

République Algérienne Démocratique et Populaire
Ministre de l'enseignement Supérieur et de la Recherche Scientifique



جامعة بجاية
Tasdawit n Bgayet
Université de Béjaïa

Université A/Mira de Béjaïa
Faculté des Sciences Exactes
Département d'Informatique

MÉMOIRE DE MASTER

En

Informatique

Option

INTELLIGENCE ARTIFICIELLE

Thème

Analyse et prédiction de la consommation
d'électricité pour l'Internet des comportements

Réalisé par : Mr. Demmouche Abdelghani

Mr. Ourak Imadeddine

Soutenu le 29 septembre 2022 devant le jury composé de :

Président	Mme S. Alloui	Maître de conf. A	U. A/Mira Béjaïa.
Rapporteur	M. A. Belaid	Professeur	U. A/Mira Béjaïa.
Rapporteur	M. E. Khennouche	Maître de conf. A	U. A/Mira Béjaïa.
Examineur	Mme D. Kessira	Maître de Assi. A	U. A/Mira Béjaïa.

Béjaïa, le 20 Septembre 2022.

** Remerciements **

Avant toute chose, on remercie dieu tout puissant pour nous avoir aidé et éclairé le chemin pour la réalisation de ce mémoire. À l'issue de ce modeste travail, on tient à exprimer nos sincères remerciements aux membres du jury pour avoir accepté d'examiner et d'évaluer ce travail. Mrs KHENNOUCHE Mouhamed et BELAÏD Ahror pour leur encadrement, leur grande disponibilité, leur confiance ainsi que leur soutien et leurs conseils qu'ils nous ont accordés durant toute cette année. À nos familles et en particulier nos parents.

✧ *Dédicaces* ✧

A mes très chers parents.

M. Demmouche Abdelghani

✧ *Dédicaces* ✧

A mes très chers parents.

M. Ourak Imadeddine

Table des matières

Table des matières	1
Liste des figures	1
Liste des Tableaux	6
1 Introduction	7
1.1 Contexte et problématique	7
1.2 Motivation et challenges	9
1.3 Contribution	10
1.4 Plan du manuscrit	10
2 Vue d'ensemble sur l'internet des comportements	11
2.1 Introduction	11
2.2 IoT	12
2.2.1 Que signifie Internet des Objets (IoT) ?	12
2.2.2 Importance de l'IoT	12
2.2.3 Utilisation de l'IoT dans divers secteurs	13
2.2.3.1 Finances	13
2.2.3.2 Vente au détail	13
2.2.3.3 Santé	14
2.2.4 Avantage de l'IoT	14
2.2.5 L'IoB et l'IoT	14
2.3 Internet des comportements	16
2.3.1 Que signifie ?	16
2.3.2 Importance de l'Internet du comportement	16

2.3.3	Utilisation de l’IoB dans divers secteurs	17
2.3.3.1	IoB en entreprise	17
2.3.3.2	IoB pendant la pandémie de Covid-19	17
2.3.3.3	IoB pour le secteur de l’assurance	18
2.3.4	Avantage de L’IoB	18
2.3.5	Comment les données IoB sont-elles collectées et utilisées?	18
2.4	Big Data	19
2.4.1	Définition de Big Data	19
2.4.2	Comment fonctionne le Big Data?	20
2.4.3	Comment le Big Data permet de prévoir le comportement du consommateur?	20
2.4.4	Analyse du comportement du consommateur à l’ère du Big Data	21
2.4.5	Caractéristiques du comportement des consommateurs à l’ère du Big Data	22
2.4.5.1	Choix de comportement des consommateurs est plus rationnel .	22
2.4.5.2	Demande des consommateurs continue d’augmenter	22
2.4.5.3	Confiance des consommateurs dans les fonctions commerciales des médias sociaux a augmenté	23
2.4.6	Dangers et potentiel du Big Data	23
2.4.7	Big Data et l’IoB	24
2.4.8	Big Data et l’IoT	25
2.5	Machine learning	26
2.5.1	Définition du machine learning	26
2.5.2	Types de machine learning	26
2.5.2.1	Machine learning supervisé	27
2.5.2.2	Machine learning non supervisé	27
2.5.2.3	Machine learning par renforcement	28
2.5.3	IoT et le machine learning	28
2.5.4	Comment le machine Learning s’applique-t-elle aux données de l’IoT? .	28
2.5.5	IoT et Machine Learning : Avantages de la combinaison des deux	29
2.5.6	L’impact du Machine learning sur l’IoT	29
2.6	Conclusion	29

3	Etat de l'art sur l'internet des comportements	31
3.1	Introduction	31
3.2	L'Internet des comportements (IoB) et son rôle dans les services clients	31
3.3	L'internet du comportement comme espace de transformation organisationnelle avec intelligence chorégraphique :	32
3.4	Internet du comportement et systèmes d'IA explicables pour influencer le com- portement de l'IoT	34
3.5	Conception de systèmes Internet des comportements	35
3.6	Etude de la pollution de l'air a Pékin affecté ou non sur les activiés de vie quotidienne des personne basé sur l'intenet des comportemnts	36
3.7	Analyse comparative	39
3.7.1	Critères de comparaison	39
3.7.2	Tableau Comparatif	40
3.8	Conclusion	42
4	Mise en œuvre et évaluation	43
4.1	Introduction	43
4.2	Environnement de développement	43
4.2.1	Matériels	43
4.2.2	Logiciels	44
4.2.2.1	Langages de programmation et packages	44
4.2.2.2	Environnement de Développement Intégré IDE	46
4.3	Métriques d'évaluation	47
4.3.1	RMSE (Racine de l'erreur quadratique moyenne)	47
4.3.2	MAE (Erreur Absolue Moyenne)	47
4.4	Visualisation et analyse des données	48
4.4.1	Description des données	48
4.4.2	pretraitement des données	49
4.4.3	Observation de la consommation d'électricité au fil du temps	50
4.4.3.1	Visualisation annuelle pour la variable la puissance active totale	52
4.4.3.2	Visualisation mensuelle pour la variable la puissance active totale	53

4.4.3.3	Visualisation quotidienne pour la variable la puissance active totale	53
4.4.3.4	Distribution annuelle pour la variable la puissance active totale	56
4.5	Modèles de résolution	57
4.5.1	Préparation des données	57
4.5.1.1	Normalisation des données	57
4.5.1.2	Données d'apprentissage et données de test	57
4.5.2	LSTM long short-term memory	57
4.5.2.1	Évaluation du modèle LSTM	58
4.5.3	Random forest	59
4.5.3.1	Évaluation du modèle Random Forest	59
4.5.4	XGBRegressor	60
4.5.4.1	Évaluation du modèle XGBRegressor	60
4.5.5	SVRegressor	61
4.5.5.1	Évaluation du modèle SVRegressor	61
4.5.6	Tableau de comparaison	62
4.5.7	Comparaison des résultats des modèles	62
4.6	Conclusion	63
	Conclusion générale	64
	Bibliographie	65

Table des figures

2.1	Architecture illustrant l'IoT et l'IoB[7]	15
3.1	Flux de travail du système IoB pour influencer le comportement de l'IoT	35
3.2	Architecture logicielle de gestion des files d'attente des Galeries des Offices à Florenc	36

