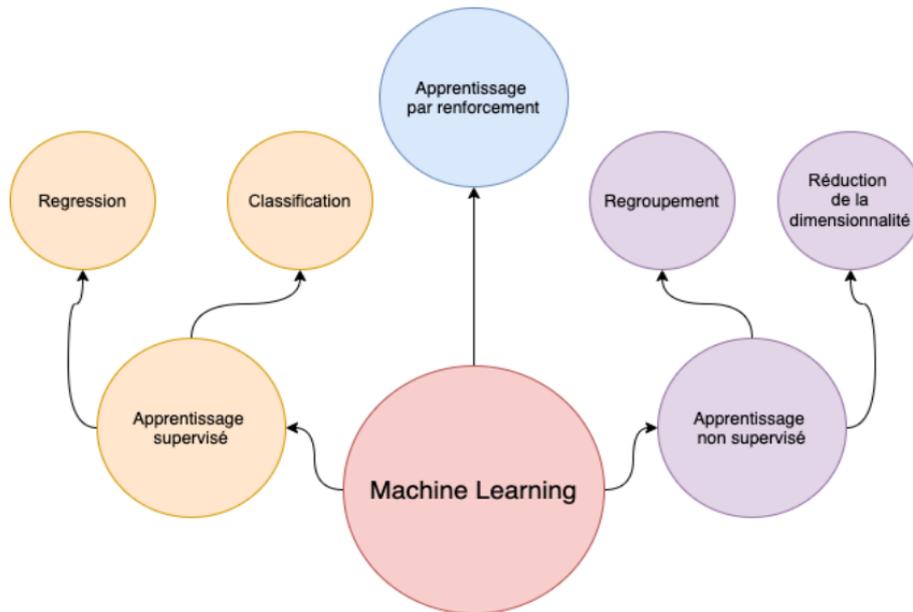


FIGURE 2.2: Les types d'apprentissage automatique [18]

2.4.1 Apprentissage supervisé

On suppose que l'on observe $(X_1, Y_1), \dots, (X_n, Y_n)$ n paires de variables aléatoires indépendantes de même loi que $(X, Y) \sim P$, On note $D_n = (X_1, Y_1), \dots, (X_n, Y_n)$ cet échantillon d'apprentissage.

- Les X_i appartiennent à un espace X , on les appelle variables d'entrées, explicatives ou encore covariables. En général $X \subset \mathbb{R}^d$, $d \geq 1$.
- les Y_i appartiennent à un espace Y , on les appelle variables de sorties, étiquettes ou encore variables à expliquer. En général $Y \subset \mathbb{R}$.

L'apprentissage supervisé consiste à entraîner un modèle en lui fournissant la réponse (label). Cette réponse permet de superviser l'apprentissage du modèle en lui disant à quel point il est loin de la bonne réponse.

Dans un apprentissage supervisé, nous avons un X (variable indépendante) et un Y (variable dépendante) lors de l'entraînement. Cette catégorie se divise en deux sous catégories principales soit la classification et la régression. [19].

- *La classification* : c'est-à-dire que la réponse que nous fournissons au modèle est une catégorie, et non une valeur continue.
- *Régression* : dans ce cas, la réponse que nous donnons au modèle est une valeur numérique continue.

Ce qui distingue la classification de la régression est le type de sortie, pas le type d'entrées. Si la sortie est une catégorie, alors c'est une classification. Et si la sortie est un fichier time-lapse, il s'agit d'une régression.

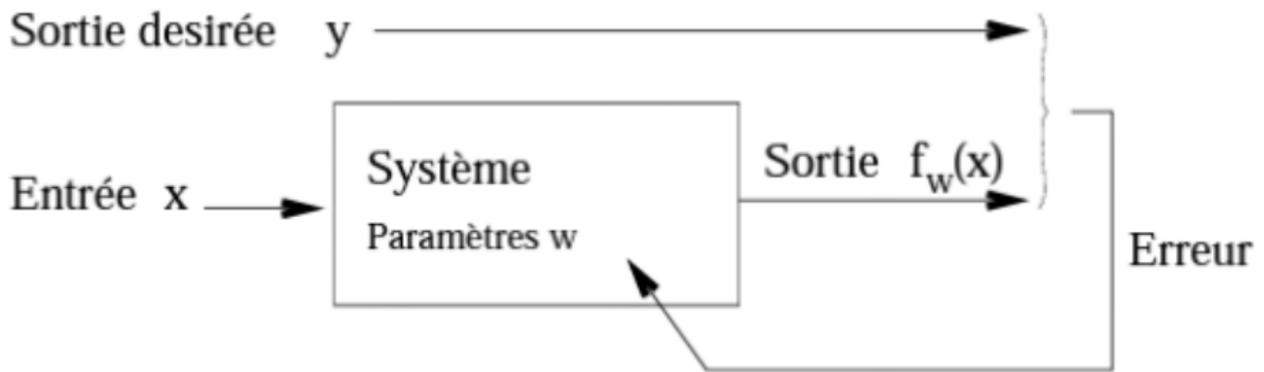


FIGURE 2.3: Schéma d'un modèle supervisé

2.4.2 Apprentissage non supervisé

L'apprentissage non supervisé implique la formation d'un modèle pour trouver des fonctionnalités et extraire des relations entre les données.

Dans un problème non supervisé, nous n'avons pas la réponse exacte que le modèle est censé trouver, nous n'avons que les données d'entrée.

Dans l'apprentissage non supervisé, lorsque former. Il existe deux principaux types d'apprentissage non supervisé Regroupement de catégories ou réduction de dimensionnalité (Réduction de dimensionnalité) [19].

1. *Regroupement* : comprend l'identification de différents groupes dans l'ensemble de données.
2. *Réduction de la dimensionnalité* : principalement utilisée de deux manières. Premièrement, il peut être utilisé pour compresser des données, ce qui a pour effet d'utiliser moins de mémoire et d'espace disque et de réduire le temps de formation. deuxièmement Peut être utilisé pour la visualisation des données.

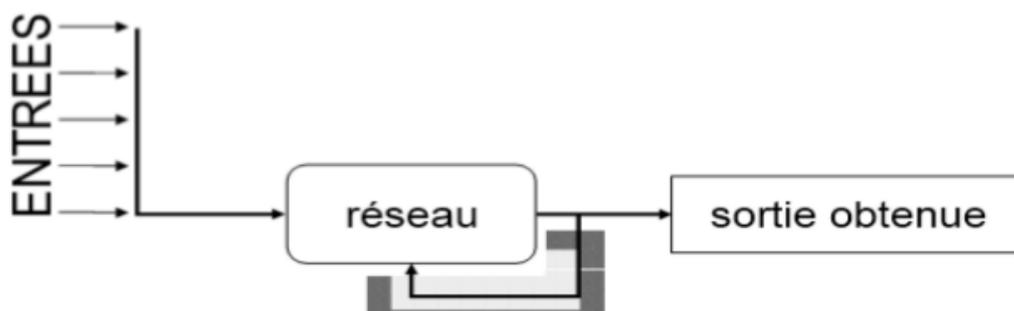


FIGURE 2.4: Schéma d'un modèle non supervisé [20]

2.4.3 Apprentissage par renforcement

L'apprentissage par renforcement consiste en un agent qui interagit avec son environnement. L'agent reçoit des observations et des récompenses de l'environnement. L'observation indique l'état de l'environnement à un moment donné, et la récompense est mesures de performance des

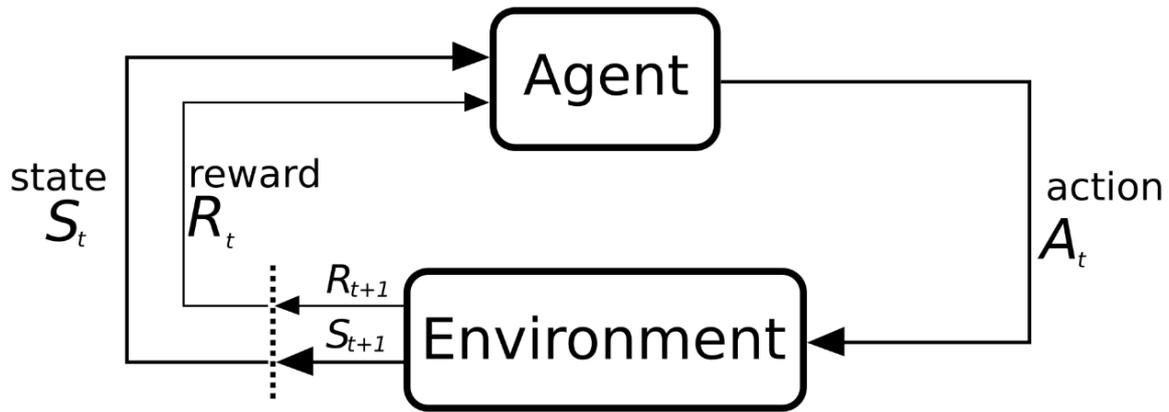


FIGURE 2.5: Schéma d'un algorithme d'apprentissage par renforcement. [21]

agents. Ensuite, l'agent analyse les observations et la récompense qu'il reçoit et décide quelle action il doit prendre.

2.5 Outils d'apprentissage

L'intelligence artificielle est principalement utilisée dans l'éducation, à travers des outils qui aident à améliorer les compétences et les tests. À mesure que les solutions d'IA pour l'éducation mûrissent, on espère que l'IA contribuera à combler les lacunes en matière d'apprentissage et d'éducation et permettra aux écoles et aux enseignants de travailler plus dur qu'auparavant. L'intelligence artificielle peut être puissante, s'autocorriger et simplifier la gestion d'entreprise, donnant aux éducateurs le temps et la liberté de mieux comprendre et de renforcer les capacités humaines que les machines ne peuvent pas. Nous les mentionnons ci-dessous après les moyens de machine learning.

2.5.1 Régression linéaire

La régression linéaire est un algorithme d'apprentissage supervisé dont le but est d'exprimer la variable y d'un ensemble d'entrée x et de l'exprimer dans une fonction linéaire où :

$$y = h(x) = \Theta_i x + \Theta_0 \quad (2.11)$$

où Θ_i est la pente et Θ_0 est la constante de l'équation.

Afin de trouver les inconnues de l'équation exprimant N paires ou triplets... les entrées qui composent un jeu de données, il existe plusieurs techniques dont les plus connues sont les moindres carrés et la méthode des régressions.

Lors de la recherche des valeurs Θ_0 et Θ_i par la méthode des moindres carrés, par exemple, ces coefficients permettent de réduire la fonction de coût $c = \sum_{i=1}^n n(h(x_i) - y_i)^2$, car cette fonction correspond à la somme des écarts au carré entre les prédictions et les valeurs attendues, et nous appelons ces écarts les valeurs résiduelles [22].

2.5.2 Régression logistique

La régression logistique est essentiellement un algorithme de classification supervisée. Dans les problèmes de classification, la variable cible (ou de sortie) y ne peut prendre que des valeurs discrètes pour un ensemble donné de caractéristiques (ou d'entrée) X .

Le modèle construit un modèle de régression pour prédire la probabilité qu'une entrée de données donnée appartienne à la catégorie numérotée "1". Tout comme la régression linéaire suppose que les données obéissent à une fonction linéaire, la régression utilise la fonction sigmoïde pour modéliser les données.

La régression logistique ne devient une technique de classification que lorsqu'un seuil de décision est introduit dans l'image. La définition du seuil est un aspect très important de la régression logistique et dépend du problème de classification lui-même.

La décision de seuil est principalement influencée par les valeurs de précision et de rappel. Idéalement, nous aimerions que la précision et le rappel soient égaux à 1, mais cela arrive rarement.

2.5.3 Les machines à vecteurs supports (SVM)

SVM est l'un des algorithmes d'apprentissage supervisé les plus populaires, qui est utilisé pour les problèmes de classification et de régression.

L'objectif de l'algorithme *SVM* est de créer des lignes de décision optimales ou des limites qui divisent l'espace à n dimensions en classes afin que nous puissions facilement placer de nouveaux points de données dans la bonne classe à l'avenir. Cette frontière de décision optimale est appelée un hyperplan.

SVM choisit les points/vecteurs extrêmes qui aident à créer l'hyperplan. Ces cas extrêmes sont appelés vecteurs de support.

Considérez le diagramme suivant, où deux classes différentes sont classées à l'aide d'une limite de décision ou d'un hyperplan [23]

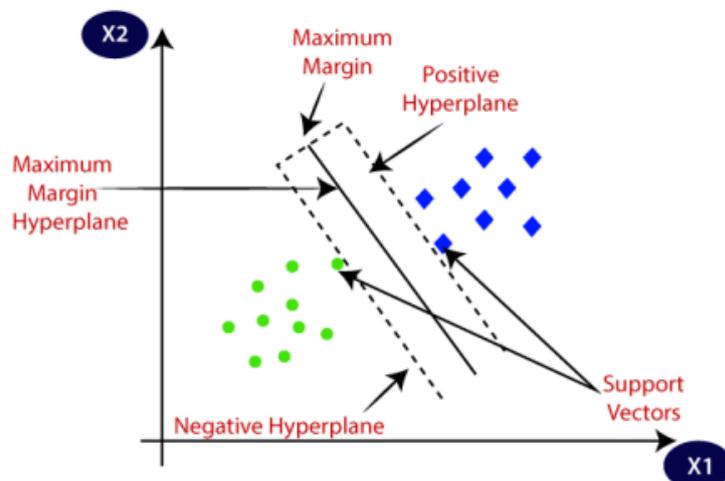


FIGURE 2.6: l'algorithme SVM [24]

Avantages :

- Il permet de traiter des problèmes de classification non linéaire complexe.
- Les *SVM* constituent une alternative aux réseaux de neurones car plus faciles à entraîner.

2.5.4 Les K plus proches voisins (KNN)

Les k plus proches voisins (en anglais k Nearest Neighbors) est une méthode utilisée pour résoudre les problèmes de régression et de classification.

Il s'agit d'une méthode très simple et directe. Pour faire des prédictions, l'algorithme ne calcule

pas de modèle prédictif à partir de l'ensemble d'entraînement. Il n'y a donc pas de phase d'apprentissage proprement dite pour lui, il fait simplement le stockage des données d'apprentissage. Le principe est le suivant : comparer les données de la classe inconnue avec les données de toutes les classes stockées. La nouvelle donnée sera la classe majoritaire parmi ses K les plus proches Voisins dans le sens de distance choisi [25].

Pour trouver le K le plus proche d'une donnée à classer, on peut utiliser la distance euclidienne. Soient deux données représentées par deux vecteurs X et Y, la distance entre ces deux données est donnée par :

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (2.12)$$

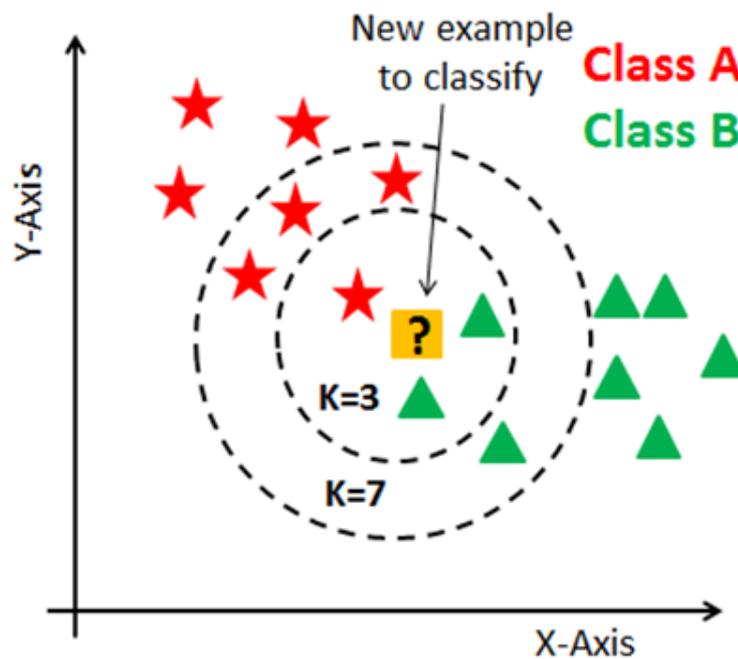


FIGURE 2.7: Les K plus proches voisins [26]

2.5.5 Les réseaux neurones

Les réseaux de neurones, également appelés réseaux de neurones artificiels, sont un sous-ensemble de l'apprentissage automatique et sont au cœur des algorithmes d'apprentissage en profondeur. Leurs noms et leurs structures sont inspirés du cerveau humain. En effet, ces réseaux imitent la façon dont les neurones biologiques s'envoient des signaux les uns aux autres.

Un réseau de neurones artificiels est constitué de différentes couches de neurones artificiels, dont une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie.

Caractéristiques de l'esprit humain, le cerveau humain est composé d'un grand nombre de neurones divers, au nombre d'environ 100 milliards de neurones.

Ces cellules se différencient du reste des cellules du corps par leur capacité à échanger des impulsions et à communiquer à travers un vaste réseau de fibres nerveuses reliées les unes aux autres, chaque cellule étant liée à près de 10 000 autres cellules, et se distinguant par sa capacité à stocker, classer et traiter des informations.

Dendrites : Il représente un ensemble d'entrées par lesquelles une cellule reçoit des informations

Neuron

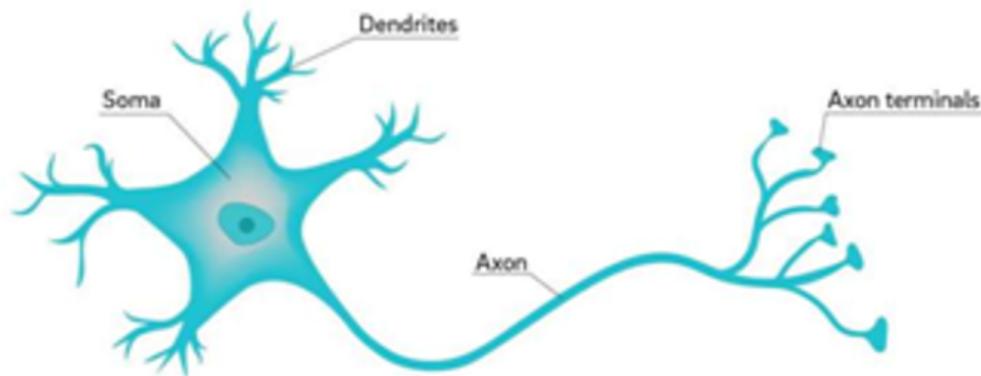


FIGURE 2.8: Neurone biologique [27]

sous forme d'impulsions d'autres cellules.

Axon : C'est une seule fibre nerveuse qui transmet la sortie de la cellule aux fibres nerveuses des autres cellules.

Soma : À l'intérieur de la cellule, les informations qui se présentent sous la forme de signaux provenant des cinq sens sont traitées et stockées.

Synapses : c'est la connexion entre l'axone et les autres dendrites neuronales

Par conséquent, nous disons que les réseaux de neurones artificiels sont des simulations de réseaux de neurones biologiques, ou tels que définis par **Teuvo Kohonen**, qui est l'un des scientifiques les plus éminents. Il dit dans sa définition :

" Les réseaux de neurones artificiels sont des réseaux densément inter-connectés qui comprennent des éléments simples, parallèles et organisés de manière hiérarchique, où ils interagissent avec des objets du monde réel de la même manière qu'un système nerveux naturel interagit avec le monde réel".

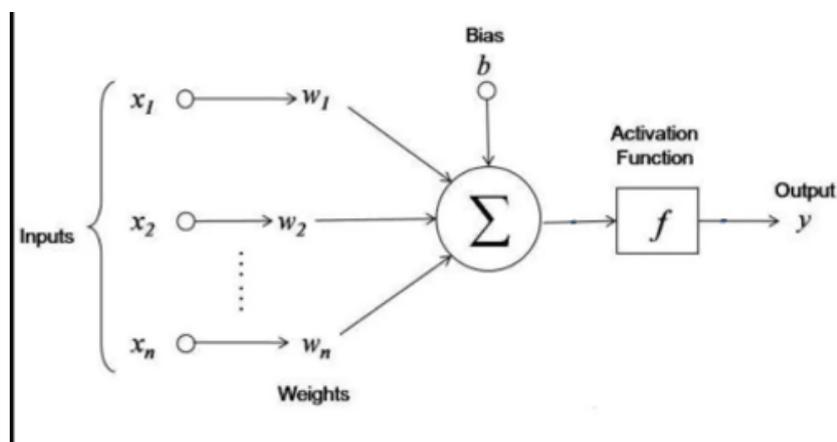


FIGURE 2.9: Un réseau neuronal électronique simulant un réseau neuronal biologique

— L'entrée (input) consiste en un certain nombre d'entrées exprimées par le vecteur X , et elles

sont représentées mathématiquement en imposant des poids W . Ces poids expriment l'importance relative de chaque entrée pour le neurone.

- Le corps cellulaire (l'élément de traitement PE) effectue le processus de collecte des entrées pondérées pour prendre la forme suivante :

$$Total\ input = W_1x_1 + W_2x_2 + \dots + W_nx_n = \sum_{i=1}^n W_ix_i \quad (2.13)$$

Il compare ensuite le résultat à la valeur limite d'une fonction connue sous le nom de fonction (Heaviside), de sorte que si la somme des entrées est supérieure ou égale à cette fonction, la sortie sera 1, et si elle est inférieure, la sortie sera 0.

- Sortie (Output) : la sortie est 0 ou 1 et est transmise au neurone suivant. Nous écrivons la valeur de la sortie y comme suit :

$$y = f\left(\sum_{i=1}^n W_ix_i\right) \quad (2.14)$$

Analogie entre réseau de neurones biologique et artificiel dans le tableau suivant :

Réseau neuronal biologique	Réseau neuronal artificiel
Soma	Neurone
Dendrite	Entrée
Axon	Sortie
Synapse	Poids

2.6 fonction d'activation

Définition biologique

Une fonction d'activation est une abstraction qui représente le taux de déclenchement d'un potentiel d'action dans une cellule. Dans sa forme la plus simple, cette fonction est binaire, c'est-à-dire que le neurone se déclenche ou non.

Définition artificiel

La fonction d'activation (ou fonction de seuil, voire fonction de transfert) d'un neurone artificiel définit les performances de ce neurone donné à partir d'une entrée ou d'un ensemble d'entrées. "énumère les différentes fonctions de transfert pouvant être utilisées comme fonctions d'activation des neurones, les trois les plus couramment utilisées sont les fonctions de "seuil" ("hard limit" en anglais

Nom de la fonction	Relation d'entrée/sortie	Icône	Nom Matlab
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlim
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlims
linéaire	$a = n$		purelin
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$		satlin
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$		satlins
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$		poslin
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$		logsig
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
compétitive	$a = 1$ si n maximum $a = 0$ autrement		compet

FIGURE 2.10: Les fonctions d'activation $a = f(x)$ [19]

2.7 Types de réseaux de neurones

Les réseaux neuronaux peuvent être classés en plusieurs types plus courants qui sont utilisés à des fins différentes.

2.7.1 Peceptron

Histoire et Définition :

C'est la forme la plus simple d'un réseau de neurones, car il se compose d'un seul neurone, avec des poids synaptiques ajustables, un nom donné par le neuroscientifique "Frank Rosenilat" à un groupe d'expériences qui ont commencé à simuler l'esprit humain dans le processus de pensée entre les années 1957 et 1962, qui ont conduit à son innovation du premier réseau de neurones artificiels de l'histoire, "peptron Network".

Qui est devenu plus tard la base de l'intelligence artificielle et de l'apprentissage profond tel que nous le connaissons aujourd'hui.

Comment est née l'idée de pesptron?L'histoire commence avec " Waterpets", un enfant sans abri qui passe la moitié de sa journée à étudier à la bibliothèque et l'autre moitié à errer dans les rues,Water a étudié tous les travaux du mathématicien et logicien "Bertrand Russell" et a corrigé ses erreurs à l'âge de 12 ans.

"Russell" est donc entré dans sa classe universitaire après avoir été émerveillé par son intelligence. Dans le même collège, des années plus tard, Waterpets a rencontré le professeur Warren McCulloch, et ils ont convenu que le cerveau n'est rien d'autre qu'une machine, et ensemble, ils ont décidé de travailler pour le comprendre et le déchiffrer,Et en 1943.

les deux ont publié leur article de recherche (The Logical Calculus and Integration of Basic Ideas in Neuronal Activity).

Où j'ai proposé un modèle mathématique logique pour représenter le processus de pensée dans les neurones biologiques. 15 ans plus tard, le scientifique "Frank Rosenblatt" a passé en revue les recherches de Warren et Bates, et il a eu une idée folle (puisque l'on sait comment fonctionnent les réseaux neuronaux biologiques et que l'esprit n'est qu'une machine, pourquoi ne pas créer un réseau neuronal électronique à l'aide d'un ordinateur?) et en fait, en 1958, il a lancé son innovation s'appelle Perceptron Synthetic Network.

1. Réseau mono-couche :

Il se compose d'une seule couche de nœuds qui connectent directement les entrées du réseau à ses sorties.

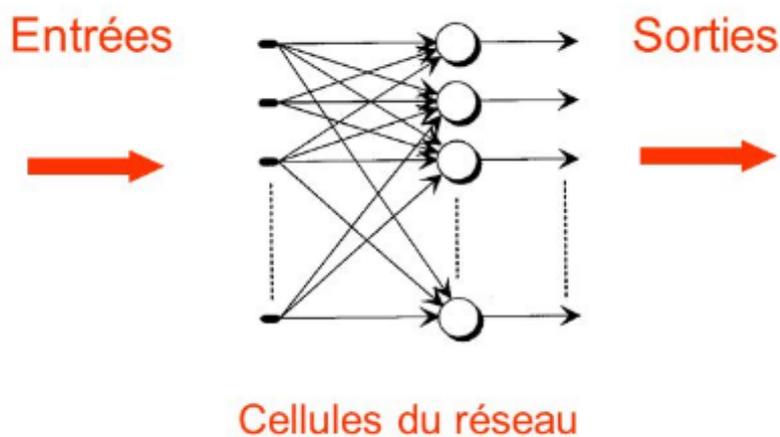


FIGURE 2.11: Réseau mono-couche

2. Le réseau comporte deux couches :

Il contient deux couches de nœuds inter-connectés, dont l'un est spécialisé dans la réception des données et l'autre dans la sortie des résultats. Ce modèle est appelé modèle matriciel en raison de la possibilité de représenter les poids sous la forme d'une matrice.

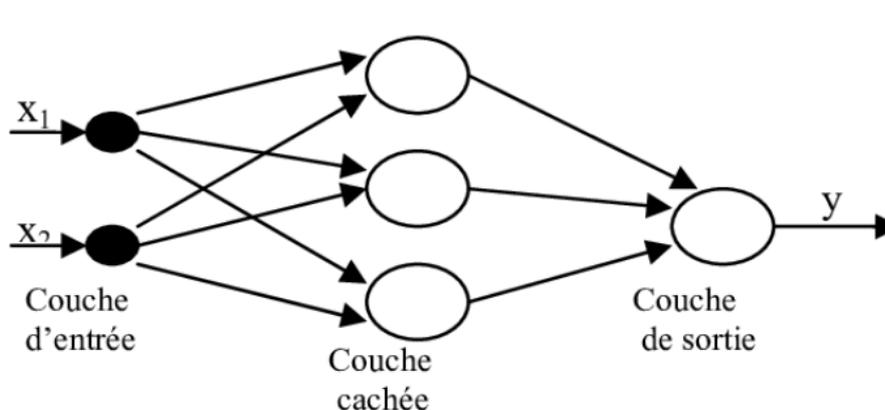


FIGURE 2.12: Réseau de neurones statique à deux couches

3. **Perceptron multicouche :**

Il contient au moins trois couches d'éléments interdépendants, la première couche étant appelée couche d'entrée (input layer), la dernière couche étant appelée couche de sortie (output layer) et les couches intermédiaires étant appelées couches cachées (Hidden layers).

Ces couches cachées sont responsables du traitement des données.

Ce modèle est plus efficace que le précédent et est utilisé pour résoudre des problèmes complexes [28], Nous expliquons ces couches dans la figure suivante :

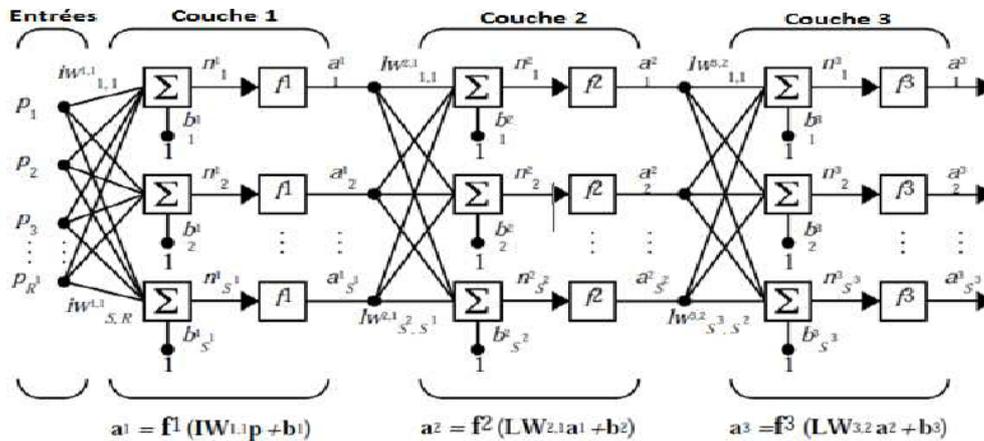


FIGURE 2.13: Perceptron multicouche [29]

Couche masquée : La couche cachée présente des couches intermédiaires d'entrée et de sortie. Il effectue tous les calculs pour trouver les fonctionnalités cachées et motifs.

2.7.2 **Modèle Hopfield**

Dans un article scientifique publié par Hopfield en 1982, il a montré que les réseaux auto-connectés ont deux propriétés importantes que nous mentionnons brièvement :

Ces réseaux doivent atteindre un état stable à la fin, quel que soit leur état initial.

Les états de stabilité peuvent être construits en modifiant les poids des connexions entre les éléments du réseau ou en modifiant leurs fonctions de seuil.

Nous mentionnons certaines des caractéristiques les plus importantes de ce modèle :

Mémoire nappée au contenu (Content-addressable Memory), Contrôle distribué et synchrone (Distributed Asynchronous Control), Représentation distribuée de l'information, ...etc

2.7.3 **Modèle Boltzmann**

Parce qu'il y a beaucoup de problèmes dans le modèle de **Hopfield**, comme le problème de la fausse stabilité, et de ce point de vue, un groupe de scientifiques, dont le scientifique HOUTON, a développé un modèle basé sur la théorie d'un physicien autrichien Boltzmann, et ils ont atteint le utilisation d'une nouvelle fonction de seuil qui a résolu les problèmes du modèle mentionné précédemment.

2.7.4 **Modèle Back Propagation**

Ce modèle consiste en un réseau de neurones artificiels multicouches, et il est enseigné de manière dirigée et itérative jusqu'à ce que les performances du réseau atteignent un niveau élevé et le niveau requis.

2.7.5 Modèle de Rétropopagation

Ce modèle se caractérise par contenir un ensemble de couches, mais chaque couche a sa propre méthode d'apprentissage, qui diffère des autres couches.

2.8 Apprentissage neuronal

Sans aucun doute, reproduire la capacité d'apprentissage humaine est l'un des objectifs les plus importants de la modélisation de réseaux de neurones artificiels. L'apprentissage est l'un de leurs attributs de base. C'est le processus qui permet au réseau de se spécialiser sur un problème spécifique en fonction de son expérience. Les règles de base d'apprentissage des réseaux de neurones sont les suivantes .

Quelques règles différentes d'apprentissage

Dans la figure suivante, nous résumons certaines des règles les plus populaires et les plus fréquemment utilisées.

Règles d'apprentissage	Supervisé	Non supervisé
Recherche du minimum local	Règle du perceptron (Rosenblatt 1958) $W(k+1) = W(k) + eX^T$	Règle d'Hebb (Changeux et Danchin 1976) $W(k) = W(k-1) + \rho y(k)X^T(k)$
	Règle de Widrow-Hoff (1960) $W(k+1) = W(k) + \alpha(t-y)_k \frac{X(k)}{\ X(k)\ ^2}$	Règle de Hinton (1989) $W(k) = (1-\gamma)W(k-1) + \rho y(k)X^T(k)$
	L'Algorithme de retropropagation (Werbos 1974) $W^m(k+1) = W^m(k) - \rho s^m (y^{m-1})^T$ $\frac{\partial \hat{F}}{\partial W^m} = s^m y^{m-1}$ $s^m = \frac{\partial \hat{F}}{\partial n^m} = \left(\frac{\partial n^{m+1}}{\partial n^m} \right) \frac{\partial \hat{F}}{\partial n^{m+1}} = \left(\frac{\partial n^{m+1}}{\partial n^m} \right) s^{m+1}$	Règle d'Oja (1982) $W(k+1) = W(k) + \rho y(k)X(k) - \rho y(k)^2 W(k)$
		Règle de Yuille et al. (1989) $W(k+1) = W(k) + \rho (y(k)X(k) - \ W(k)\ ^2 W(k))$
Recherche du minimum global	Technique d'adoucissement simulé Algorithme de Métropolis et al. (1953) : $W(X \rightarrow X') = \begin{cases} 1 & \text{si } \Delta E < 0 \\ e^{-\Delta E/KT} & \text{autre} \end{cases}$ Règle de Geman et Geman (1984) : $T(t) = \frac{T_0}{1 + \ln t}$	Règle de Hassoum (1995) $W(k+1) = \rho \left[y(k)X(k) - \lambda W(k) \left(1 - \frac{1}{\ W(k)\ } \right) \right]$
	Algorithmes génétiques (Holland 1975; Goldberg 1989)	

FIGURE 2.14: Présentation quelques types d'apprentissage [20]

2.9 Apprentissage profond

Le Deep learning ou apprentissage profond est un type d'intelligence artificielle. L'apprentissage profond est un ensemble de méthodes d'apprentissage automatique ayant pour but la modélisation à haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires. Ces techniques ont permis des progrès importants et rapides dans les domaines de l'analyse du signal sonore ou visuel et notamment de la reconnaissance faciale, de la reconnaissance vocale... etc [30].

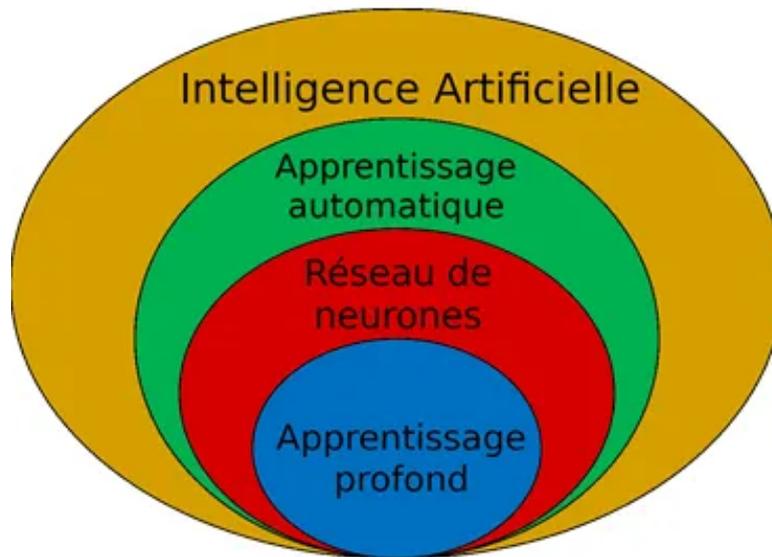


FIGURE 2.15: Intelligence artificielle, l'apprentissage automatique et l'apprentissage en profondeur [31]

2.9.1 Domaines d'applications

Ces techniques se développent dans le domaine de l'informatique appliquée aux *NTIC* (Reconnaissance visuelle et vocale notamment) à la robotique, à la bio-informatique, la reconnaissance ou comparaison de formes, la sécurité, la santé, etc..., la pédagogie assistée par l'informatique, et plus généralement à l'intelligence artificielle.