3.3	Le framwork IoB pour Etude de la pollution de l'air a Pékin	39
4.1	Ensemble de données sur la consommation électrique des habitations	48
4.2	Ensemble de données sur la consommation électrique des ménages	49
4.3	La puissance active totale	50
4.4	La puissance réactive totale	50
4.5	Tension moyenne	50
4.6	Intensité moyenne du courant	50
4.7	Énergie active pour la cuisine	50
4.8	Énergie active pour la Buanderie	50
4.9	Énergie active pour les systèmes de climatisation	51
4.10	Visualisation annuelle de l'énergie active	52
4.11	Visualisation mensuelle de l'année 2008	53
4.12	Visualisation quotidienne de les 15 premiers jours de l'année 2008	54
4.13	Distribution annuelle de la puissance active totale	56
4.14	evolution du modèle LSTM	58
4.15	Performances du modèle de prédiction LSTM	58
4.16	Performances du modèle de prédiction Random forest	59
4.17	Performances du modèle de prédiction XGBRegressor	61
4.18	Performances du modèle de prédiction SVRegressor	62

Liste des tableaux

3.1	Table de comparaison des approches utilisant IoB	41
4.1	Spécifications des ordinateurs utilisés	44
4.2	Tableau de comparaison des tests d'erreur des modèles	62

Chapitre 1

Introduction

1.1 Contexte et problématique

Dans notre monde hyperconnecté dominé par les connexions web, il est difficile d'imaginer que des experts considèrent que l'internet aura besoin d'une "refonte massive". Pourtant, c'est précisément cette omniprésence qui a poussé de plus en plus de technologues à travailler sur ce qu'ils appellent une "nouvelle phase" de l'internet. Ils insistent sur le fait que cet internet de "deuxième génération" doit changer pour être beaucoup plus intelligent ; doit évoluer pour devenir un "web sémantique" qui, en plus d'être plus efficace, nous offre plus de contrôle sur nos données. C'est ce qu'ils prévoient avec l'arrivée du Web 3.0, que de nombreux acteurs du secteur considèrent comme la "grande révolution de l'internet". Cette révolution de l'internet crée plus de 30 milliards d'objets connectés en 2010 (plus que d'humains)[1]. De là, une nouvelle technologie est apparue, qui est Internet des Objets (IoT) qu'est un réseau d'objets physiques reliés entre eux qui collectent et échangent des informations et des données via Internet. L'équipement est lié et traité de manière autonome par ces objets, et les données dans le cloud se développent de plus en plus de manière complexe grâce à l'IoT. Les objets connectés produisent de grandes quantités de données dont le stockage et le traitement entrent dans le cadre de ce que l'on appelle les Big Data. En logistique, il peut s'agir de capteurs qui servent à la traçabilité des biens pour la gestion des stocks et les acheminements. Dans le domaine de l'environnement, il est question de capteurs surveillant la qualité de l'air, la température, le niveau sonore, l'état d'un bâtiment, etc. Lorsque le professeur de psychologie « Gothe Nyman » [2] a suggéré l'idée de recueillir des données précises sur l'utilisation et le comportement des consommateurs lorsqu'ils

s'engagent avec l'Internet des objets (IoT) en 2012, plusieurs personnes l'ont considéré comme le début de Internet of Behaviour (IoB). Cependant, pour évaluer les données client pour les objectifs de l'entreprise.

L'Internet comportemental est considéré comme l'une des principales tendances technologiques pour l'année 2021. Pendant la pandémie de COVID-19, l'analyse des images des caméras de surveillance et des photos partagées sur les réseaux sociaux a fourni des statistiques sur l'utilisation des masques, souvent avec des données GPS exploitables indiquant, par exemple, où placer davantage d'affiches. Suite à l'épidémie, des systèmes informatiques de "passeport" de santé ont été développés. Ils vous permettent d'identifier les zones à haut risque, votre statut vaccinal et les personnes que vous avez rencontrées au cours d'une période donnée. Un exemple est l'application française TousAntiCovid[4]. IoB collecte de grandes quantités de données comportementales humaines à partir d'appareils IoT et les transforme en influences précieuses afin d'améliorer l'expérience utilisateur et l'expérience de recherche en modifiant le comportement, les intérêts et les préférences des utilisateurs. L'Internet des objets (IoT), l'expérience utilisateur/de recherche et le comportement sont les trois domaines dans lesquels IoB opère. Il relie les appareils à Internet afin de suivre l'activité des utilisateurs, qui est ensuite analysée afin d'améliorer la recherche et l'expérience utilisateur et d'influencer le comportement des utilisateurs afin d'atteindre un objectif spécifique, d'un point de vue psychologique, IoB essaie de comprendre correctement les données et d'appliquer cette compréhension pour créer de nouveaux produits, commercialiser des articles existants, réorganiser la chaîne de valeur, améliorer les revenus ou réduire les dépenses. Les comportements des consommateurs ne sont pas les seuls pris en compte. Il vise également à améliorer la santé, le bien-être, la nutrition, les performances sportives, mieux maîtriser le diabète, la consommation de tabac, la gestion du stress, changer le comportement et l'empreinte écologique des gens. Là se pose le problème, y a-t-il une relation entre psychologie et IoB?. L'exploitation de l'internet des comportements sur divers projets pose toujours le problème de prédiction du comportement des utilisateur, les modèles de série chronologique sont des défis scientifique importants qui trouvent leurs applications dans des domaines aussi variés que la finance, la production d'électricité, la climatologie,etc. Parmi les problèmes posés, c'est que comment prédire le comportement humain sachant que le comportement humain n'est pas absolu et constant pour une seul comportement tout le temps ?

Dans ce travaille nous définirons ce qu'est l'IoB et les différentes techniques liées à ce terms

et expliquerons les différentes approches qui construisent IoB pour leurs études se terminant finalement par la mise en œuvre où nous analysons et prédisons le comportement de la consommation d'électricité pour une seule ménage durant 4 ans.

1.2 Motivation et challenges

La prévision désigne le processus de prédiction de la valeur numérique des observations à venir d'une série temporelle. Il s'agit d'un problème difficile car il s'agit d'un processus d'extrapolation et il existe une incertitude importante dans les estimations. Les modèles reposent sur des hypothèses dérivées de données passées qui souvent ne tiennent pas. En plus de cela, le processus de prévision lui-même pose des défis supplémentaires. Par exemple, dans la plupart des applications, nous sommes intéressés à prédire de nombreuses valeurs à l'avance (par exemple, les valeurs pour chacune des prochaines 24 heures). Bien que de grands horizons de prévision soient souhaitables, ils compliquent considérablement la tâche en raison de l'incertitude accrue il est plus facile de prédire l'avenir à court terme ou immédiat. Parfois, nous sommes intéressés par la prévision, non pas du comportement général d'une série chronologique, mais d'événements d'intérêt spécifiques et rares. L'objectif principal derrière la détection d'anomalies de séries chronologiques est la détection en temps opportun d'événements intéressants mais rares, qui peuvent être perturbateurs dans le domaine d'application particulier. Cela peut être crucial pour la prise de décision car cela permet aux professionnels de prendre les mesures appropriées pour prévenir ces événements ou atténuer leurs conséquences. Le deuxième défi est que les événements d'intérêt sont rares. Par conséquent, l'apprentissage du concept derrière ces événements représente un problème d'apprentissage déséquilibré.

Les séries chronologiques poursuivent deux objectifs majeurs : identifier la nature du phénomène représenté par une séquence d'observations, et réaliser des prévisions (les valeurs que la variable de la série chronologique devrait prendre dans le futur). Ces deux objectifs nécessitent une identification, et une description plus ou moins formelle des composantes de la série chronologique observée. Une fois ces composantes établies, il nous reste à les interpréter et à les intégrer dans d'autres données (c'est-à-dire, les utiliser dans notre approche du phénomène étudié, par exemple, la consommation d'électricité). Sans se soucier de la bonne compréhension ni de la pertinence de notre interprétation du phénomène, nous pouvons extrapoler la composante

identifiée afin de prédire les événements futurs.