L'apprentissage profond peut par exemple permettre à un ordinateur de mieux reconnaître des objets hautement déformables et/ou analyser par exemple les émotions révélées par un visage photographié ou filmé, ou analyser les mouvements et position des doigts d'une main, ce qui peut être utile pour traduire le langage des signes, améliorer le positionnement automatique d'une caméra, etc...

Elles sont utilisées pour certaines formes d'aide au diagnostic médical (ex : reconnaissance automatique d'un cancer en imagerie médicale), de prospective ou de prédiction.

2.10 Quelques réseaux profonds

2.10.1 Réseaux de neurones récurrents *RNN*

Recurrent Neural Network (*RNN*) est considéré comme un réseau avec mémoire .

Un réseau neuronal récurrent est un type de réseau neuronal qui contient des boucles utilisées

pour stocker des informations dans le réseau. Les *RNN* utilisent les sorties précédentes comme entrées supplémentaires, ce qui leur permet de prédire très précisément ce qui va se passer ensuite [32].

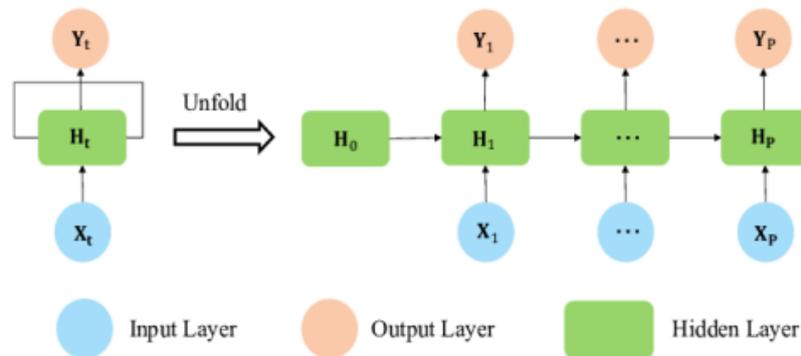


FIGURE 2.16: Architecture de *RNN* [33]

- X : couche d'entrée
- H : couche cachée
- Y : couche de sortie

A tout moment, d'entrée actuelle est la combinaison des entrées à $x(t)$ et $x(t-1)$.

2.10.2 Réseaux de neurones profonds *DNN*

Les réseaux de neurones profonds sont également connus sous le nom de perceptrons multicouches.

Un réseau de neurones est dit profond s'il contient au moins une couche cachée; il peut contenir des millions de neurones organisés en couches où l'information ne circule que de la couche d'entrée vers la couche de sortie. Ils sont utilisés en apprentissage profond pour concevoir des mécanismes d'apprentissage supervisés et non supervisés.

2.10.3 Réseaux Neurones Convolutifs *CNN*

Le réseau de neurones convolutifs est l'un des modèles de classification d'images connus Le plus puissant de l'apprentissage en profondeur. Il permet d'attribuer automatiquement à chaque image donnée en entrée sous forme de matrice de pixels un label correspondant à sa classe [34].

L'architecture du *CNN* possède deux parties :

- **Partie convolutive** : Son but est d'extraire des informations spécifiques Appelez chaque image une "fonctionnalité" en appliquant une opération de filtrage convolution. Ce processus de filtrage produit de nouvelles images, appelées cartes de contraction, en répétant l'opération, on obtient les caractéristiques de rétrécissement de l'image par rapport à sa taille initiale. Ces valeurs de la dernière caractéristique sont concaténées dans Un vecteur nommé codes *CNN*. Une image se représente en 3 dimensions :

1. Deux dimensions pour une image en niveaux de gris qui correspondent à la largeur et à la hauteur de l'image.
2. Une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales (Rouge, Vert, Bleu).

— **Partie classification** : Le vecteur obtenu dans la section précédente est l'entrée de ce bloc consiste en un réseau multi-couche. les valeurs dans le vecteur Transformer l'entrée à l'aide de plusieurs combinaisons et fonctions linéaires Activez pour renvoyer un nouveau vecteur en sortie.

Le but du *CNN* est de résoudre le problème des réseaux de neurones traditionnels qui souffrent du problème de dimension de ses couches .

2.10.4 Le LSTM (Long Short-Term Memory)

Le LSTM a été proposé en 1997 par Sepp Hochreiter et Jürgen Schmidhuber ils ont introduit des cellules mémoires, une forme de mémoire intermédiaire permettant de stocker des informations importantes sur une période de temps plus longue que les RNN existants [35].

Les LSTM sont des blocs de construction spéciaux des réseaux de neurones récurrents (RNN) doté d'une mémoire à court terme à long terme. C'est une évolution de RNN qui résout le problème du gradient disparaissant c'est-à-dire que le gradient de poids diminue progressivement pendant l'entraînement, de sorte que le réseau ne stocke plus d'informations utiles. C'est-à-dire que le gradient de poids diminue progressivement pendant l'entraînement, de sorte que le réseau ne stocke plus d'informations utiles.

Les cellules LSTM ont trois types de portes :

- la porte d'entrée régule le flux d'informations .
- la porte de mémorisation et d'oubli veille à ce que les informations non importantes soient oubliées.
- la porte de sortie détermine quelles informations sont transmises à l'étape suivante.

CONCLUSION

Les réseaux de neurones sont considérés comme un outil très puissant dans la prédiction, ils sont connus par leur puissance d'apprentissage. Dans ce chapitre nous avons présenté l'importance de prédiction et la prédiction basée sur L'IA, et illustrons les outils d'apprentissage. Dans le chapitre suivant, nous allons présenter le lieu de stage. Ensuite nous présentons les données acquises de la consommation d'énergie. Pour une conversion d'une série non stationnaire en une série stationnaire.

Chapitre 3

Acquisition et pré-traitement de données

INTRUCTION

Ce présent chapitre montre les différentes phases de modélisation et d'acquisition de données afin d'implémenter le système de prédiction. Dans un premier lieu, nous allons présenter le lieu de stage qui est le service Division Relations Commerciales de Sonelgaz. Ensuite, nous présentons les données acquises qui concernent les communes Bordj Bou Arreridj et Ras El Oued. Un test de stationnarité est nécessaire pour avoir des jeux de données normalisées. La méthode est différences logarithmiques.

3.1 Présentation de lieu de stage

3.1.1 Lieux

La formation s'est déroulée dans la société centrale Sonelgaz située dans la zone dite (Bab Al-Souk) en face de l'école Ben Badis, nous montrons l'emplacement plus précisément dans la photo suivante :



FIGURE 3.1: Site de l'entreprise de formation

3.1.2 Présentation de l'entreprise

Sonelgaz, ou Société Algérienne de l'électricité et du gaz, a été créée en 1969 en remplacement de l'entité précédente Électricité et Gaz Algérie (EG A), créée en 1947. Ses principales missions sont le transport et la distribution de l'électricité, ainsi que le transport et la distribution de gaz naturel au moyen de conduites et de pipelines sur le marché national.

3.1.3 Service distribution

Nous expliquons cette section des services comme suit :

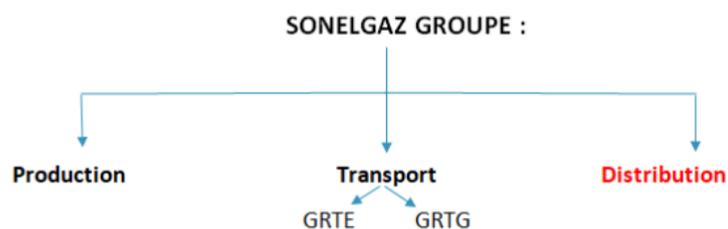


FIGURE 3.2: SONELGAZ GROUPE

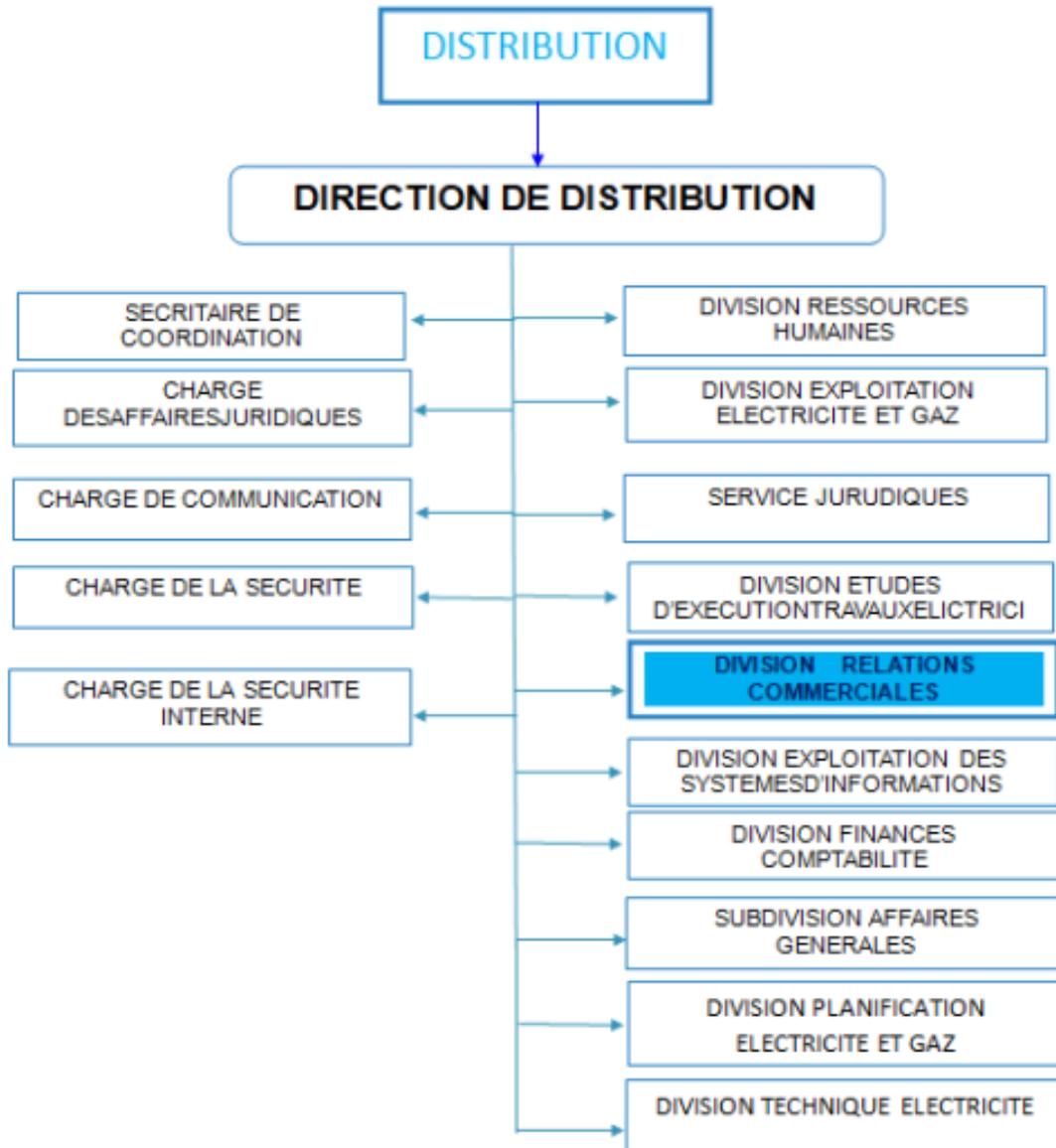


FIGURE 3.3: Source :Régie de distribution d'énergie Sonelgaz,BBA

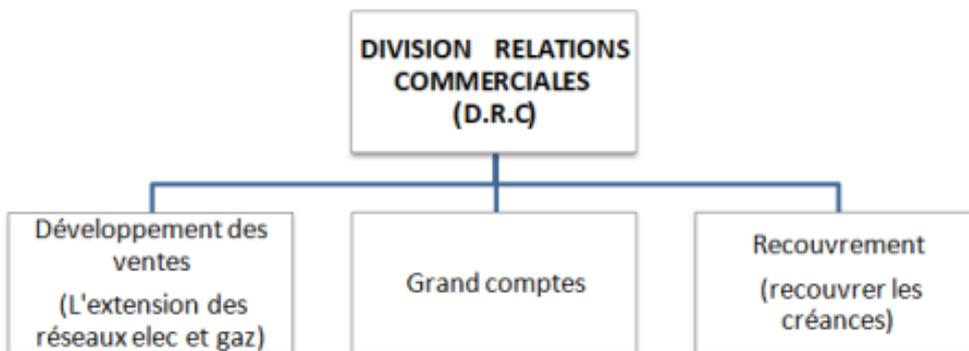


FIGURE 3.4: Source :Régie de distribution d'énergie Sonelgaz,BBA

3.1.4 L'intermédiaire entre les clients et sonelgaz :

Sonelgaz Central ne traite qu'avec ses agences affiliées. Quant à la médiation entre les citoyens, cet intérêt s'exerce à travers les cinq agences de Sonelgaz qui sont les suivantes :

- BORDJ BOU ARRERIDJ 01 (BENBADIS)
- RAS EL OUED
- MANSOURAH
- BORDJ BOU ARRERIDJ 02 (MOKRANI)
- BORDJ EL GHDIR

Chaque agence a une direction que nous expliquons dans la figure suivante :

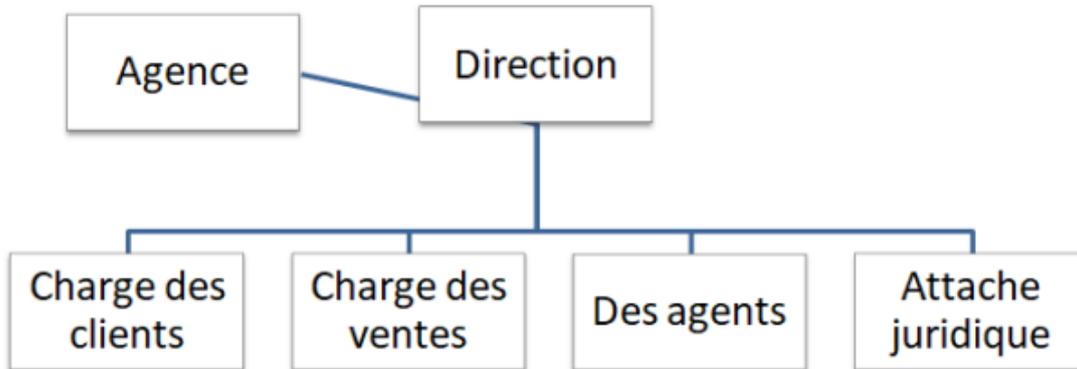


FIGURE 3.5: tâches de chaque agence

3.1.5 Type d'énergie

La société de négoce Sonelgaz traite différents types d'énergie, elle produit ou achète de l'énergie à haute intensité, puis réduit sa gradation pour atteindre le consommateur à moyenne ou faible intensité, ce que nous expliquons dans le tableau suivant :

ELECTRICITE	GAZ
<i>Basse tension</i>	<i>Basse pression</i>
<i>Moyen tension</i>	<i>Moyen pression</i>
<i>Haute tension</i>	<i>Haute pression</i>

FIGURE 3.6: Formes d'énergie

Au final, nous avons obtenu du service distribution de la société Sonelgaz les données de consommation d'énergie pour toutes les communes de Bordj Bou Arreridj de début 2019 jusqu'à 2022, qui se trouve dans la figure suivante : les variables utilisées dans l'étude peuvent être classées

tiem	region1	region2	region3	region4	region5	region6	region7	region8	region9	region10	region11	region12	region13	region14	region15	region16	region17
2019/M1	15421931	6895044	1061	3011053	12244	11581	2972861	2600	2 753 274	2539849	826333	7707	2190183	186733	182083	0	
2019/M2	12979242	770374	0	442538	2805666	1032743	14233	2346778	3713378	2160	2875088	674	1153612	2145	309888	0	1192
2019/M3	13 574 280	1 903 145	1 144 401	993	643	342 315	1 736 752	3 006	871 331	3 381	835 585	1 954 552	2 555 613	888 854	518 357	870 282	322
2019/M4	14 891 070	7 068 692	824	2 734 165	0	0	2 719 448	0	2 533 969	2 531 728	823 033	6 690	2 041 665	209 000	203 636	15 587	
2019/M5	12602691	756851	0	418195	2572735	951713	598	2345369	3 509 913	0	2764149	2183	1168332	0	313693	0	1183
2019/M6	16339031	2075522	1309535	932	0	335211	2044641	0	934 152	103	900325	2128985	3040821	800427	531349	789827	371
2019/M7	21599330	9166457	2315	4860737	353	496	3783135	2600	3 737 340	3851686	1108333	5695	2697838	240748	222440	1902	
2019/M8	24858816	1253600	0	939276	5312225	1250871	8878	3858838	5 073 994	8580	5034747	4519	1802665	0	517443	1 710	1802
2019/M9	25 181 032	2 684 627	1 866 886	595	3 403	379 464	2 724 905	27	1 218 628	213	1 153 977	3 122 871	3 307 467	1 041 383	615 248	1 161 791	426
2019/M10	22 388 472	10 644 738	2 363	4 630 262	807	1 291	4 563 846	2 947	3 482 389	3 584 742	1 250 136	11 888	3 148 965	240 934	245 253		
2019/M11	14 787 186	873 687	1 060	588 989	3 467 481	1 090 572	341	2 534 172	3 931 108	15 960	3 642 876	761	1 369 263	322	408 999	12 826	1 317
2019/M12	17 416 131	2 159 262	1 537 339	12 886	46 077	536 371	1 860 588	4 500	1 014 219	4 172	813 558	2 318 819	2 807 756	937 232	779 212	938 181	359
2020/M1	13 052 553	6 839 457	-	2 898 756	-	-	3 045 417	-	3 665 330	2 553 346	855 056	-	2 216 381	180 778	186 134	-	-
2020/M2	12 649 571	763 227	-	463 968	2 367 626	841 243	12 283	2 271 016	3 491 697	-	2 608 056	1 753	1 986 967	287 183	-	-	1 148
2020/M3	15 777 599	1 886 655	1 278 895	6 422	7 332	296 276	1 814 986	-	911259	340	844 943	1 780 227	2 311 547	669 992	404 078	965 612	312
2020/M4	12 823 748	7 114 871	384	2 826 516	19 196	-	2 896 362	-	2 493 128	2 698 582	741 142	-	2 990 555	141 888	153 079	-	-
2020/M5	14 430 543	809 013	-	529 642	3 143 156	996 780	3 567	2 393 855	3 449 003	-	2 752 221	8 255	1 226 907	3 217	354 114	1	1 267
2020/M6	16 226 049	1 811 844	1 411 773	22 963	8 877	360 669	1 577 566	4 675	901 640	179	730 598	1 948 675	2 832 707	818 498	604 194	751 729	313
2020/M7	13 227 812	7 362 422	16 943	3 675 276	70 672	4 205	3 079 980	29 554	3 175 820	3 126 844	1 096 815	6 686	2 548 684	196 010	226 638	744	2
2020/M8	15 741 686	1 031 699	10 435	712 040	3 822 495	1 228 110	5 493	3 221 691	4 327 616	3 227	3 833 462	1 254	1 372 628	8 596	424 074	-	316
2020/M9	23 914 280	2 506 936	1 683 761	8 640	2 776	380 667	2 652 706	1 680	1 117 696	515	1 139 856	2 807 042	3 905 124	1 057 498	607 748	1 188 858	438
2020/M10	20 207 296	10 458 501	-	4 878 131	1 433	1 550	4 196 937	-	4 063 042	4 026 477	1 404 545	3 292	3 910 527	259 315	297 445	270	-
2020/M11	15 926 447	954 187	1 316	536 597	3 762 364	1 033 885	2 477	2 664 682	4 769 094	1 439	4 024 836	-	1 398 949	3 281	426 877	1 066	1 439
2020/M12	19 215 246	2 068 819	1 462 912	11 840	1 337	333 707	2 033 907	2 277	1 063 334	-	894 567	2 148 351	3 113 125	1 014 534	733 967	1 036 969	356
2021/M1	13 088 496	7 148 737	1	2 818 507	-	-	2 843 026	1	3 088 228	2 406 911	825 927	-	2 486 106	180 117	199 607	11 114	-
2021/M2	12 176 287	827 282	1 115	430 496	2 672 019	866 462	6 772	2 295 129	4 074 786	1 763	2 964 774	5 837	1 186 225	827 353	359 863	-	980
2021/M3	17 623 122	1 700 122	1 416 138	419	43 083	281 137	1 648 863	137	1 082 546	21 495	801 519	1 982 473	2 995 769	823 422	556 076	946 505	334
2021/M4	13 656 592	7 334 435	12 769	3 071 704	2 034	190	3 024 665	-	3 817 765	2 772 819	808 354	194	2 580 202	214 706	201 198	-	-

FIGURE 3.7: Base de données de consommation d'énergie ,Source :Régie de distribution d'énergie Sonelgaz,BBA

comme suit :

- la variable indépendante est le temps.
- la variable dépendante est la consommation.

Nous symbolisons ces variables comme suit :

- T : Il représente le temps, c'est-à-dire les mois (du 1 janvier 2019 au 12 décembre 2022).
- Y : Il représenté les consommations mensuelles.

Discussion :

En raison d'un défaut dans les données de certaines communes dont la consommation d'énergie n'a pas été enregistrée dans certains mois. Nous avons décidé de ne pas l'étudier, alors nous n'avons pris que la commune de BOEDJ BOU ARRERIDJ et RAS EL OUED .

3.2 Objectif (Prédiction à moyen terme)

Notre base de données est constituée de données enregistrées sur une base mensuelle, Par conséquent,la prédiction sera jusqu'à trois mois ou plus, et cela est classé dans une prédiction à moyen terme, qui est un type de prédiction proportionnée à la planification de la production d'énergie, au budget et à l'analyse de divers plans d'exploitation. Aussi, prévoir à moyen terme nous aide beaucoup pour planifier l'entretien du réseau électrique et regarder les prix de l'électricité ainsi que pour contrôler la production et réduire la perte de cette énergie et préserver l'environnement....etc.

3.3 Présentation des différentes séries

Comme nous l'avons dit précédemment, l'étude sera limitée à deux communes, Bordj Bou Arreridj et Ras El Oued, en raison de la présence de données complètes et claires, contrairement au reste des communes qui manquent de données. les données des deux séries peuvent être représentées comme suit :

— Commune de Bordj Bou Arreridj

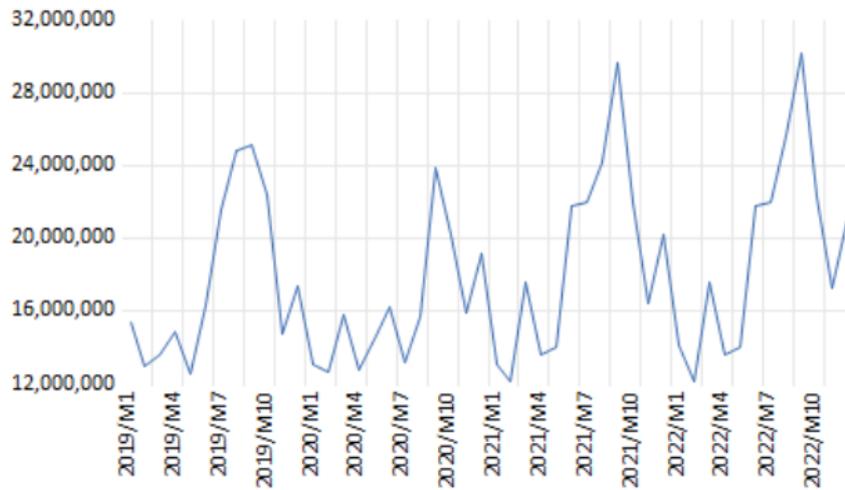


FIGURE 3.8: La consommation d'énergie des 4 dernières années pour la commune de BBA

Analyse graphique;

A partir de la courbe graphique, on remarque une fluctuation des valeurs de consommation de début mars 2019 à octobre 2019, à partir de cette date une grande stagnation de la consommation commence jusqu'en septembre 2020.

Ensuite, les valeurs recommencent à augmenter lentement et avec un forte probabilité que cela soit dû à la crise résultant de la pandémie de COVID-19 dont le monde a été témoin et la faible possibilité reportez-vous à d'autres raisons secondaires.

— Commune de Ras El Oued

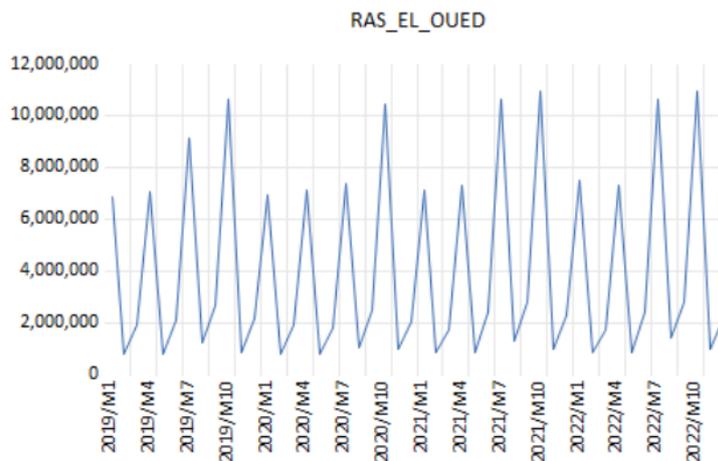


FIGURE 3.9: La consommation d'énergie des 4 dernières années pour la commune de RAS EL OUED

Analyse graphique;

Nous remarquons qu'il y a une augmentation et une diminution continues de cette série chronologique au cours de ces quatre années.

3.4 Pré-traitement des deux séries

Tous les résultats obtenus dans ce qui suit ont été préparés par nous selon le programme Eviews 12 .

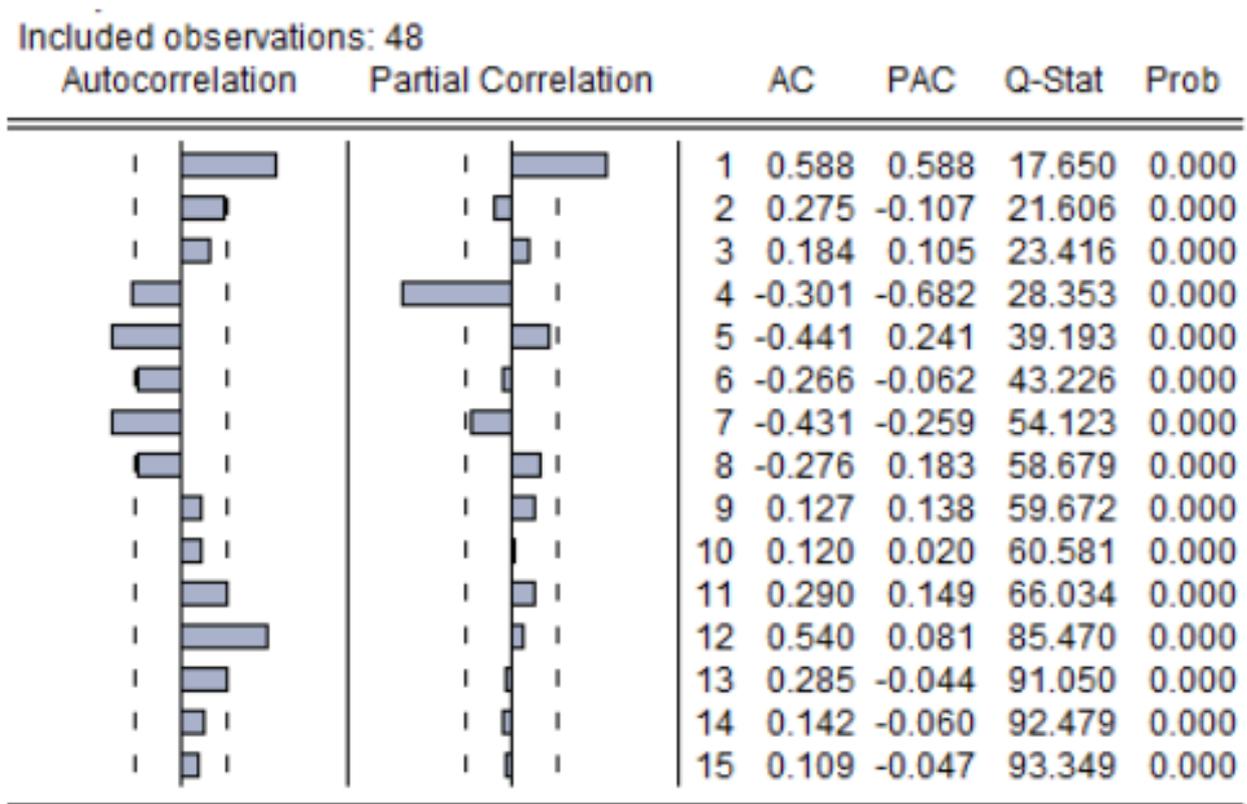
3.4.1 Pré-traitement de série chronologique Bordj Bou Arreridj**— Correlogram**

FIGURE 3.10: Correlogram of BBA

A travers la figure et ci-dessus, on remarque que la corrélation n'est pas significative car elle est loin de zéro et sa décroissance est très lente, et à partir de là on accepte l'hypothèse nulle que la série non stationnaire.

— **Dickey-Fuller test with trend and intercept;**

Augmented Dickey-Fuller Unit Root Test on BBA				
Augmented Dickey-Fuller test statistic		-4.694734	0.0021	
Test critical values:	1% level	-4.198503		
	5% level	-3.523623		
	10% level	-3.192902		
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(BBA)				
Method: Least Squares				
Date: 04/18/23 Time: 13:12				
Sample (adjusted): 8 48				
Included observations: 41 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
BBA(-1)	-1.117640	0.238062	-4.694734	0.0000
D(BBA(-1))	0.911097	0.207587	4.388979	0.0001
D(BBA(-2))	0.531159	0.208372	2.549089	0.0158
D(BBA(-3))	0.790465	0.153082	5.163671	0.0000
D(BBA(-4))	0.161401	0.197911	0.815525	0.4208
D(BBA(-5))	0.215921	0.168803	1.279130	0.2100
D(BBA(-6))	0.517703	0.166789	3.103943	0.0040
C	17721442	3968997.	4.464968	0.0001
@TREND("1")	84798.23	35801.62	2.368559	0.0241
R-squared	0.761619	Mean dependent var	-13472.20	
Adjusted R-squared	0.702024	S.D. dependent var	4518738.	

FIGURE 3.11: Dickey-Fuller test

A travers le tableau, il nous apparaît clairement que la valeur de la probabilité du coefficient de tendance générale (0,0241) est inférieure à (0,05), et cela témoigne d’une tendance générale dans la série, et à partir de laquelle nous acceptons l’hypothèse nulle que la série a une tendance générale et qu’elle est non stationnaire.

3.4.2 Pré-traitement de série chronologique Ras El Oued

— **Correlogram;** À travers la figure, nous remarquons que les coefficients sont significativement

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.281	-0.281	3.0019	0.083
		2	0.003	-0.082	3.0024	0.223
		3	-0.099	-0.132	3.3978	0.334
		4	-0.036	-0.115	3.4518	0.485
		5	-0.100	-0.177	3.8828	0.566
		6	-0.067	-0.205	4.0849	0.665
		7	0.189	0.069	5.7316	0.571
		8	-0.158	-0.161	6.9336	0.544
		9	0.011	-0.152	6.9396	0.643
		10	-0.002	-0.100	6.9397	0.731
		11	0.071	-0.028	7.2095	0.782
		12	-0.003	-0.018	7.2100	0.843
		13	-0.049	-0.108	7.3513	0.883
		14	0.038	-0.091	7.4426	0.916
		15	-0.015	-0.021	7.4579	0.944
		16	-0.054	-0.110	7.6594	0.958

FIGURE 3.12: Correlogram of Ras El Oued

égaux à zéro, c’est-à-dire qu’ils diminuent progressivement vers zéro, et à partir de là, nous rejetons l’hypothèse nulle que la série est non stationnaire, et nous acceptons l’alternative que la série est stationnaire, c’est-à-dire que les coefficients de corrélation sont dans le domaine de confiance.

— **Dickey-Fuller test with trend and intercept;**

À travers la figure, il est clair que la valeur de la probabilité du coefficient de tendance géné-

Augmented Dickey-Fuller Unit Root Test on RAS_EL_OUED				
Test Critical values:	1% level	-4.211000		
	5% level	-3.529758		
	10% level	-3.196411		
*Mackinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(RAS_EL_OUED)				
Method: Least Squares				
Date: 04/18/23 Time: 13:53				
Sample (adjusted): 10 48				
Included observations: 39 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
RAS_EL_OUED(-1)	-2.104650	0.809672	-2.599386	0.0147
D(RAS_EL_OUED(-1))	0.913773	0.698796	1.307639	0.2016
D(RAS_EL_OUED(-2))	0.676665	0.598191	1.131185	0.2676
D(RAS_EL_OUED(-3))	1.098073	0.501061	2.191498	0.0369
D(RAS_EL_OUED(-4))	0.878841	0.470272	1.868794	0.0721
D(RAS_EL_OUED(-5))	0.728485	0.435111	1.674253	0.1052
D(RAS_EL_OUED(-6))	0.272059	0.409658	0.664112	0.5121
D(RAS_EL_OUED(-7))	-0.020630	0.301389	-0.068449	0.9459
D(RAS_EL_OUED(-8))	-0.343832	0.188948	-1.819716	0.0795
C	7517421.	2942970.	2.554366	0.0164
@TREND("1")	28872.21	17798.47	1.622173	0.1160

FIGURE 3.13: Dickey-Fuller test

rale (0,11) est supérieure à (0,05), et cela indique qu'il n'y a pas de tendance générale dans la série, et à travers elle nous rejetons le nul hypothèse que la série a une tendance générale et accepter l'hypothèse alternative que la série n'a pas de tendance générale et donc qu'elle est stationnaire.