1.3 Contribution

Nous avons pris pour étude la prédiction du comportements des ménages selon leur consommation d'électricité. Nous mesurons donc quotidiennement l'électricité consommée dans chaque maison pour pouvoir juger son comportement selon la valeur obtenue. Ce comportement peut varier sur plusieurs niveaux de consommation partant du bas, allant jusqu'au haut niveau.

1.4 Plan du manuscrit

Ce travail est divisée en Quatre chapitres chapitres.

Le chapitre 1 contient le contexte et la motivation de ce travail ainsi qu'un aperçu des contributions proposées et le plan du manuscrit.

Le chapitre 2 contient d'abord, une définition de l'internet des comportement et quelques notions de base liées à ce concept ; Internet des objets, Big Data.

Le chapitre 3 contient l'état de l'art, Nous présentons les différentes approches basées sur l'internet des comportements et on définit les avantages et limites de chaque approche.

Le chapitre 4 contient les différents outils nécessaires lors de la processus de mise en œuvre du modèle. Ensuite, nous expliquerons toutes les étapes que nous avons suivies à partir en définissant les logiciels et matériels utilisés. Enfin, nous discuterons de notre cadre expérimental et les résultats.

Chapitre 2

Vue d'ensemble sur l'internet des comportements

2.1 Introduction

Internet est devenu un moyen d'information et de communication intégré à notre quotidien. Les développements de l'Internet des objets, des mégadonnées et de l'intelligence artificielle ont porté ce type de recherche à un nouveau niveau et ont conduit à l'émergence de l'Internet des comportements, qui analyse les modèles de comportement. L'intelligence artificielle et l'Internet des objets sont devenus très populaires dans divers aspects de notre vie. Les applications de l'Internet des objets ont radicalement changé nos vies, apportant une immense valeur aux activités des particuliers et des entreprises. Cette profusion a donné naissance aux appareils portables, aux applications domestiques intelligentes, aux systèmes de soins de santé avancés, et à l'automatisation industrielle près des années d'incertitude, L'adoption généralisée de ces technologies dans notre société moderne et leurs développements récents conduisent naturellement et inévitablement à l'émergence de l'Internet du comportement. L'objectif de ce chapitre est d'approfondir des notion de base telle que l'internet du comportement, l'internet des objets, Big Data et machine learning.

2.2 IoT

2.2.1 Que signifie Internet des Objets (IoT) ?

L'Internet of Things (IoT) décrit le réseau de terminaux physiques, les « objets », qui intègrent des capteurs, des logiciels et d'autres technologies en vue de se connecter à d'autres terminaux et systèmes sur Internet et d'échanger des données avec eux. Ces terminaux peuvent aussi bien être de simples appareils domestiques que des outils industriels d'une grande complexité[22]. En effet l'Internet des objets est un système d'appareils informatiques interdépendants, de machines mécaniques et numériques, d'objets, d'animaux ou de personnes qui reçoivent des identifiants uniques (UID) et la capacité de transférer des données sur un réseau sans nécessiter d'humain à humain, interaction humaine ou homme-ordinateur. Une chose dans l'Internet des objets peut être une personne avec un implant de moniteur cardiaque, un animal de ferme avec un transpondeur à biopuce, une automobile qui a des capteurs intégrés pour alerter le conducteur lorsque la pression des pneus est basse ou tout autre élément naturel ou artificiel. Objet qui peut se voir attribuer une adresse IP (Internet Protocol) et qui est capable de transférer des données sur un réseau[21].

2.2.2 Importance de l'IoT

Ces quelques dernières années, l'IoT est devenu l'une des technologies les plus importantes du 21ème siècle. Maintenant que nous pouvons connecter des objets du quotidien (appareils électroménagers, voitures, thermostats, interphones bébés) à Internet par l'intermédiaire de terminaux intégrés, des communications sont possibles en toute fluidité entre les personnes, les processus et les objets.

Grâce à des traitements informatiques peu coûteux, au cloud, au Big Data, à l'analytique et aux technologies mobiles, les objets physiques peuvent partager et collecter des données avec un minimum d'intervention humaine. Dans ce monde hyperconnecté, les systèmes digitaux peuvent enregistrer, surveiller et ajuster chaque interaction entre les objets connectés. Le monde physique rencontre le monde digital, et ils coopèrent. L'IoT permet aux entreprises d'automatiser les processus et de réduire les coûts de main-d'œuvre. Il réduit également les déchets et améliore la prestation de services, ce qui rend la fabrication et la livraison des marchandises moins

coûteuses, tout en offrant une transparence dans les transactions des clients.

En tant que tel, l'IoT est l'une des technologies les plus importantes de la vie quotidienne, et il continuera de prendre de l'ampleur à mesure que de plus en plus d'entreprises réalisent le potentiel des appareils connectés pour rester compétitives[20].

2.2.3 Utilisation de l'IoT dans divers secteurs

Alors que l'IoT continue de se déployer à l'échelle mondiale et de se répandre dans tous les secteurs, certaines industries mènent des investissements dans cette technologie révolutionnaire, qui change notre façon de vivre et de travailler parmi les plus grand secteur d'utilisation de l'IoT on peut cité :

2.2.3.1 Finances

L'IoT devient de plus en plus sécurisé. Les banques et les clients se sont habitués à gérer les transactions financières via différents appareils connectés. Il n'est pas rare de voir des distributeurs intelligents avec des distributeurs automatiques de billets connectés. Comme la quantité de données transférées et collectées est énorme avec l'IoT, les entreprises financières peuvent mesurer le risque avec précision. Avec le temps, les banques commenceront à utiliser des capteurs et des analyses de données pour collecter beaucoup plus d'informations sur les clients et ainsi proposer des services personnalisés. Cela aidera les banques à comprendre comment leurs clients achètent et dépensent leur argent.

2.2.3.2 Vente au détail

La technologie de l'Internet des objets aide les détaillants à réduire leurs coûts opérationnels et à améliorer l'expérience des clients grâce à des cas d'utilisation nouveaux et innovants. Les capteurs GPS et les étiquettes RFID offrent un parcours complet du mouvement de la marchandise de la fabrication à l'étagère du magasin jusqu'aux clients et les capteurs GPS et les étiquettes RFID offrent un parcours complet du mouvement de la marchandise de la fabrication à l'étagère du magasin jusqu'aux clients. La plupart des détaillants trouvent la gestion des stocks fastidieuse. Un suivi des stocks inexact entraîne des ruptures de stock, des surstocks et des pertes et aussi lorsque le magasin est bondé, le passage à la caisse reste le processus le plus

désagréable pour les clients. Les solutions IoT peuvent automatiser et personnaliser les caisses.

2.2.3.3 Santé

Le secteur de la santé a commencé à utiliser de manière optimale les technologies IoT. La technologie IoT dans les soins de santé est représentée par des systèmes de suivi et des systèmes de santé en temps réel et est responsable de l'amélioration du traitement des patients, du diagnostic, de la maintenance des équipements médicaux et de diagnostic et des chirurgies à distance. Les soins de santé connectés sont nécessaires pour prendre les bonnes décisions, prendre les bonnes mesures, traiter intelligemment et donc la satisfaction des patients[37].

2.2.4 Avantage de l'IoT

L'internet des objets offre plusieurs avantages aux organisations. Certains avantages sont spécifiques à l'industrie, et certains sont applicables à plusieurs industries. Certains des avantages communs de l'IoT permettent aux entreprises de :

- surveiller l'ensemble de leurs processus commerciaux.
- améliorer l'expérience client (CX).
- gagner du temps et de l'argent et améliorer la productivité des employés.
- intégrer et adapter les modèles commerciaux.
- prendre de meilleures décisions commerciales et générer plus de revenus [21].

2.2.5 L'IoB et l'IoT

L'Internet du comportement est une extension de l'IoT en effet l'IoB est un cas d'utilisation de l'IoT qui se concentre sur le fait d'inciter les utilisateurs à améliorer leur comportement (par exemple, la rentabilité du fournisseur d'IoT, la durabilité, le bien-être de la société, etc.). On peut voir IoB comme une intersection de trois piliers :

- **IoT** : Fournit les données des clients telles que l'emplacement, la routine quotidienne, l'état de santé, etc.
- **Psychologie d'utilisateur** : Cherche à comprendre la motivation derrière les actions d'un individu.

- **Analyse de données** : les algorithmes peuvent utiliser les données de l'IoT et les résultats de la psychologie pour identifier les modèles de comportement et les recommandations.

IoB est l'extension naturelle de l'IoT où ces trois composants (IoT, Psychologie d'utilisateur, Analyse de données) sont utilisés pour atteindre la satisfaction client à long terme et de plus grands profits. Les entreprises ont utilisé les mégadonnées pour comprendre les désirs et les besoins de leurs clients. Cependant, en reliant le Big Data avec la psychologie et de puissants modèles d'interprétation de données basés sur l'IA/ML, il est possible de comprendre les raisons qui sous-tendent les décisions des clients et d'obtenir de plus grands profits en influant sur les comportements des clients. Il ne s'agit pas du tout des choses lorsque les entreprises qui nous connaissent grâce aux données fournies par l'IoT peuvent désormais influencer notre comportement en utilisant les données fournies par l'IoB. Envisagez d'utiliser une application de santé pour smartphone pour vérifier votre nutrition, vos habitudes de sommeil, votre fréquence cardiaque ou votre glycémie. L'application peut vous avertir des circonstances potentiellement dangereuses et proposer des changements de comportement qui conduiraient à un résultat plus positif ou souhaitable. Pour le moment, les entreprises utilisent principalement l'IoT et l'IoB pour surveiller et tenter d'influencer notre comportement afin d'atteindre l'objectif visé par le comportement d'Allstate[7].

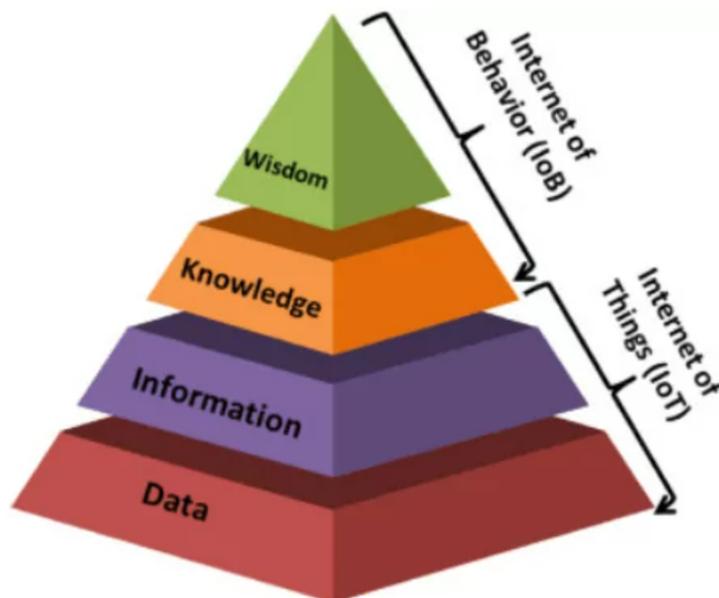


FIGURE 2.1 – Architecture illustrant l'IoT et l'IoB[7]

2.3 Internet des comportements

2.3.1 Que signifie ?

L'Internet des comportements (IoB) est un domaine de recherche et développement qui cherche à comprendre comment, quand et pourquoi les humains utilisent la technologie pour prendre des décisions. IoB combine trois domaines d'études : les sciences du comportement, science des données et l'Internet des objets (IoT)[15].

Les plates-formes IoB sont conçues pour collecter, agréger et analyser les données générées à partir d'une grande variété de sources, y compris les appareils numériques domestiques, les ordinateurs portables et les activités humaines en ligne. Les données sont ensuite analysées en termes de psychologie comportementale pour rechercher des modèles pouvant être utilisés par les équipes marketing, médicale et commerciales etc. pour influencer le comportement futur des utilisateurs.

Un objectif important d'IoB est d'aider les spécialistes à comprendre Données massives générées par les nœuds de réseau dans l'Internet des objets[15].

2.3.2 Importance de l'Internet du comportement

Il ne fait aucun doute que les données sont importantes pour les entreprises depuis la création d'Internet. Cependant, en 2021, Gather a qualifié l'IoB de l'une des principales tendances technologiques. l'Internet des comportements offre à chacun plus d'opportunités de collecter des données et de les analyser. Elle peut garantir que les entreprises poursuivent leur croissance face aux bouleversements importants liés à la pandémie actuelle et au changement climatique de l'économie mondiale. Cependant, l'objectif principal de l'Internet du comportement est de collecter, d'analyser, de répondre et de comprendre tous les types de comportements afin d'améliorer l'expérience client/utilisateur. En dehors de cela, les données comportementales aident également les entreprises à prendre des décisions plus éclairées et à améliorer la qualité de leur service et leur chaîne de valeur de la meilleure façon possible. Dans l'ensemble, La psychologie et le marketing fonctionnent ensemble depuis le début de la publicité. De cette façon, les entreprises peuvent obtenir de nouvelles informations sur les données collectées par l'IoT. L'IoB est devenu un nouvel outil puissant pour les ventes et le marketing des entreprises du monde

entier. Grâce à cela, les entreprises peuvent acquérir une compréhension approfondie de leurs clients pour les garder plus satisfaits. En bref, l'Internet du comportement est là pour générer une impulsion significative dans le développement de l'industrie des ventes[5].

2.3.3 Utilisation de l'IoB dans divers secteurs

L'IoB est l'une des 10 principales tendances identifiées dans le rapport technologique stratégique de Gartner pour 2021 et peut générer une génération de valeur significative. Les données produites par les appareils IoT peuvent être utilisées par les équipes marketing pour stimuler les ventes qui relèvent du sujet IoB. En outre, l'IoB pourrait guider les entrepreneurs et les chefs d'entreprise pour qu'ils prennent des décisions d'investissement plus axées sur les données, car l'IoB fournit plus d'informations sur les préférences des clients. L'IoB peut être utilisé dans divers secteurs parmi eux on cite :

2.3.3.1 IoB en entreprise

La publicité en ligne est de plus en plus utilisée par diverses entreprises pour atteindre leurs clients. Ils peuvent découvrir et cibler certaines personnes ou groupes qui pourraient bénéficier de leurs produits ou services avec l'aide d'IoB. Google et Facebook utilisent des données comportementales pour fournir des publicités aux utilisateurs sur leurs sites. Cela permet aux entreprises d'interagir avec leurs consommateurs cibles et de mesurer leur comportement en réponse aux publicités via des "taux de clics". De même, Youtube utilise l'analyse comportementale pour améliorer l'expérience du spectateur en recommandant ou en mettant en évidence uniquement les vidéos et les sujets qui l'intéressent.

2.3.3.2 IoB pendant la pandémie de Covid-19

L'épidémie a permis de mieux connaître les précautions à prendre. Les employeurs peuvent utiliser des capteurs ou des étiquettes RFID pour voir s'il y a des incohérences dans le respect des normes de sécurité. Les restaurants et les applications de livraison de nourriture, par exemple, utilisent les informations du protocole pour guider leurs décisions. Swiggy et Zomato, par exemple, ont tous deux exposé et promu les procédures de sécurité des restaurants. Ils ont également enregistré et diffusé la température du livreur pour rassurer les consommateurs sur

leur sécurité.