3.5 Conversion d'un série non stationnaire en une série stationnaire

En analysant les résultats précédents des deux séries, nous avons conclu que la série chronologique de consommation d'énergie pour la commune de Ras El Oued est stable, tandis que la série pour la commune de Bordj Bou Arreridj est instable. Nous ne travaillerons donc que sur la stabilité de ce dernier. Pour résoudre ce problème, nous nous appuyerons sur la prise de (d) des différences logarithmiques de la série jusqu'à ce qu'il soit stable, ce qui entraînera une diminution du nombre d'observations enregistrées, Le travail est automatique dans le programme Eviews comme suit :

1. **Processus de différence**

Nous avons fait la première différence selon l'équation suivante :

$$\ln(y_t) - \ln(y_{t-1}) = \delta \tag{3.1}$$

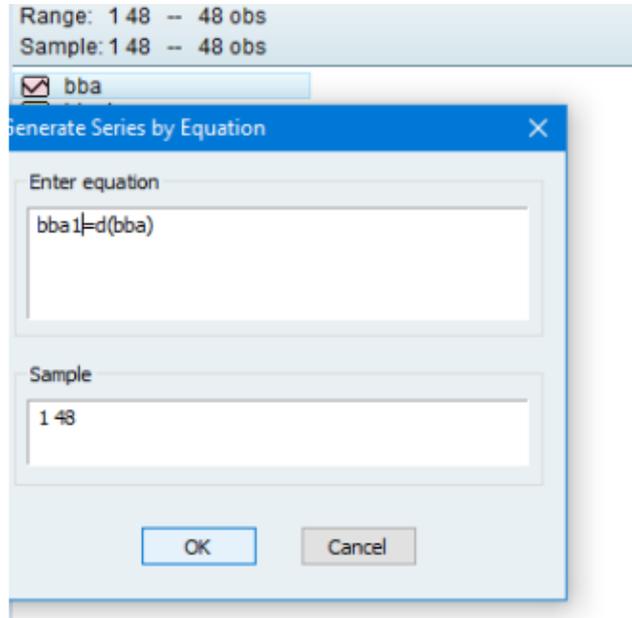


FIGURE 3.14: Processus de différence première

2. **Tester la stationnarité de la série obtenue**

Nous exécutons un Dickey-Fuller test et nous obtenons les résultats suivants

Augmented Dickey-Fuller Unit Root Test on BBA1				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-5.646649	0.0002
Test critical values:				
	1% level		-4.211868	
	5% level		-3.529758	
	10% level		-3.196411	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(BBA1)				
Method: Least Squares				
Date: 04/18/23 Time: 18:16				
Sample (adjusted): 10 48				
Included observations: 39 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
BBA1(-1)	-3.076816	0.544892	-5.646649	0.0000
D(BBA1(-1))	1.879853	0.464901	4.043553	0.0004
D(BBA1(-2))	1.673184	0.410297	4.077979	0.0003
D(BBA1(-3))	1.716372	0.338525	5.070145	0.0000
D(BBA1(-4))	1.143928	0.316651	3.612578	0.0011
D(BBA1(-5))	0.863540	0.289807	2.979703	0.0058
D(BBA1(-6))	0.783216	0.219966	3.560629	0.0013
D(BBA1(-7))	0.518875	0.156224	3.321361	0.0024
C	-0.046640	0.062706	-0.743789	0.4630
@TREND("1")	0.001887	0.002086	0.904734	0.3731

FIGURE 3.15: Dickey-Fuller test

3. La série obtenue est la suivante

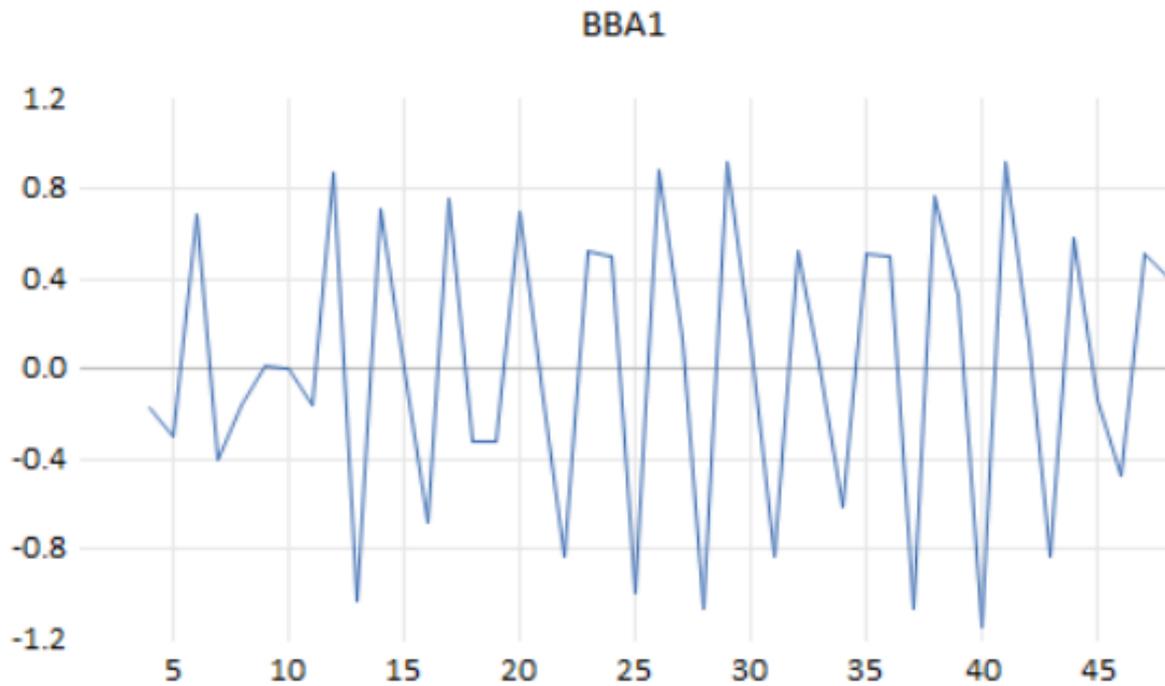


FIGURE 3.16: La nouvelle série de valeurs de consommation d'énergie pour la commune de Bordj Bou Arreridj

4. Certaines des valeurs obtenues sont les suivants

1	NA			
2	-0.172439			
3	0.044826			
4	0.092585			
5	-0.166851			
6	0.259646			
7	0.279106			
8	0.140550			
9	0.012879			
10	-0.118439			
11	-0.413893			
12	0.163637			
13	-0.288405			
14	-0.031368			
15	0.220968			
16	-0.207293			
17	0.118048			

FIGURE 3.17: Quelques chiffres de nouvelles données pour les valeurs de consommation d'énergie pour la commune de Bordj Bou Arreridj

Nous remarquons que nous n'avons perdu qu'une seule scène dans le processus de diff. La série est devenue stationnaires et a pu être exploitée pour la prédiction.

CONCLUSION

Dans ce chapitre, nous avons mené une étude des données de consommation d'énergie à l'aide du logiciel EVIEWS, en exécutant plusieurs tests pour tester la stabilité de chacune des deux séries en utilisant plusieurs méthodes bien connues, dont le test de Dickey-Fuller, Convertissez ensuite la série chronologique instable en une série chronologique stable à l'aide du même logiciel, Ceci est basé sur la méthode des différences logarithmiques, Dans le chapitre suivant, nous utilisons ces données obtenues comme paramètres pour les modèles LSTM et NARX, sur lesquels nous baserons pour faire la prédiction.

Chapitre 4

Application

INTRUDUCTION

Ce dernier chapitre présent notre application de prédiction qui a été fait par deux modèles le premier NARAX (Nonlinear autoregressive network with exogenous inputs) qui est l'un des algorithmes les plus célèbres et les plus efficaces avec les séries chronologique dans les réseaux neuronaux et le deuxième est l'algorithme LSTM (Long-Short-Term-Memory) qui est également très bien connu dans le domaine de la prédiction de série chronologique et est l'un des modèles de l'apprentissage profonde. Enfin nous comparons l'efficacité de ces modèles avec les résultats obtenus.

4.1 Prédiction de la consommation d'énergie à l'aide d'un réseau de neurones non linéaire auto-régressif

4.1.1 Auto régressif non linéaire

Il s'agit d'un puissant réseau neuronal moteur récurrent, utilisé pour la prédiction de séries temporelles, il consiste donc en un retard temporel. Il est construit sur la base d'un réseau neuronal multicouche.

Il se compose de trois couches de cellules : une couche d'entrée, une couche cachée et une couche de sortie.

Ces couches sont connectées les unes aux autres par une alimentation électrique directe, et la sortie supplémentaire (externe) avec des connexions de rétroaction aux cellules de la couche d'entrée, cela conduira à une stabilité plus rapide du réseau car il utilise l'algorithme de rétro-propagation pour corriger les erreurs. L'équation de définition du modèle *NARX* est :

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-n_y), u(t-1), u(t-2), \dots, u(t-u)) \quad (4.1)$$

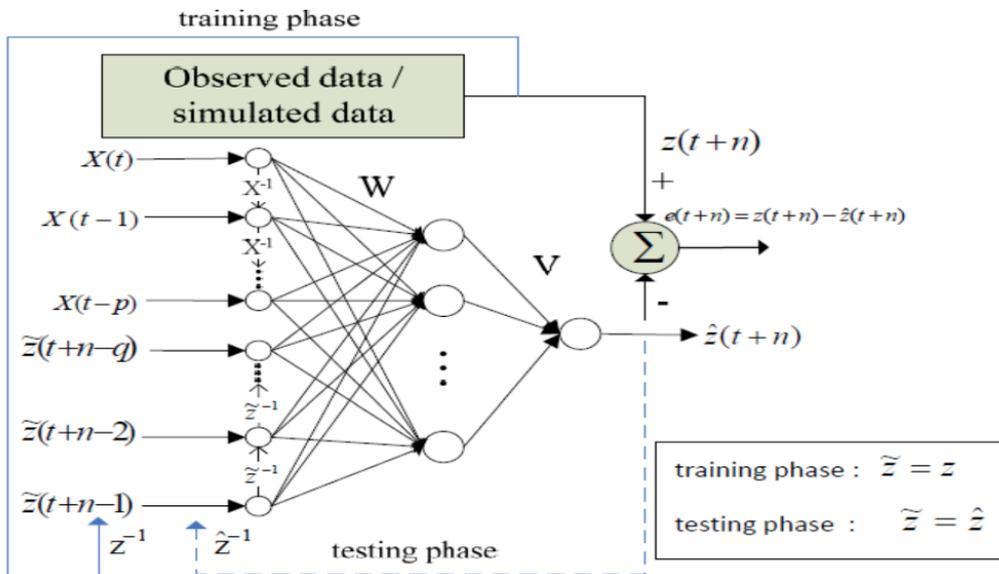


FIGURE 4.1: L'architecture du réseau de neurones artificiels *NARX*

4.1.2 Matériel et méthodes

— Structure des données :

Nous disposons de données mensuelles sur la consommation d'énergie de 2019 à 2022. Nous effectuons la normalisation de la loi suivante

$$X_{norm} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (4.2)$$

puis divisons les données en quatre parties et les organisons en fonction de la nature de l'algorithme.

— Méthodologie :

1. Application de réseau neurones artificiels (RNA) (Séries chronologiques de réseaux de neurones) , de Matlab R2015b.
2. L'autorégressif non linéaire avec entrée externe (NARX)

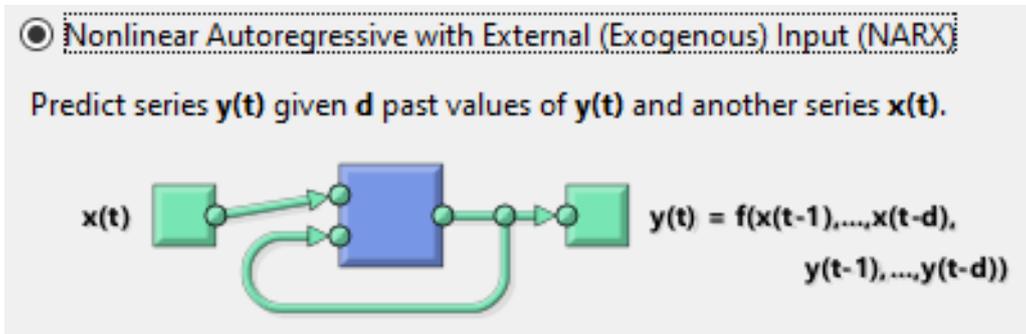


FIGURE 4.2: Réseau NARX dans un programme Matlab R2015b

3. Déterminer le pourcentage de chaque donnée (formation, validation, test) par rapport au total des données.
4. Ensuite, choisissez l'algorithme d'entraînement et à la fin, nous obtenons les résultats requis et le format de réseau spécifié.

4.2 Architecture des réseaux neurones

4.2.1 Conception des deux réseaux

Après plusieurs tests nous avons atteint les réseaux suivants

— **Réseaux 1**

Il représente le réseau de neurones des données de la commune de Bordj Bou Arreridj , est se présente comme suit :

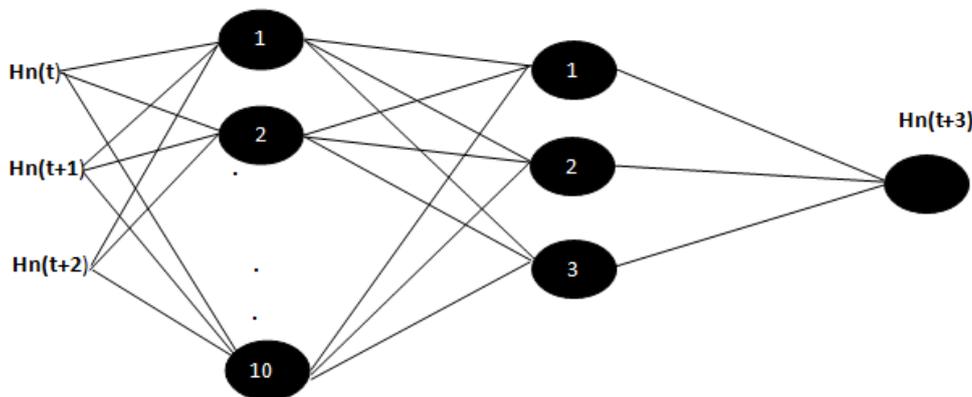


FIGURE 4.3: Réseau Commune de Bordj Bou Arreridj

— **Réseaux 2**

Il représente le réseau de neurones des données de la commune de Ras El Oued , est se présente comme suit :

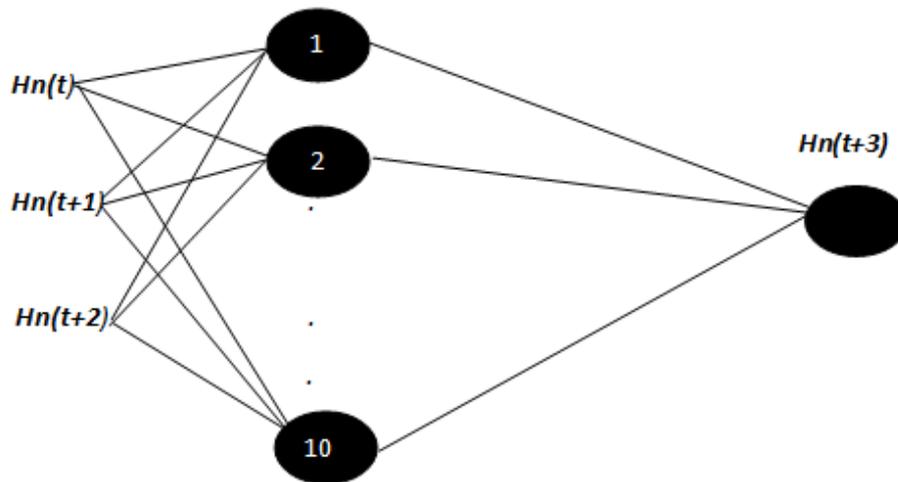


FIGURE 4.4: Réseau Commune de Ras El Oued

4.2.2 Discuter des deux réseaux

Dans le tableau suivant, nous montrons quelques-unes des décisions et des critères qui ont été prises dans le processus d'apprentissage au sein du programme **Matlab**, ainsi que l'erreur qui en résulte.

D'après le tableau ci-dessus, nous remarquons que :

Réseau	Structures	Algorithme d'entraînement	Pourcentage (Training\validation\Testing)	MSE (mean Squared erreur)
1	3-10-3-1	Scaled Conjugat Gradient	80/20	0,0096843
2	3-10-1	Levenberg - Marquardt	70/15/15	0,0063351

FIGURE 4.5: Les paramètres et résultats d'apprentissage

- L'erreur tend toujours vers zéro.
- Dans les deux réseaux, il y a trois entrées (T , $T+1$, $T+2$) et une sortie ($T+3$).
- Le premier réseau contient deux couches cachées, la première avec 10 neurones et la seconde avec 3 neurones.
- Le deuxième réseau contient une couche cachée avec 10 neurones.
- Nous nous sommes appuyés sur l'algorithme Scaled Conjugate Gradient dans le réseau 1.
- Dans le réseau 2, nous nous sommes appuyés sur l'algorithme de Levenberg-Marquardt.
- Dans les deux réseaux, nous sommes appuyés sur la fonction d'activation redSigmoide
- Les performances des deux réseaux de neurones artificiels ont été évaluées à l'aide de l'erreur quadratique moyenne (Mean Squared Error), le taux d'erreur absolue moyenne (Mean Absolute

Percentage Error) et de l'erreur absolue moyenne (Mean Absolute Error).

$$MSE = \frac{\sum_{i=1}^n (y_i - y_i')^2}{n} \quad (4.3)$$

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - y_i'}{y_i} \right|}{n} \quad (4.4)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - y_i'|}{n} \quad (4.5)$$

Où y est la valeur d'origine de la consommation, t : la valeur attendue, n est le nombre d'échantillons et i est la séquence de mois.

- Les premier et deuxième réseaux sont basés sur un réseau de type **Feed-forward backpropagation** qui ajuste les poids et les biais du réseau très rapidement.

4.3 Test de performance du réseau neurones artificiels

Pour connaître les performances du réseau de neurones artificiels, si l'entraînement est meilleur, il existe un ensemble de tests et de chiffres qui montrent les performances du réseau de neurones artificiels.

4.3.1 Évolution des carrés moyens des erreurs à travers les étapes de formation

La courbe de performance représente la relation entre l'évolution des carrés d'erreur lors des phases d'apprentissage (erreur et le nombre de tentatives (Epoche) pour les courbes (Training, validation, testing)).

1. Réseau 1

On remarque dans le graph ci-dessous que plus le nombre de tentatives d'apprentissage est grand, plus le nombre de carrés d'erreur (mes) est faible, et le dernier point de tentatives d'apprentissage se situe au point 209, et le meilleur point sur lequel le réseau de neurones artificiels a été entraîné est aussi le point 209.

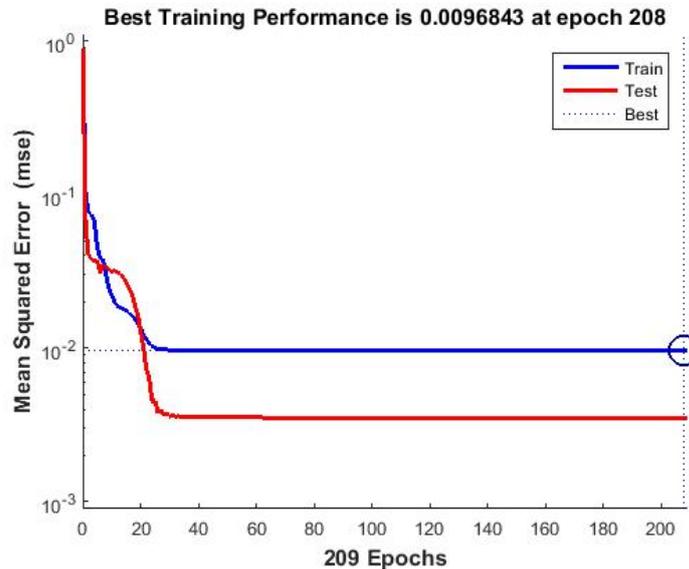


FIGURE 4.6: Evolution des carrés moyens des erreurs)

2. Réseau 2

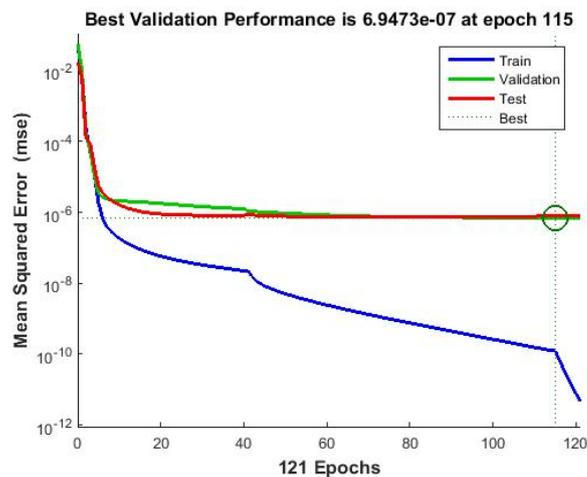


FIGURE 4.7: Evolution des carrés moyens des erreurs

L'erreur (mes) diminue avec les tentatives d'entraînement pour atteindre 121 au dernier point tenté de l'entraînement, et le meilleur point sur lequel le réseau de neurones s'est entraîné est de 115 pour toutes les courbes (Training, validation, testing).

4.3.2 Corréler la sortie du réseau neuronal avec les valeurs réelles

Ce diagramme montre la relation entre la corrélation de la sortie du réseau de neurones avec les valeurs réelles pour chacune des courbes (Training, validation, testing).

Et comme on sait précédemment que la corrélation des valeurs réelles est confinée entre 0 et 1, et dès qu'elle est proche de 0, cette corrélation est faible, et dès qu'elle est proche de 1, cette corrélation est forte, et si elle est égal à 1, alors cette corrélation est complète.

Ce test est l'un des meilleurs tests pour savoir si le réseau neuronal est mieux formé sur la façon d'utiliser la corrélation des valeurs réelles.

1. Réseau 1

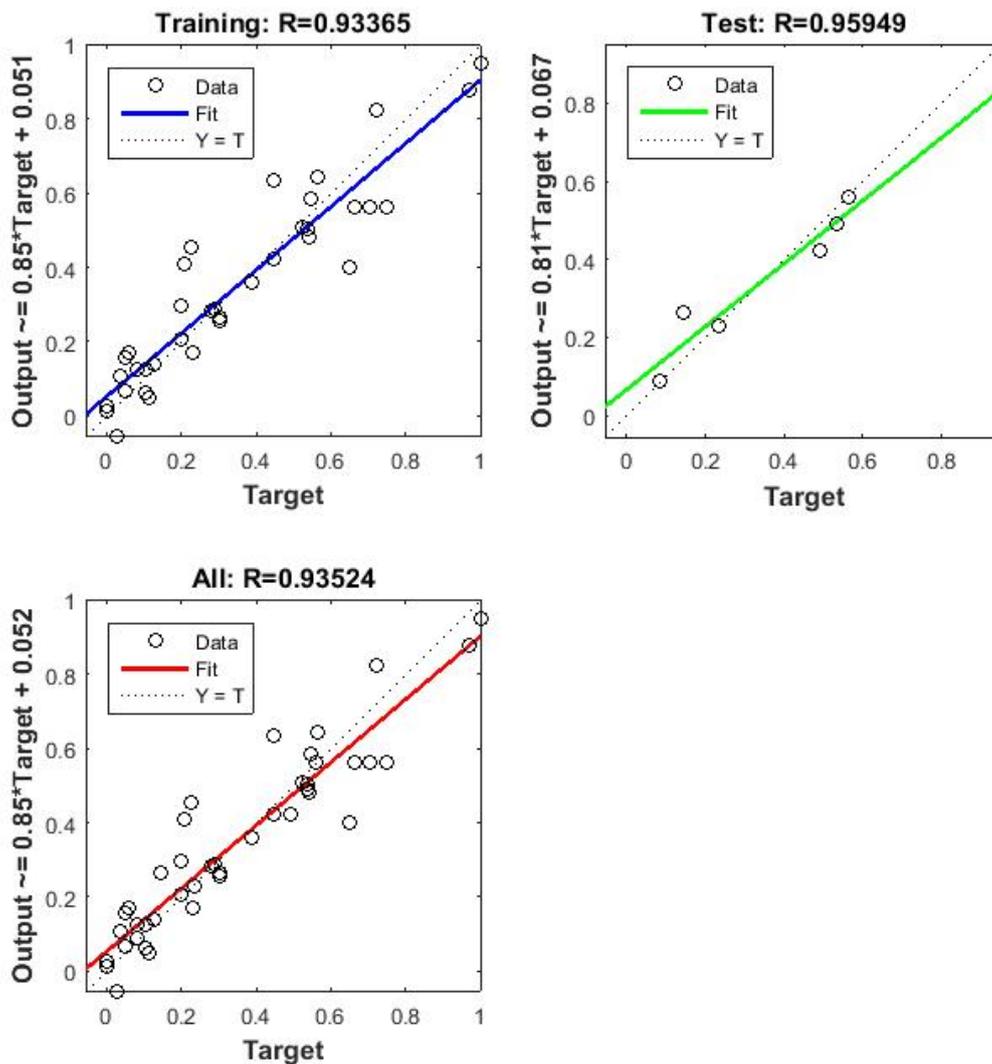


FIGURE 4.8: Illustration des lignes de régression et des valeurs estimées et réelles

D'après les graphiques, nous remarquons le rassemblement de points qui représentent les valeurs réelles sur les droites de régression montrant une forte corrélation proche de un, c'est-à-dire avec une valeur de 0,9.

2. Réseau 2

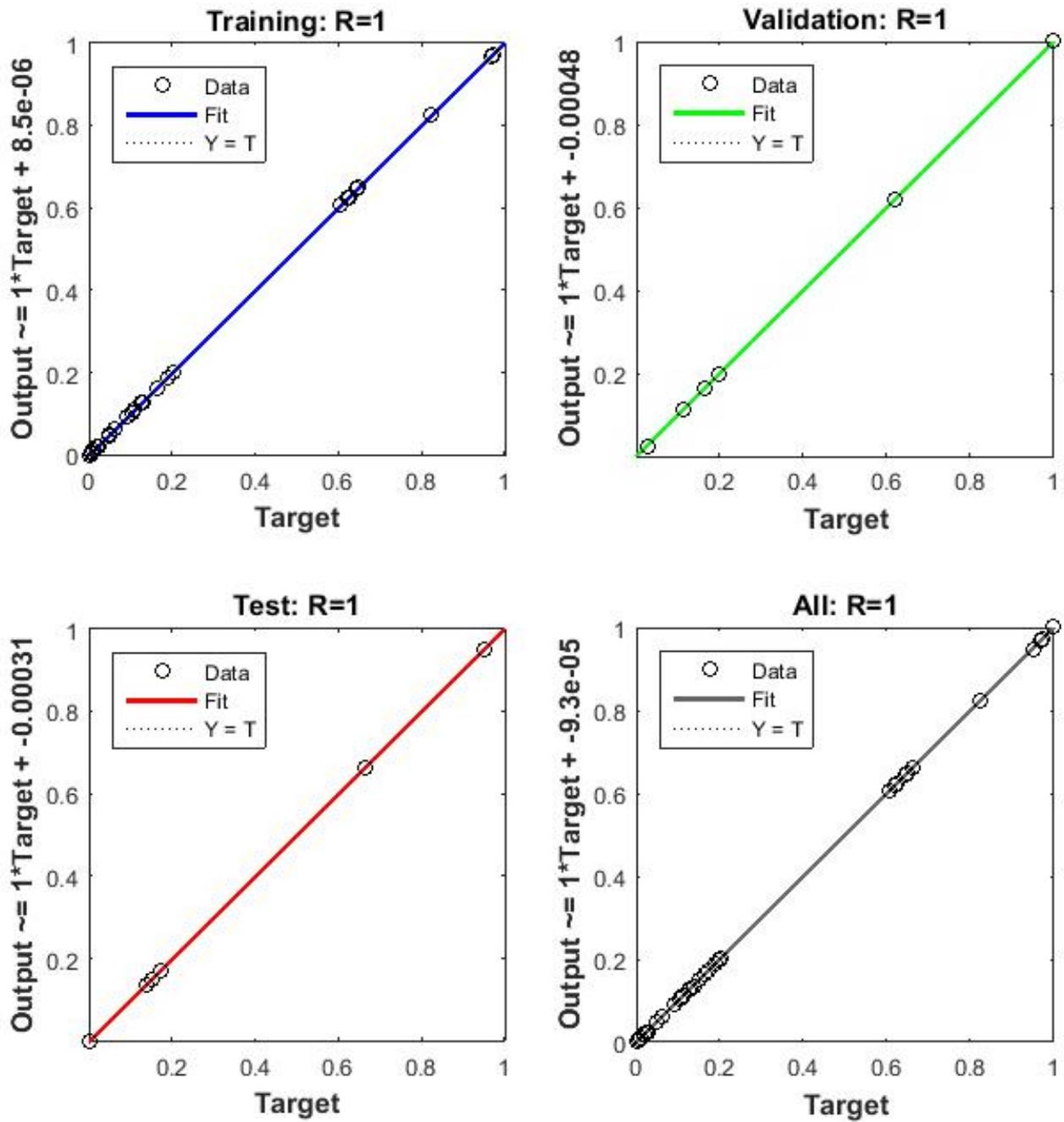


FIGURE 4.9: Illustration des lignes de régression et des valeurs estimées et réelles

A travers les courbes suivantes pour la formation, l'essai et la vérification, nous remarquons la combinaison de points représentant les valeurs réelles sur les lignes de régression montre qu'il y a une corrélation forte ou complète où la valeur a atteint 1.

4.3.3 Représentations graphiques de la sortie d'un réseau de neurones

1. Réseau 1

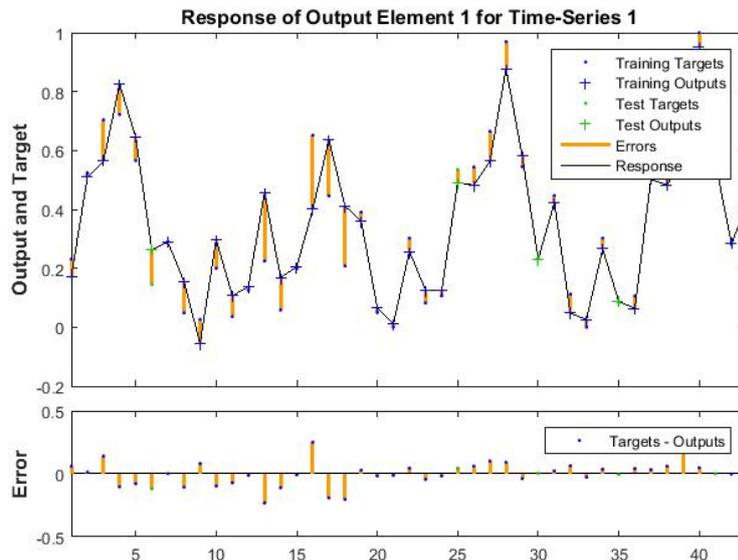


FIGURE 4.10: La sortie du réseau de neurones

Le graphe montre la réponse des éléments avec les sorties de la série temporelle en fonction du nombre d'observations.

On peut voir que la valeur d'erreur entre (objectifs) et (résultats) est très faible dans toutes les données sauf le mois 16 qui a une augmentation significative de l'erreur ainsi que le mois 14, mais en général le champ d'erreur est assez petit et acceptable, ce qui indique que le réseau de neurones synthétique pourrait bien s'entraîner.

2. Réseau 2

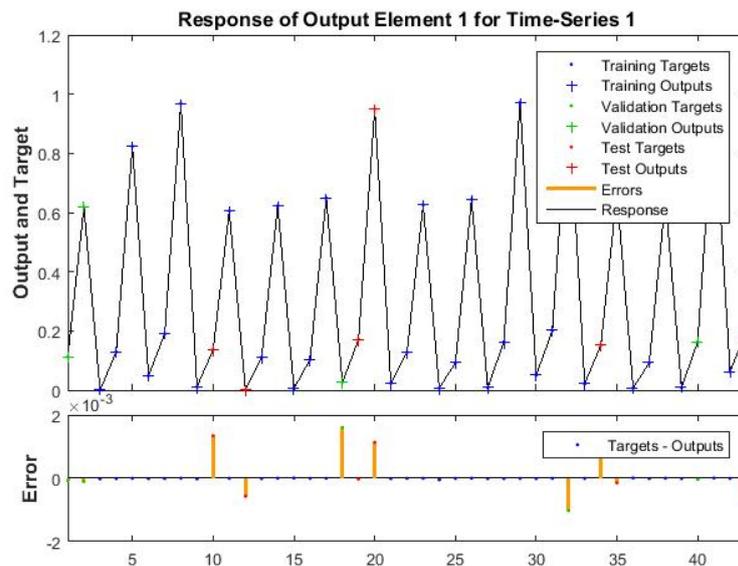


FIGURE 4.11: La sortie du réseau de neurones

De la déclaration, nous concluons que le réseau de neurones artificiels est bien formé. Comme nous pouvons le voir, la valeur d'erreur entre les cibles et les sorties est généralement très faible, proche de zéro. Ce n'est qu'au 18e mois que la valeur d'erreur la plus élevée a été observée, puis elle a de nouveau diminué, et à partir de là, nous pouvons dire que l'erreur diminue à mesure que le nombre de vues augmente.

4.4 Les résultats de modèle NARX

Après avoir construit, formé et validé le réseau de neurones, nous pouvons maintenant l'utiliser pour prédire les valeurs mensuelles de consommation d'énergie et les comparer avec les valeurs réelles obtenues de Sonelgaz.

Afin d'obtenir les résultats de prédiction, nous nous sommes appuyés sur le programme Matlab R2015b après avoir formé le réseau et vérifié leurs compétences dans le même programme, et les résultats étaient les suivants :

— Réseau 1

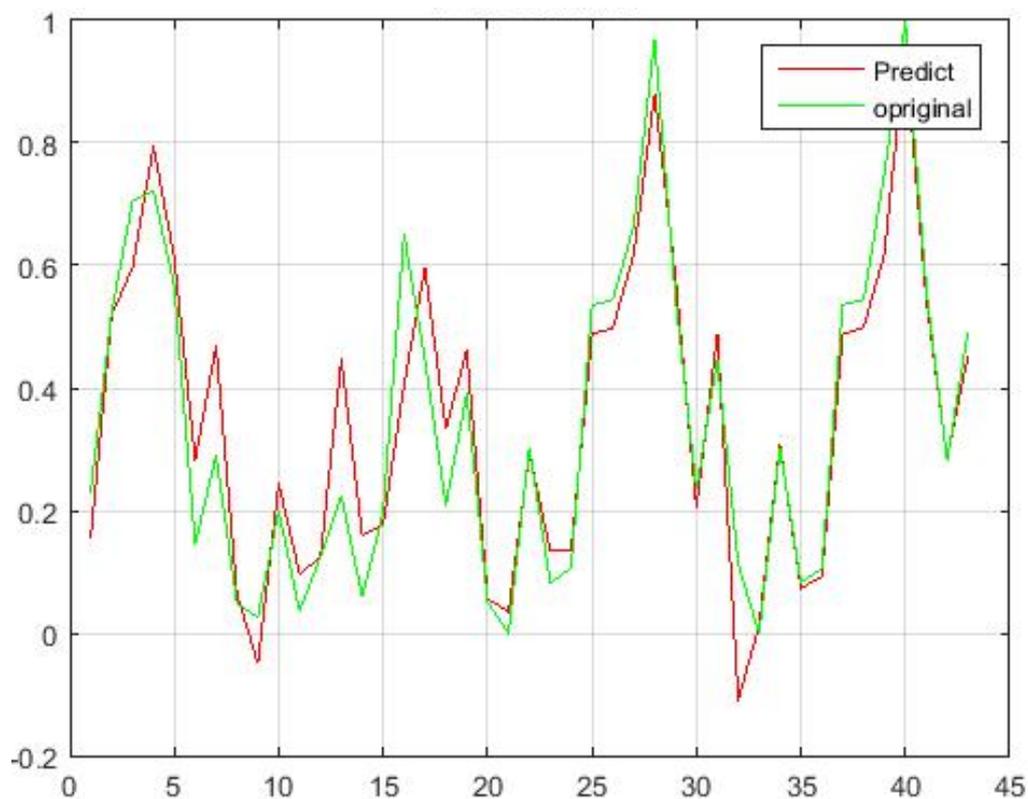


FIGURE 4.12: Comparaison des consommations énergétiques réelles et prévisionnelles de la commune de Bordj Bou Arreridj

Discussions

Les valeurs de prédiction obtenues à partir du réseau de neurones entraîné (**NARX**) ne sont très proches des valeurs réelles qu'à certains points, on remarque une augmentation de la valeur d'erreur entre la valeur attendue et vraies valeurs, mais ce n'est pas très important car ce n'est pas loin de la valeur réelle. Cela indique que le réseau a été bien formé et a donné des résultats

satisfaisants malgré le petit nombre d'observations dans la base de données.