2.3.3.3 IoB pour le secteur de l'assurance

Dans le secteur de l'assurance, l'IoB peut être très bénéfique. Les outils de suivi des conducteurs sont déjà utilisés par des compagnies d'assurance comme Allstate et StateFarm pour suivre et sécuriser la conduite d'un conducteur. Avec l'aide d'IoB, ils peuvent évaluer le comportement et peut-être déterminer si un certain événement était un accident ou une hypothèse erronée de la part de l'assuré. Cela peut aider à prévenir les incidents de conduite en état d'ébriété, de conduite sous l'influence de drogues et même les mineurs ou les retraités de prendre le volant et de provoquer un accident [12].

2.3.4 Avantage de L'IoB

L'Internet du comportement peut être bénéfique et très utile on va cité ici quelques avantages spécifiques de l'IoB[14].

- IoB peut aider les entreprises à résoudre leurs problèmes en concluant plus de ventes et en gardant leurs clients très satisfaits[2].
- Cela peut même vous aider à remplacer plusieurs enquêtes auprès des clients, qui prennent du temps à la fois pour les clients et les entreprises[14].
- IoB peut vous permettre d'analyser le comportement de vos prospects et leurs habitudes d'achat sur différentes plateformes[14].
- Vous pouvez étudier des données impossibles à obtenir sur la façon dont vos clients potentiels interagissent avec votre entreprise, vos produits et vos services[5].
- Une source de business models innovants.
- Cela vous aidera même à fournir des notifications en temps réel à vos clients sur toute dernière offre, point de vente ou même publicité ciblée[14].

2.3.5 Comment les données IoB sont-elles collectées et utilisées ?

Les données des utilisateurs peuvent être recueillies à partir d'une gamme de sites et de technologies, y compris le site Web d'une entreprise, les profils de médias sociaux, les capteurs,

la télématique, les balises, les moniteurs de santé (tels que Fitbit) et une variété d'autres appareils. Chacun de ces sites rassemble différents types d'informations. Par exemple, un site Web peut garder une trace du nombre de fois qu'une personne visite une certaine page ou combien de temps elle y reste. De plus, la télématique peut suivre la force de freinage du conducteur d'un véhicule ou la vitesse typique du véhicule. Les données sont collectées et analysées par les entreprises à diverses fins. Ces raisons incluent l'aide aux entreprises à prendre des décisions commerciales éclairées, la personnalisation des techniques de marketing, Obtenir des informations plus détaillées sur le comportement des utilisateurs. Les entreprises établissent des normes pour faciliter l'analyse de ces données. Lorsqu'un utilisateur effectue une ou plusieurs actions spécifiques, l'entreprise commence alors à convaincre l'utilisateur de modifier son comportement. Par exemple, si un utilisateur visite trois fois la page d'une entreprise vendant des jeans slim pour hommes, la boutique numérique peut lui montrer une publicité pop-up lui offrant 25 % de réduction sur une paire de jeans. Combiner des données provenant de nombreuses sources et les évaluer pour prendre une décision est une autre composante de l'Internet des comportements. Les entreprises peuvent développer des profils d'utilisateurs détaillés pour chaque utilisateur en combinant des données provenant de diverses sources. Ces profils peuvent ensuite être examinés pour voir quel est le meilleur plan d'action pour la personne[12].

2.4 Big Data

2.4.1 Définition de Big Data

Les Big Data ou mégadonnées désignent l'ensemble des données numériques produites par l'utilisation des nouvelles technologies à des fins personnelles ou professionnelles. Cela recoupe les données d'entreprise (courriels, documents, bases de données, historiques de processeurs métiers...) aussi bien que des données issues de capteurs, des contenus publiés sur le web (images, vidéos, sons, textes), des transactions de commerce électronique, des échanges sur les réseaux sociaux, des données transmises par les objets connectés (étiquettes électroniques, compteurs intelligents, smartphones...), des données géolocalisées, etc.

L'expression « Big Data » date de 1997 selon l'Association for Computing Machinery. En 2001, l'analyste du cabinet Meta Group (devenu Gartner) Doug Laney décrivait les Big Data

d'après le principe des « trois V » :

- le Volume de données de plus en massif.
- la Variété de ces données qui peuvent être brutes, non structurées ou semi-structurées.
- la Vélocité qui désigne le fait que ces données sont produites, récoltées et analysées en temps réel.

. Certaines entreprises ajoutent un quatrième « V » à cette définition pour la Véracité qui évoque la nécessité de vérifier la crédibilité de la source et la qualité du contenu afin de pouvoir exploiter ces données[8].

2.4.2 Comment fonctionne le Big Data ?

Le Big Data permet de relever un immense défi technologique : stocker une grande quantité de données provenant de différents canaux sur un immense disque dur, facilement accessibles depuis les quatre coins de la planète. Des données stockées en lieu sûr et récupérables à tout moment en cas d'incident quelconque.

Pour y parvenir, les fichiers sont découpés en plusieurs fragments nommés « chunks ». Ces fragments sont ensuite répartis sur plusieurs ordinateurs, et il existe différentes façons de les reconstituer. Une panne survient ? Une autre machine vient prendre le relais en empruntant un autre chemin. Ainsi, les différentes pièces du puzzle restent disponibles en permanence, et peuvent être assemblées d'une façon ou d'une autre.

La duplication massive des données est l'une des clés de voûte de l'architecture du Big Data. Le cloud computing, les supercalculateurs hybrides (high performance computing ou HPC) et les systèmes de fichiers distribués (DFS ou distributed files system) figurent parmi les principaux modèles de stockage actuellement disponibles [8].

2.4.3 Comment le Big Data permet de prévoir le comportement du consommateur ?

Le Big Data inclut les informations publiées sur les réseaux sociaux (âge et coordonnées, avis, commentaires, photos, vidéos), la navigation web (à travers les fameux cookies) et les achats en ligne. Ces données fournissent des indices très intéressants sur les comportements des consommateurs et les tendances du marché.

Si vous venez d'acheter des chaussures de running sur un site de e-commerce et si vous avez publié une photo de vous en train de courir le marathon de Paris, vous êtes une cible alléchante pour les enseignes de sport. Ces marques pourront ainsi vous envoyer des newsletters ou des offres promotionnelles sur des produits associés à vos chaussures afin de compléter votre équipement.

Le Big Data permet de connaître votre profil, mais aussi votre comportement global : fréquence d'utilisation des réseaux sociaux et de vos achats en ligne (historique des transactions, dépenses effectuées), canaux utilisés, heures de connexion, etc.

Le Big Data est un outil essentiel pour les entreprises BtoB et BtoC. Les données collectées les aident à confectionner des campagnes marketing personnalisées adaptées aux besoins, aux préférences et aux comportements des consommateurs. Ces informations contribuent à améliorer l'expérience client, à attirer les prospects et à fidéliser les clients existants.

Un ciblage amélioré permet de rendre plus efficaces les campagnes marketing et de toucher le segment désiré, celui qui est le plus susceptible d'être intéressé par les produits et/ou services de l'entreprise. Le Big Data constitue par ailleurs un avantage compétitif pour les professionnels qui détiennent une multitude de données, car ils peuvent anticiper les changements de comportements et mieux comprendre pourquoi les consommateurs se sont tournés vers tel ou tel prestataire.

Le Big Data est un outil précieux dans de vastes domaines privés et publics allant de la vente en ligne à la recherche scientifique en passant par la culture, la politique (campagnes électorales), le transport, les assurances, l'industrie, le secteur bancaire et énergétique [8]

2.4.4 Analyse du comportement du consommateur à l'ère du Big Data

À l'heure actuelle, l'analyse des mégadonnées a été largement utilisée dans tous les domaines de la vie et est progressivement devenue une force productive indispensable, favorisant l'allocation et l'utilisation efficaces des moyens de production et favorisant rapidement l'amélioration de l'efficacité de la production sociale. Les mégadonnées font référence à la grande quantité de données impliquées dans la collecte de données, et les outils d'analyse conventionnels ne peuvent pas terminer l'acquisition, le traitement et la collation des données dans un court laps

de temps. Différente de l'analyse d'échantillons, l'analyse de données volumineuses consiste à traiter toutes les données de manière exhaustive, complète et professionnelle et à en tirer des informations efficaces [11].

2.4.5 Caractéristiques du comportement des consommateurs à l'ère du Big Data

2.4.5.1 Choix de comportement des consommateurs est plus rationnel

L'avènement de l'ère des mégadonnées a changé la façon dont les consommateurs obtiennent des informations sur les produits, et les informations qu'ils connaissent sont plus suffisantes et précises. Dans le modèle de marché traditionnel, les consommateurs connaissent principalement un certain produit ou une certaine marque par le biais de publicités et manquent d'autres supports d'information, ce qui limitera la prise de décision rationnelle des consommateurs. À l'ère du Big Data, les consommateurs peuvent saisir pleinement les informations sur les produits grâce à des données d'analyse massives, comprendre en profondeur les attributs du produit et passer continuellement de l'implication situationnelle des produits à une implication à long terme. Par conséquent, les consommateurs généreront en permanence une perception interne positive et favoriseront l'apparition d'un comportement d'achat des consommateurs. De nos jours, il existe de nombreux consommateurs avec une forte implication des produits sur le marché. Ces acheteurs utiliseront la recherche et la comparaison d'informations sur le réseau, ainsi que d'autres évaluations d'utilisateurs, pour évaluer de manière exhaustive des facteurs tels que la performance des coûts des produits, les avantages de la marque et leurs propres besoins, et enfin prendre des décisions d'achat plus rationnelles[11].

2.4.5.2 Demande des consommateurs continue d'augmenter

La valeur des activités commerciales comprend la valeur d'usage du produit lui-même et la valeur d'expérience d'achat. De plus, parfois l'utilité apportée par l'expérience aux consommateurs joue un rôle déterminant dans la décision d'achat. Le Big Data favorise le marketing personnalisé du e-commerce, alors que les consommateurs demandent de plus en plus de services innovants et personnalisés[11].

2.4.5.3 Confiance des consommateurs dans les fonctions commerciales des médias sociaux a augmenté

De nos jours, les fonctions commerciales des médias sociaux sont continuellement explorées et utilisées, et la valeur commerciale est de plus en plus importante. Grâce aux médias sociaux, les entreprises peuvent maîtriser des informations personnelles de plus en plus complètes sur les consommateurs, afin de pouvoir analyser avec précision leurs préférences personnelles, leurs habitudes et d'autres informations, afin de mieux répondre aux besoins profonds des consommateurs ou d'exploiter les besoins potentiels des consommateurs. Les entreprises peuvent analyser les habitudes, les croyances et les préférences des gens par le biais des médias sociaux, et peuvent être précises dans une certaine mesure, formant ainsi une relation intime presque invasive avec les consommateurs et répondant mieux aux besoins profonds des consommateurs. Et les consommateurs s'attendent également à ce que leurs besoins soient de plus en plus pris en compte, découverts et satisfaits, de sorte qu'ils font également de plus en plus confiance et soutiennent la promotion commerciale des médias sociaux[11].

2.4.6 Dangers et potentiel du Big Data

Les mégadonnées peuvent être particulièrement utiles. De plus, il est capable de produire des informations plus précises, opportunes et détaillées que celles glanées à partir de sources de données plus traditionnelles. Dans le monde en développement, où les données sont relativement rares, le Big Data est une opportunité en plein essor. Par exemple, des entreprises telles que le projet de Nairobi exploitent les données des téléphones portables et d'autres types d'appareils de communication mobile pour comprendre les besoins des pauvres et pour aider à développer et suivre des programmes et des politiques destinés à répondre à ces besoins. Dans le même temps, le Big Data n'est pas exempt de problèmes. Les données provenant des médias sociaux sont également sujettes à des biais. Les utilisateurs de services de partage de localisation, par exemple, sont généralement des hommes à revenu faible à moyen âgés de 19 à 29 ans. Comme c'est le cas pour tout type de données, il est essentiel de comprendre l'étendue et la nature du biais dans les données et, si possible, de le corriger. Dans le cas contraire, son utilisation pourrait conduire à la mise en place de politiques inappropriées ou inéquitables. Il est également important de reconnaître que les techniques statistiques traditionnelles, qui supposent la

normalité et le caractère aléatoire, peuvent ne pas convenir aux mégadonnées. L'utilisation du traitement automatique du langage naturel et des méthodes d'apprentissage automatique pour l'analyse des sentiments, sans tenir compte de ces types de facteurs et tendances contextuels sous-jacents, peut produire des résultats qui peuvent ne pas coïncider correctement avec ou ne pas représenter avec précision les véritables intérêts et valeurs des parties prenantes concernées, y compris le public. Par conséquent, sortir des mots et des phrases de leur contexte dans un corpus de texte peut également conduire à des résultats et des conclusions trompeurs. La protection de la vie privée est une autre préoccupation soulevée par rapport à la collecte et à l'utilisation des mégadonnées. Un excellent exemple de la façon dont les mégadonnées portent potentiellement atteinte à la vie privée des individus est le suivi sur Internet. Des entreprises comme Double Click et Google AdSense, collectent d'énormes quantités d'informations détaillées et personnellement identifiables sur le comportement de navigation et les activités numériques des individus sur le Web, qu'ils utilisent ensuite à des fins publicitaires et autres. Le « retrait » n'est pas toujours une solution viable ou même pratique. Bien que la plupart des navigateurs aient en fait des paramètres de confidentialité, ces ajustements nécessitent un certain niveau de connaissances en matière d'Internet et de connaissances technologiques [14].

2.4.7 Big Data et l'IoB

Le cloud et les Big Data sont considérablement renforcés par la croissance de l'internet des objets. L'Internet des objets se rapporte à un système composé de capteurs installés tant dans les objets de tous les jours que dans les outils de production qui collectent des données en continu[6].

Lorsque l'Internet des objets (IoT) a introduit le concept de collecte de données via la connectivité des appareils en ligne, les gens ont plaisanté sur les robinets du gouvernement et « Big Brother » surveillant chacun de nos mouvements. Même avant l'Internet des objets (ou Internet, d'ailleurs), les gens partageaient des histoires de mondes dystopiques où l'intelligence artificielle et la vie des machines avaient déplacé l'humanité. Et pourtant, nous voici des années plus tard avec des assistants virtuels nommés Alexa et Echo ornant nos bureaux et nos tables basses ; les appareils portables qui suivent et transmettent nos données biométriques ; et des appareils électroménagers intelligents qui peuvent être automatisés à partir d'une application

et servir également de paniers de commerce électronique. Nous savourons à quel point il est agréable de pouvoir synchroniser nos appareils et de les faire tous se parler.

C'est peut-être de l'apprentissage automatique, ou peut-être n'est-ce rien de plus que l'anthropomorphisme du code binaire et des circuits, faisant écho à l'infusion littéraire des qualités humaines dans les animaux et l'abstrait. Quelle que soit votre vision de l'écosystème IoT, l'Internet of Behavior (IoB) était un résultat inévitable du Big Data et de notre volonté d'en tirer le meilleur parti[9].