— Réseau 2

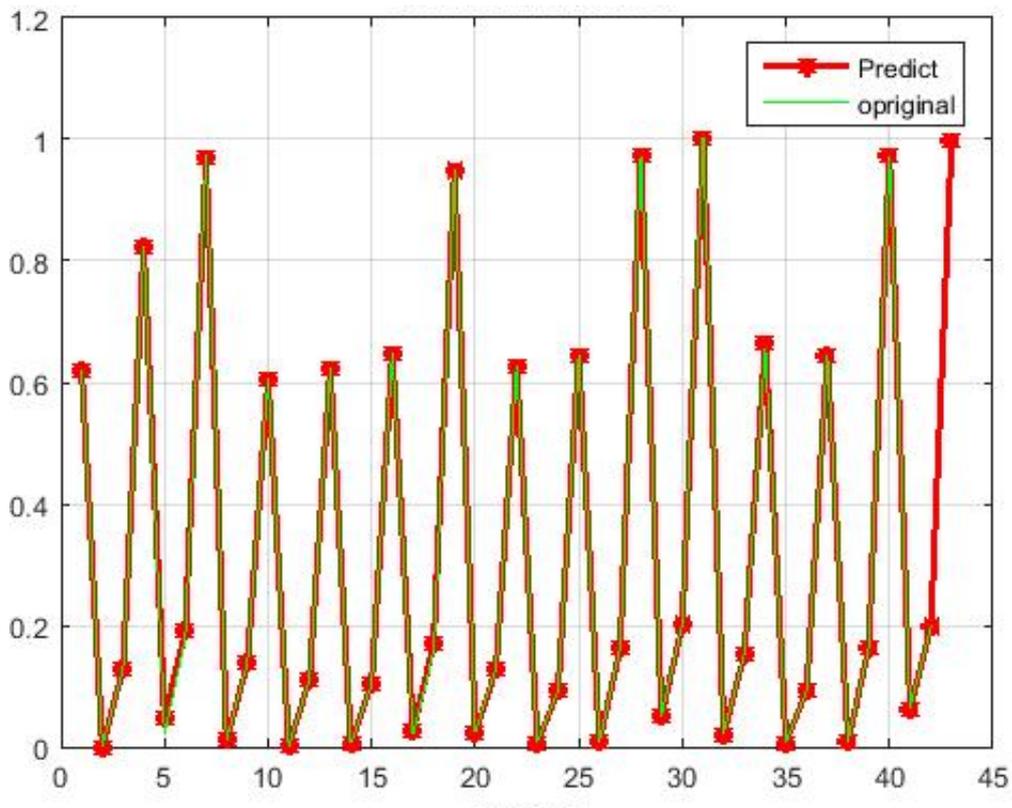


FIGURE 4.13: Comparaison des consommations énergétiques réelles et prévisionnelles de la commune de Ras El Oued

Discussions

Les valeurs de prédiction obtenues à partir du réseau neuronal formé (NARX) sont presque identiques aux valeurs réelles des valeurs de consommation d'énergie pour la municipalité de Ras El ouad, et cela est une preuve claire que le réseau neuronal a été formé très bien et les résultats sont largement satisfaisants.

4.5 Prédire la consommation d'énergie à l'aide du réseau de neurones profonds LSTM

Le réseau de neurones profonds LSTM, sur lequel nous nous appuyerons dans la suite, pour prédire les séries chronologiques de consommation d'énergie pour chacune des deux communes de Bordj Bou Arreridj et Ras El Oued, à partir de la même base de données obtenue pendant la période de formation dans le Société Sonelgaz, qui a été traitée dans le programme Eviews, puis nous avons fait le processus de normalisation sur ces données. Nous expliquerons comment fonctionne ce réseau avancé à travers un ensemble d'images qui expliquent mieux ce processus. qui est le suivant :

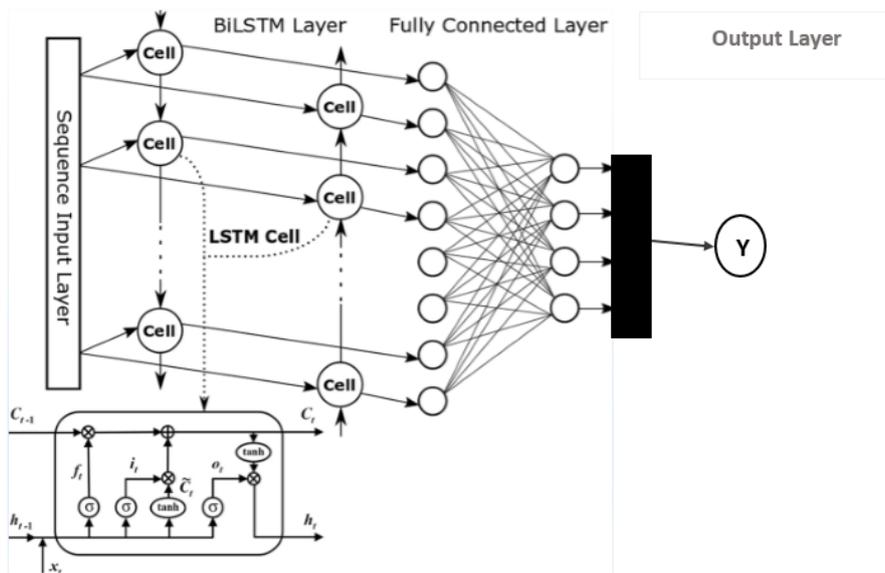


FIGURE 4.14: Long Short-Term Memory Model

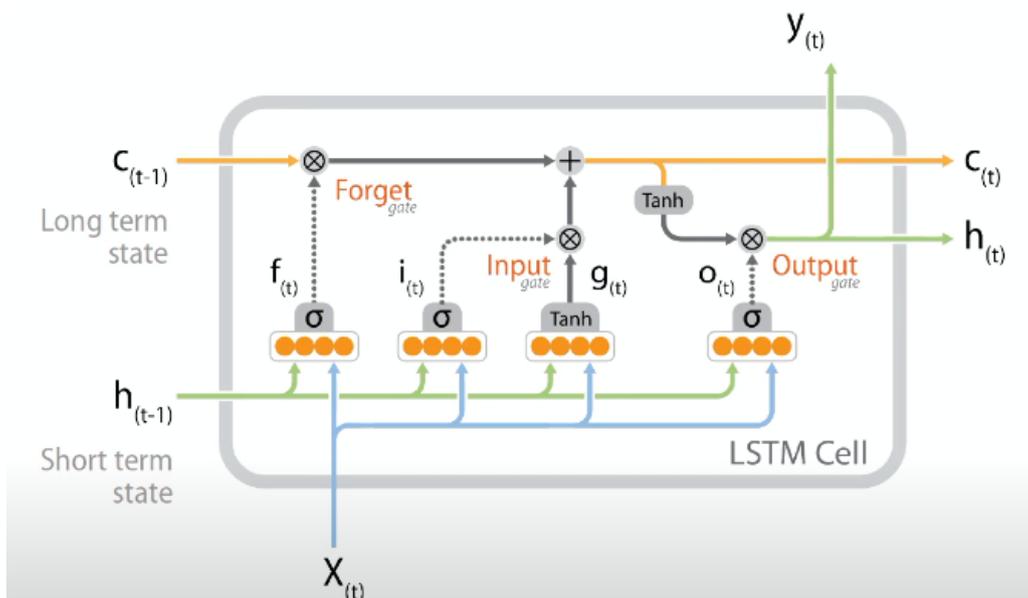


FIGURE 4.15: LONG SHORT-TERM MEMORY [36] (LSTM cell)

Avec :

$x_t \in R^d$	input vector
$f_t \in R^h$	forget gate's activation vector
$i_t \in R^h$	input gate's activation vector
$o_t \in R^h$	output gate's activation vector
$g_t \in R^h$	current entry vector
$h_t, y_t \in R^h$	hidden state or output vector
$c_t \in R^h$	cell stat vector
\otimes	Hadamard product
σ	Sigmoid function
w_k	weights matrix
b_k	bais vector

$$f(t) = \sigma(W_{xf}^T X(t) + W_{hf}^T h_{(t-1)} + b_f) \quad (4.6)$$

$$i(t) = \sigma(W_{xi}^T X(t) + W_{hi}^T h_{(t-1)} + b_i) \quad (4.7)$$

$$g(t) = \tanh(W_{xg}^T X(t) + W_{hg}^T h_{(t-1)} + b_g) \quad (4.8)$$

$$o(t) = \sigma(W_{xo}^T X(t) + W_{ho}^T h_{(t-1)} + b_o) \quad (4.9)$$

$$c(t) = f(t) \otimes c_{(t-1)} + i(t) \otimes g(t) \quad (4.10)$$

$$y(t) = h(t) = o(t) \otimes \tanh(c(t)) \quad (4.11)$$

4.6 Résultats de l'algorithme LSTM pour les données de consommation d'énergie pour chacune des deux communes

Au lieu d'utiliser les initialisations de poids et de biais par défaut, ce modèle utilise des initialisations pour s'assurer que la variance du gradient de poids entre les couches reste constante. La fonction d'activation passe de ReLU à la fonction hyperbolique tangente (tanh) pour éviter le phénomène de gradient explosif observé lors de la phase d'apprentissage.

L'erreur quadratique moyenne est utilisée comme fonction de perte et comme technique de régularisation L2 pour éviter un éventuel phénomène de surajustement du modèle (Ir = 0,01, époque = 250 dans le réseau 1 et , époque=180 Dans le réseau 2).

Dans cet travail , nous choisissons Stochastic Gradient Descent (SGD) et Adaptive Moment Estimation (Adam) comme optimiseurs pour mon modèle LSTM.et Dans les deux réseaux 1 et 2, nous avons 3 entrées et 1 sortie.

Dans les résultats, nous avons obtenu des valeurs RMSE et MAPE faibles lors de l'utilisation de

l'optimiseur Adam. Par conséquent, le modèle LSTM amélioré par Adam a une bonne précision et est plus raisonnable.

Nous avons utilisé le programme MATLAB 2023 pour compléter cet algorithme avec 48 observations pour chaque pale sur 4 ans, et les résultats obtenus sont les suivants :

4.6.1 Réseau 1

— Corréler la sortie du réseau neuronal avec les valeurs réelles

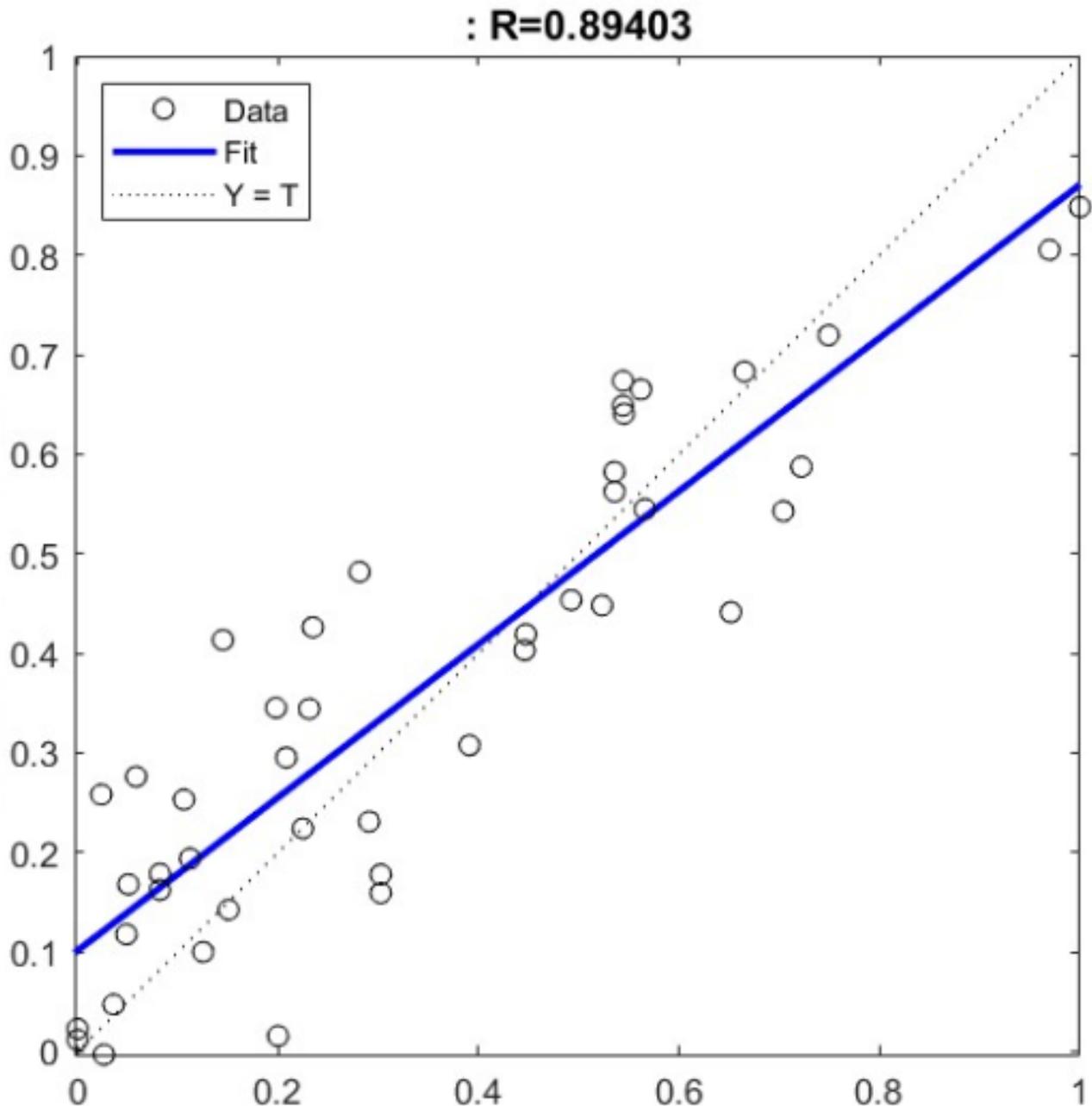


FIGURE 4.16: Diagramme schématique de la ligne de régression

D'après les graphiques, nous remarquons que l'ensemble des points représentant les valeurs réelles sur la droite de régression montre une forte corrélation avec une valeur de 0,8.

— Résultat de prédiction

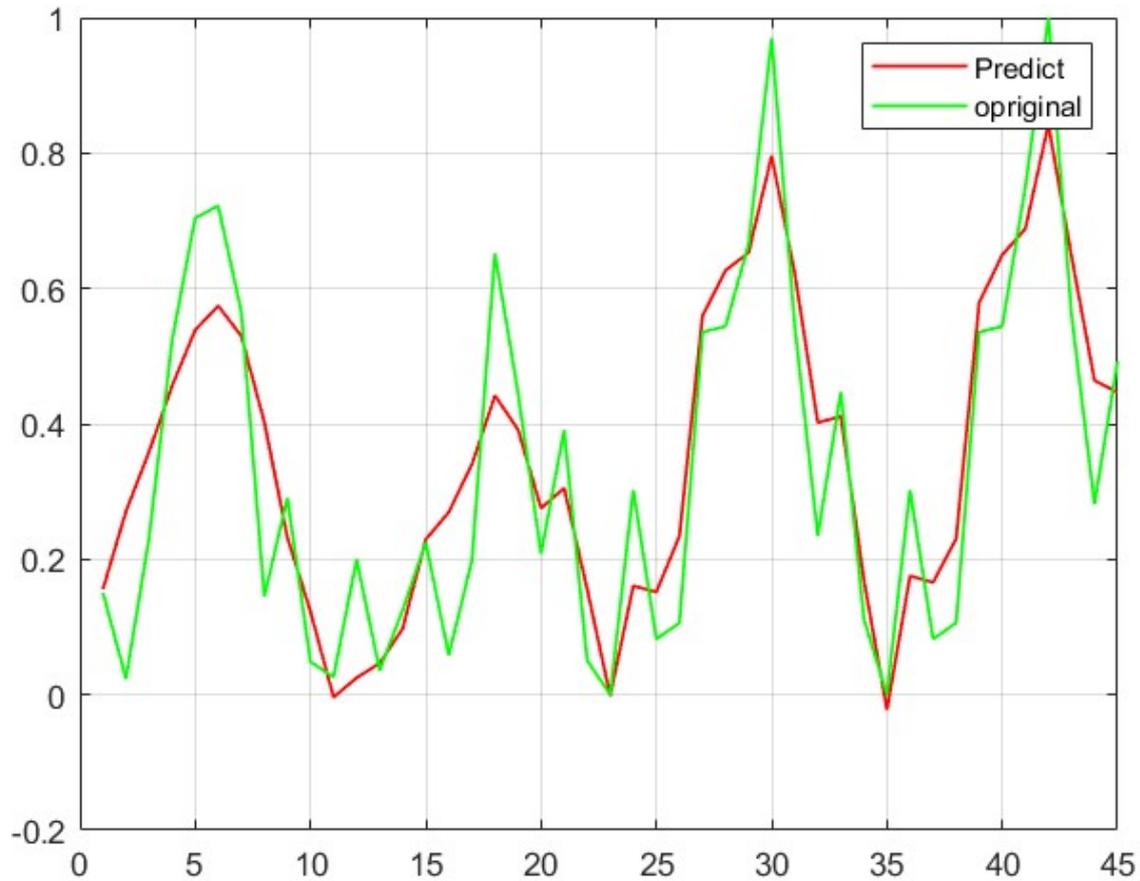


FIGURE 4.17: Comparaison des consommations énergétiques réelles et prévisionnelles de la commune de Bordj Bou Arreridj

Discussions

Les valeurs de prédiction obtenues à partir du réseau de neurones profonds (*LSTM*) sont quelque peu proches des valeurs réelles, ce qui indique que le réseau est bien formé et a donné des résultats satisfaisants malgré le petit nombre d'observations dans la base de données.