2.4.8 Big Data et l'IoT

L'Internet des objets est ancré dans le monde technologique d'aujourd'hui et fait déjà l'actualité dans de nombreux secteurs. Les capteurs IoT peuvent être trouvés dans des appareils allant des smartphones et des boîtes noires aux voitures et aux navires, ce qui souligne le besoin d'analyse de données.

Il est juste de dire que si l'IoT est un phénomène relativement nouveau, il existe déjà de nombreuses opportunités. Cependant, cela ne veut pas dire que cette sphère de la technologie est sans défis, car il y a beaucoup de choix. Le plus difficile de tous doit être le fait que l'IoT crée des masses de données qui n'ont nulle part où aller. En effet, selon un rapport de Cisco, l'Internet des objets créera 400 zégaoctets de données par an d'ici 2018. Par conséquent, les organisations doivent décider comment elles vont gérer et donner un sens au grand volume de données qu'elles collectent à partir de l'IoT. Il y a, pourrait-on dire, un grand défi avec le Big Data.

Le Big Data aura un rôle important dans l'efficacité de traitement des informations et permettra ainsi aux développeurs d'IoT, une optimisation de ces outils afin d'élargir la perspective d'évolution actuelle, Le Big Data devrait permettre d'analyser en temps réel les données générées par les IOT et ainsi, optimiser l'usage de cette technologie. Pour ce faire, le Big Data procède en 4 étapes :

- La collecte des données générées par les IoT en suivant les 3 principes premiers du Big Data : vitesse, volume et variété.
- Le stockage des données dans des fichiers au sein de la base de données du Big Data.
- L'analyse des données par des systèmes analytiques complexes et performants, tels que

Spark ou encore Hadoop.

- La mise en place de rapport des données analysées.

L'interaction entre IoT et Big Data n'est cependant pas à sens unique. L'IoT pourrait également beaucoup apporter au Big Data. Plus les IoT auront de l'importance dans la vie quotidienne et celle de la ville, plus les développeurs seront demandeurs de plus grande capacité en termes de Big data et plus ce business prendra de l'ampleur. Il sera ainsi important d'améliorer les technologies de stockage de données pour pouvoir développer des systèmes capables de traiter encore plus de données. Cette interaction pourrait ainsi permettre une croissance technologique dans les deux domaines simultanément [10]

2.5 Machine learning

2.5.1 Définition du machine learning

Le machine learning (ML) est une forme d'intelligence artificielle (IA) qui est axée sur la création de systèmes qui apprennent, ou améliorent leurs performances, en fonction des données qu'ils traitent. L'intelligence artificielle est un terme large qui désigne des systèmes ou des machines simulant une forme d'intelligence humaine. Le machine learning et l'IA sont souvent abordés ensemble et ces termes sont parfois utilisés de manière interchangeable bien qu'ils ne renvoient pas exactement au même concept. Une distinction importante est que, même si l'intégralité du machine learning repose sur l'intelligence artificielle, cette dernière ne se limite pas au machine learning.

Aujourd'hui, nous utilisons le machine learning dans tous les domaines. Lorsque nous interagissons avec les banques, achetons en ligne ou utilisons les médias sociaux, des algorithmes de machine learning entrent en jeu pour optimiser, fluidifier et sécuriser notre expérience. Le machine learning et la technologie qui l'entoure se développent rapidement, et nous commençons seulement à entrevoir ses capacités [24].

2.5.2 Types de machine learning

Il existe trois types principaux d'apprentissage automatique : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage semi-supervisé

2.5.2.1 Machine learning supervisé

Le machine learning supervisé est une technologie élémentaire mais stricte. Les opérateurs présentent à l'ordinateur des exemples d'entrées et les sorties souhaitées, et l'ordinateur recherche des solutions pour obtenir ces sorties en fonction de ces entrées. Le but est que l'ordinateur apprenne la règle générale qui mappe les entrées et les sorties. Les algorithmes de machine learning supervisé sont les plus couramment utilisés. Avec ce modèle, un data scientist sert de guide et enseigne à l'algorithme les conclusions qu'il doit tirer. Tout comme un enfant apprend à identifier les fruits en les mémorisant dans un imagier, en apprentissage supervisé, l'algorithme apprend grâce à un jeu de données déjà étiqueté et dont le résultat est prédéfini. Comme exemples de machine learning supervisé, on peut citer des algorithmes tels que la régression linéaire et logistique, la classification en plusieurs catégories et les machines à vecteurs de support

2.5.2.2 Machine learning non supervisé

Dans le cadre du machine learning non supervisé, l'algorithme détermine lui-même la structure de l'entrée (aucune étiquette n'est appliquée à l'algorithme). Cette approche peut être un but en soi (qui permet de découvrir des structures enfouies dans les données) ou un moyen d'atteindre un certain but. Cette approche est également appelée « apprentissage des caractéristiques ». Le machine learning non supervisé utilise une approche plus indépendante dans laquelle un ordinateur apprend à identifier des processus et des schémas complexes sans un quelconque guidage humain constant et rigoureux. Le machine learning non supervisé implique une formation basée sur des données sans étiquette ni résultat spécifique défini. Pour continuer avec l'analogie de l'enseignement scolaire, le machine learning non supervisé s'apparente à un enfant qui apprend à identifier un fruit en observant des couleurs et des motifs, plutôt qu'en mémorisant les noms avec l'aide d'un enseignant. L'enfant cherche des similitudes entre les images et les sépare en groupes, en attribuant à chaque groupe sa propre étiquette. Comme exemples d'algorithmes de machine learning non supervisé, on peut citer la mise en cluster de k-moyennes, l'analyse de composants principaux et indépendants, et les règles d'association.

2.5.2.3 Machine learning par renforcement

Dans le machine learning par renforcement, un programme informatique interagit avec un environnement dynamique dans lequel il doit atteindre un certain but, par exemple conduire un véhicule ou affronter un adversaire dans un jeu. Le programme-apprenti reçoit du feedback sous forme de « récompenses » et de « punitions » pendant qu'il navigue dans l'espace du problème et qu'il apprend à identifier le comportement le plus efficace dans le contexte considéré [24].

2.5.3 IoT et le machine learning

L'IoT et le Machine Learning ont fait lentement leurs chemins dans tous les aspects de notre vie. Qu'il s'agisse de téléviseurs intelligents, de réfrigérateurs, de thermostats ou de cafetières intelligentes, les dispositifs IoT ont infiltré la vie quotidienne. Ils gagnent lentement la reconnaissance du grand public avec le début de la nouvelle année. En combinaison avec les assistants-domestiques comme Google Home et Alexa d'Amazon Echo qui ouvrent la voie à la Machine Learning, cette avancée technologique rend notre vie plus facile, efficace et améliorée. Toutefois, ces appareils ne sont pas considérés comme idéalement automatisés. En effet, ils nécessitent encore beaucoup d'intervention manuelle pour l'attribution des tâches. Tous ces appareils sont programmés pour suivre un ensemble particulier de commandes d'instruction. Une saisie manuelle doit donc se faire pour leur fournir des instructions personnalisées. Ils n'interagissent pas entre eux

2.5.4 Comment le machine Learning s'applique-t-elle aux données de l'IoT ?

Selon une étude, il y aura plus de 55 milliards d'appareils IoT d'ici 2025[33]. Le machine learning pour les capacités prédictives s'intègre désormais à la plupart des plateformes IoT industrielles. On le trouve effectivement dans Microsoft Azure IoT, Amazon AWS IoT ou Google Cloud IoT Edge. Par exemple, si un capteur détecte une chaleur ou une vibration excessive, il déclenche une alerte. Si ce même capteur est connecté à Internet, alors les données rencontrées par le capteur peuvent également être exploitées. L'objectif étant d'en tirer plus d'enseignements et effectuer des analyses pour une utilisation future. L'intelligence artificielle (IA) et

l'Internet des objets (IoT) séparément sont tous deux puissants. Toutefois, la combinaison des deux technologies émergentes présente d'immenses avantages pour les entreprises désireuses de réaliser une véritable transformation numérique [26].