4.6.2 Réseau 2

— Corréler la sortie du réseau neuronal avec les valeurs réelles

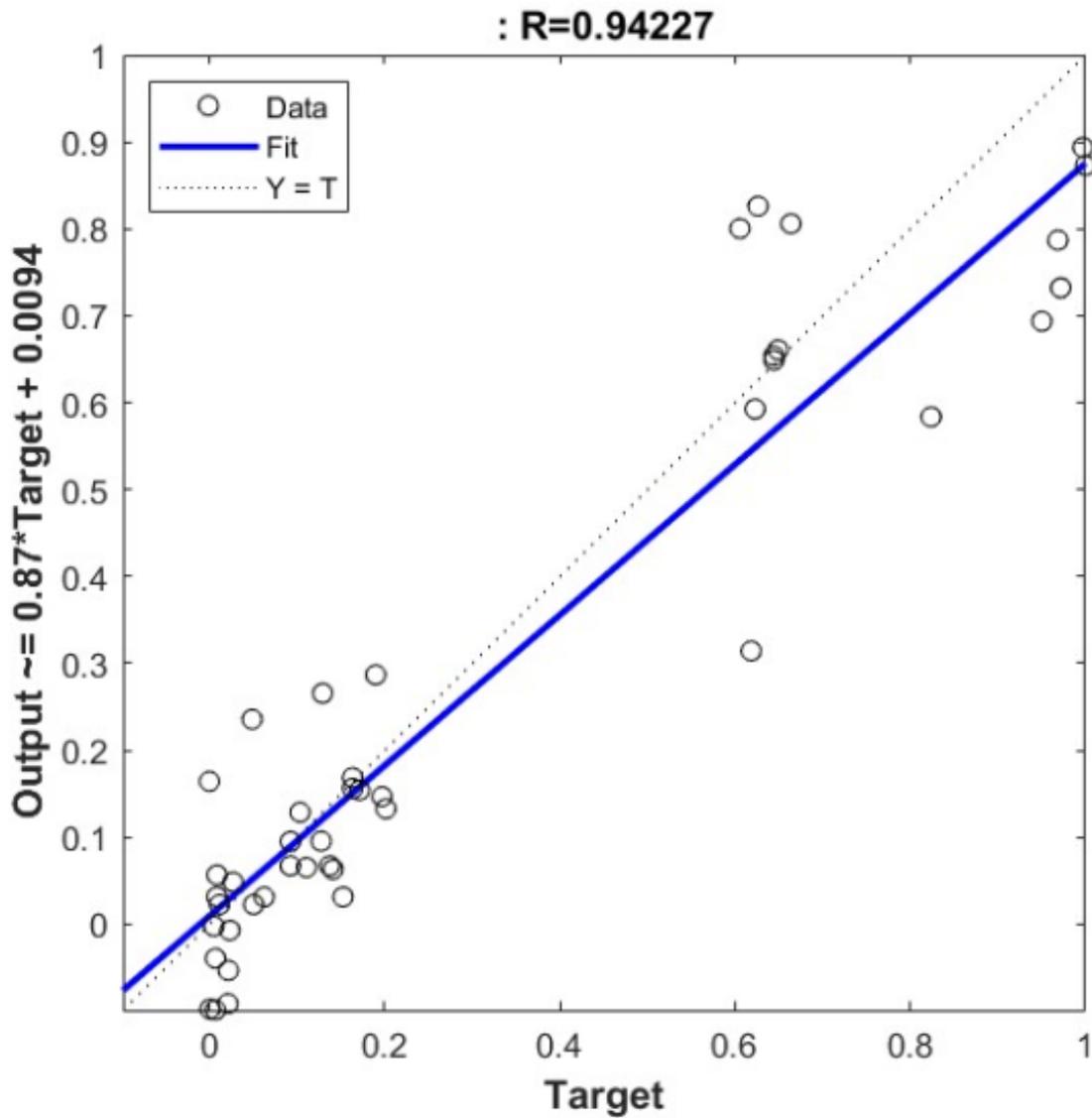


FIGURE 4.18: Diagramme schématique de la ligne de régression (Ras El Oued)

A travers le graphique, on remarque que le groupe de points représentant les valeurs réelles sur la droite de régression montre une corrélation très, très forte de 0,94, et c'est un résultat très impressionnant et satisfaisant dans une large mesure.

— Résultat de prédiction

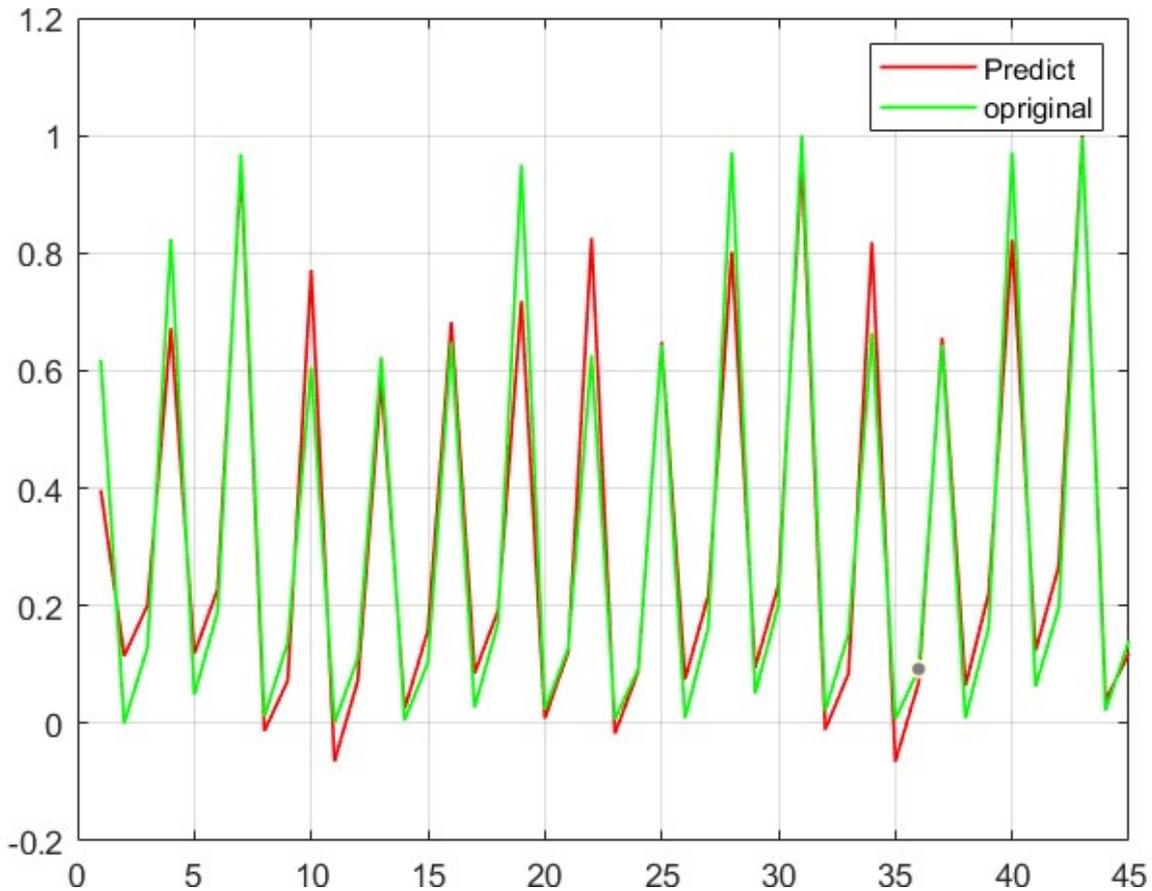


FIGURE 4.19: Comparaison des consommations énergétiques réelles et prévisionnelles de la commune de Ras El Oued

Discussions

Les valeurs de prédiction obtenues à partir du réseau de neurones profonds formés (LSTM) sont identiques avec un pourcentage très élevé avec les valeurs réelles des valeurs de consommation d'énergie pour la municipalité de Ras El-Oued, et cela est une preuve claire que le réseau de neurones a été très bien formé et les résultats sont largement satisfaisants.

Remarque

Tous les résultats de prédiction que nous avons obtenus à partir des modèles *NARX* et *LSTM* sont normalisés, cela explique l'existence de valeurs négative.

4.7 Comparaison entre les deux réseaux

Les deux modèles de réseaux de neurones utilisés, *NARX* non linéaire (pour le machine learning) et deep LSTM (pour le deep learning), fournissent d'excellents résultats de prédiction. D'après les résultats de cette étude, il est également conclu que le réseau neuronal *NARX* est plus efficace que les résultats LSTM et fournit des résultats relativement meilleurs. Malgré l'attente inverse (les LSTM donnent de meilleurs résultats avec de très grandes quantités de données).

CONCLUSION

À la fin de ce chapitre, nous avons appris comment traiter les algorithmes d'intelligence artificielle NARX (Nonlinear autoregressive exogenous model) et LSTM (Long-Short-Term-Memory), et nous avons pu obtenir des résultats de prédiction pour chacune des deux séries temporelles, et atteindre l'objectif de l'étude dans ce sujet. Où nous avons atteint une prédiction acceptable dans une large mesure avec le moins de champ d'erreur possible, Nous en concluons l'efficacité et la capacité de l'intelligence artificielle dans la prévision et sa concurrence avec les méthodes de prédiction statistique.

CONCLUSION GÉNÉRAL

De nos jours, les systèmes de gestion et de traitement des données sont censés stocker et traiter de grandes séries temporelles. Une série temporelle est une suite de valeurs représentant l'évolution d'une quantité spécifique au cours du temps. De telles suites peuvent être exprimées mathématiquement afin d'en analyser le comportement, généralement pour comprendre son évolution passée et pour en prévoir le comportement futur.

Plusieurs méthodes de prévision sont développées et utilisées : Méthodes statistiques, systèmes experts, logique floue, approche connexionnistes ...etc. Leurs principales spécificités par rapport aux domaines les plus courants de l'apprentissage automatique sont leur dépendance dans le temps et leurs comportements saisonniers pouvant apparaître dans leur évolution. Dans le cas de mesures du taux de change monétaire par exemple, intuitivement, nous comprenons que la valeur de demain sera affectée par la séquence des jours passés. On dit de ces données qu'elles ne sont donc pas indépendantes.

Notre but était d'implémenter deux techniques de prédiction utilisant les réseaux de neurones artificiels de type Perceptron multi couches et les réseaux de neurones convolutifs utilisés dans le domaine de deep learning, afin de réaliser une application de prévision à court terme de la charge électrique au sein de la wilaya de Bordj bouarreridj.

Dans ce travail, nous avons étudié différents modèles de prédiction des séries temporelles. Nous avons discuté mathématiquement leurs principes, puis nous avons distingué leurs avantages et inconvénients. Nous avons commencé par les modèles linéaires auto-régressifs uni-variés et multi-variés, qui sont parmi les plus utilisés dans la littérature, et quelques modèles basés sur les réseaux de neurones artificiels, qui sont de plus en plus utilisés dans la prédiction des séries temporelles.

Pour la démonstration et la validation du système, nous avons choisi comme étude de cas, la prédiction de la consommation de l'énergie au sein de la wilaya de bordj bouarreridj et en particulier les communes de bordj et Ras eloued. D'après les tests effectués, il est nettement observable que les résultats obtenus sur la base de validation sont très satisfaisants du fait que les valeurs prédites sont très proches des valeurs réelles, ce qui atteste l'exactitude et l'adéquation de la solution proposée.

En guise de conclusion, nous arrivons aux conclusions suivantes :

- La stationnarité est un facteur très important, d'après notre étude, la série de la commune de Ras El Oued est plus stationnaire que celle de la commune de Bordj Bou Arreridj. Ce qui explique que les résultats des prévisions pour la commune de Ras El Oued sont meilleurs que ceux de la commune de Bordj Bou Arreridj.
- Les deux modèles de réseaux de neurones utilisés, le NARX non linéaire (pour la machine learning) et le LSTM profond (pour le deep learning), ont fourni d'excellents résultats de prédiction.
- Il a également été conclu à partir des résultats de cette recherche que le réseau de neurones NARX est plus efficace et fournit des résultats relativement meilleurs que les résultats du LSTM. Bien qu'attendez le contraire (LSTM donne de meilleurs résultats dans le cas d'une très grande masse de données).

En perspectives, ce travail peut être amélioré ainsi :

- Considérer une large masse de données afin d'avoir de bons résultats utilisant le deep learning.
- Lier les différents réseaux développés à un système accessible sur le web afin de consulter en temps réel les besoins de l'énergie électriques.

Bibliographie

- [1] M. Selim, R. Zhou, W. Feng, and P. Quinsey, "Estimating Energy Forecasting Uncertainty for Reliable AI Autonomous Smart Grid Design," *Energies* .2021
- [2] M. Q. Raza, N. Mithulananthan, J. Li, and K. Y. Lee, "Multivariate Ensemble Forecast Framework for Demand Prediction of Anomalous Days," *IEEE Trans. Sustain. Energy* .2020
- [3] A. Gupta and P. K. Sarangi, "Electrical load forecasting using genetic algorithm based backpropagation method," *ARPN J. Eng. Appl. Sci* .2012
- [4] M. Lekshmi and K. N. Adithya Subramanya, "Short-term load forecasting of 400kV grid substation using R-tool and study of influence of ambient temperature on the forecasted load," *2019 2nd Int. Conf. Adv. Comput. Commun. Paradig*.2019
- [5] Fabrice Mazerolle - Statistique descriptive - Série statistique à une et deux variables, séries chronologiques, indices -Gualino Editeur (2005)
- [6] Fabrice Mazerolle - Statistique descriptive - Série statistique à une et deux variables, séries chronologiques, indices -Gualino Editeur (2005)
- [7] <https://www.aquiladata.fr/insights/mieux-comprendre-les-methodes-de-prevision-des-series-chronologiques/> (visité au 23/02/2023)
- [8] <https://blog.insee.fr/prix-de-lenergie-carburants-gaz-electricite-un-tierce-en-hausse/> (visité au 25/02/2023)
- [9] <http://www.jybaudot.fr/Previsions/trend.html> (visité au 04/03/2023)
- [10] <https://www.i3s.unice.fr/crescenz/publications/Florence/introduction-series-chronologiques>. (visité au 10/03/2023)
- [11] Hemsas Ouahiba , Prédiction dans les Séries Chronologiques , Faculté de Mathématiques , Département Mathématiques appliquées, Mémoire de Master , Année 2014
- [12] N.D. Lewis - *Neural Networks for Time Series Forecasting with R* (2017, N.D. Lewis).
- [13] LOUDJANI Nawal, Développement d'algorithmes pour l'analyse des Séries temporelles des données de production d'eau potable , Université Mohamed Khider – Biskra , Faculté de Sciences et de la technologie, Département Génie civil et hydraulique , Thèse Doctorat en

Sciences.

[14] Ahmed BENSALMA , Introduction aux tests de racine unitaire, PÙle Universitaire de Koléa , Tipaza, Algérie , école Nationale Supérieure de Statistique et d'économie Appliquée

[15] Beida Mohammed , Ferhat Taleb Amar ,les outils de gestion Prévision

عرباوي عبد الستار، التنبؤ بدرجات الحرارة لولاية ورقلة باستخدام طريقة بوكس- جنكيز للسلاسل الزمنية للفترة، كلية الرياضيات و علوم المادة، اختصاص احتمالات و إحصاء

[16]

الدكتور مصطفى شعراوي، مقدمة في التحليل الحديث للسلاسل الزمنية، كلية العلوم جامعة الملك عبد العزيز المملكة العربية السعودية

[17]

[18] <https://fr.linkedin.com/pulse/les-types-de-machine-learning-william-simetin-grenon> (visité au 15/03/2023)

[19] OTMANI IMANE , L'ANALYSE DISCRIMINANTE ET LE PERCEPTRON MULTICOUCHE , UNIVERSITE MENTOURI CONSTANTIN , FACULTE DES SCIENCES EXACTES , DEPARTEMENT DE MATHEMATIQUE

[20] Karima BENLEMBAREK "Systèmes de Prédiction Basée Web Applications aux Séries Temporelles",Faculté des Sciences de l'ingénieur,Département d'Informatique , diplôme de MAGISTER en (2010), UNIVERSITE BADJI MOKHTAR-ANNABA.

[21] <https://www.wikiwand.com/fr/Apprentissagepar-renforcement-profond> Media . (visité au 17/03/2023)

[22] Kelly H ZOU, Kemal TUNCALI et Stuart G SILVERMAN. " Correlation and simple linear regression " (2003).

[23] Vikramaditya JAKKULA. " Tutorial on support vector machine (svm) " In : School of EECS, Washington State University (2006).

[24] <https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm> (visité au 24/03/2023)

[25] Ian GOODFELLOW, Yoshua BENGIO et Aaron COURVILLE. Deep learning. MIT press, 2016.

[26] <https://ai.plainenglish.io/introduction-to-k-nearest-neighbors-knn-algorithm-e8617a448fa8> (visité au 06/04/2023)

[27] <http://comprendre-machine-learning.fsil.fr/introduction-aux-reseaux-de-neurones/> (visité au 11/04/2023)

- [28] Leonardo NORIEGA. " Multilayer perceptron tutorial ". In : School of Computing. Staffordshire University (2005).
- [29] <https://www.researchgate.net/figure/Architecture-du-Perceptron-Multicouche> (visité au 25/04/2023)
- [30] Ian GOODFELLOW, Yoshua BENGIO et Aaron COURVILLE. Deep learning. MIT press, 2016.
- [31] <https://eduscol.education.fr/sti/si-ens-paris-saclay/ressources-pedagogiques/introduction-a-lapprentissage-automatique> (visité au 29/04/2023)
- [32] Tomas MIKOLOV et al. " Recurrent neural network based language model. " In : Interspeech. T. 2. 3. Makuhari. 2010 .
- [33] <https://www.researchgate.net/figure/The-folded-and-unfolded-structure-of-recurrent-neural-networks-1-RNN-Similar> (visité au 21/05/2023)
- [34] Rafael C GONZALEZ. " Deep convolutional neural networks [Lecture Notes] ". In : IEEE Signal Processing Magazine (2018).
- [35] Ikram Chraïbi Kaadoud, "Apprentissage de séquences et extraction de règles de réseaux récurrents : application au traçage de schémas techniques ", thèse de doctorat en informatique, sous la direction de Frédéric Alexandre, Université de Bordeaux, 2018.
- [36] Sepp Hochrieter, Jurgen Schmidhuber(1997)[LSTM].