2.5.5 IoT et Machine Learning : Avantages de la combinaison des deux

Le Machine Learning (ML) et l'Internet des objets (IoT) séparément sont tous deux puissants. Toutefois, la combinaison des deux technologies émergentes présente d'immenses avantages pour les entreprises désireuses de réaliser une véritable transformation numérique [26].

2.5.6 L'impact du Machine learning sur l'IoT

L'IoT développé et renforcé par le machine learning multiplie efficacement l'aide et l'influence de l'apparition des entreprise la possibilité révolutionnaire de cet échange qui s'empare de ces technologies intégrales. En réalité, l'IA (Intelligence Artificielle) est aujourd'hui un élément complet du succès. Le Machine Learning présente divers avantages, tels que la planification de systèmes de capteurs, l'offre d'analyses en temps réel, l'augmentation de la sécurité, la minimisation du flux de données et toutes les applications utilisateur personnalisées utilisant une grande quantité de données sur Internet. Au-delà du Machine Learning, l'objectif principal est d'automatiser la création de divers modèles analytiques pour permettre aux algorithmes d'apprendre en continu à l'aide des connaissances disponibles. L'impact du Machine Learning sur les progrès de l'IoT joue un rôle essentiel dans l'amélioration des applications intelligentes réelles [23]

2.6 Conclusion

Avec le développement de la technologie d'analyse des mégadonnées dans l'économie de réseau moderne, de nombreuses plateformes de réseau ou entreprises participant au commerce électronique recueilleront des informations personnalisées sur le comportement des consommateurs et utiliseront la technologie d'analyse et de traitement des mégadonnées pour intégrer et extraire des informations efficaces, et faire des recommandations ciblées aux consommateurs

avec précision. Sous l'attrait de ce mode de promotion et de ces méthodes de commercialisation "sur mesure", il est plus facile de stimuler l'intérêt des consommateurs et même de modifier les préférences de la demande des consommateurs, améliorant ainsi considérablement l'effet de substitution des produits connexes. Dans le même temps, alors que l'ère des mégadonnées rend les informations sur les produits quantitatives, transparentes et accessibles, les consommateurs peuvent également collecter indépendamment des informations pertinentes sur les produits visés, lire les conclusions pertinentes de l'analyse des mégadonnées, réaliser une consommation rationnelle et atteindre la valeur d'utilité attendue pour le la plus grande mesure. L'évolution des technologies numériques est marquée à la fois par de grandes opportunités et des défis complexes pour les entreprises. Compte tenu de l'intérêt croissant pour l'IoT et le Big Data, cette étude visait à présenter ces notions en matière de numérisation des entreprises et à enrichir les théories actuelles. Ce qui nous a permis de traiter ces notions et de déduire la relation qui est entre elles pour bien connaître les termes de notre thème.

Chapitre 3

Etat de l'art sur l'internet des comportements

3.1 Introduction

Dans ce chapitre nous allons présenter les différents travaux et études sur les approches utilisant l'internet de comportement. En effet l'IoB permet de développer une compréhension approfondie des comportements utilisateurs dont chaque entreprise a besoin, l'objectif de l'IoB est d'enregistrer, d'analyser, de comprendre et de répondre à toutes les formes de comportement humain dans une façon qui permet aux gens d'être suivis et interprétés en utilisant les progrès technologiques en développement et les progrès dans algorithmes d'apprentissage automatique. Nous allons introduire les défis les plus représentatifs qui seront utilisés pour comparer les approches étudiées telles que la méthodologie, le modèle de résolution et l'évolutivité. Nous donnons également les avantages et les limites de chaque approche, et pour finir nous discuterons des approches présentées dans ce chapitre en soulignant leurs limites.

3.2 L'Internet des comportements (IoB) et son rôle dans les services clients

Javaid et al [27] propose une approche pour l'amélioration de la qualité des services clients, L'approche commence par la collecte de données à partir de toutes les sources pertinentes, à savoir les réseaux sociaux, les sites Web, les appareils basés sur la télévision, etc., puis procède

à l'analyse de ces données collectées. Les faits basés sur les données ont ensuite été analysés de manière critique, projetés et combinés de manière approfondie pour résumer en une interférence significative, En suite la phase semi-finale et conclusive consistent à mettre en œuvre le front de collecte et d'analyse des données pour la réalisation de l'IoB pour la satisfaction du client dans tous les services concernés. Les données peuvent être obtenues à partir de plusieurs points de vente différents, des données commerciales sur les consommateurs aux médias sociaux, et davantage de données peuvent être accessibles à tous.

Cela combine diverses technologies de reconnaissance visuelle, de surveillance de localisation et de données étendues, qui se concentrent directement sur la personne et relient les données résultantes aux événements comportementaux. IoB concerne la façon dont les données peuvent être mieux comprises et comment elles peuvent produire et vendre de nouveaux biens à partir de la psychologie humaine. Cela a conduit à de nouvelles connaissances sur les preuves de l'IoT grâce à la recherche comportementale et à la psychologie. Avec une grande quantité de données, le mécanisme de distribution pour vendre des biens et des services aux personnes intéressées peut être stratégiquement mis en œuvre.

3.3 L'internet du comportement comme espace de transformation organisationnelle avec intelligence chorégraphique :

Stary et al [38] ont présenté l'internet des comportements comme un espace de transformation organisationnelle avec intelligence chorégraphique, Les systèmes IoB nécessitent une représentation de modèle (appelée jumeau numérique). les systèmes IoB sont considérés comme des systèmes adaptatifs complexes et analysés en fonction des flux de valeur entre les entités encapsulant le comportement. La carte résultante peut être affinée du point de vue de la fonction et du comportement de communication. Ainsi, la gestion des processus métier orientée sujet (SBPM) et son schéma de représentation chorégraphique jouent un rôle crucial : les modèles S-BPM enrichis constituent la base de l'ingénierie intégrée à la conception. les étapes de conception :

— **Représentation et analyse de la chaîne de valeur :** La recherche d'artefacts basés

sur des entités de comportement et leur interaction. Ce type d'entité constitue l'espace de conception de transformation, conçu pour mettre en oeuvre les applications IoB et les relier à l'analyse prédictive. L'intervention méthodologique est basée sur la connaissance opérationnelle de l'entreprise et sa représentation structurée des flux de valeur entre les parties prenantes impliquées, et entre les systèmes de soutien et les parties prenantes. La VNA(Value Network Analysis) se veut un instrument de développement au delà de l'ingénierie, car elle vise à comprendre la dynamique organisationnelle, et donc à gérer les connaissances structurelles dans une perspective de recherche de valeur, pour les parties prenantes individuelles et l'organisation dans son ensemble.

- **Raffinement orienté sujet et achèvement de l'exécution** : l'affinement des représentations de conception, telles que la carte BeE, vers des modèles numériques de systèmes IoB servant de référence pour l'ingénierie. L'approche présentée affine les cartes BeE du point de vue de la fonction et du comportement de communication, en utilisant la représentation chorégraphique et le schéma d'ingénierie de la gestion des processus d'affaires orientée sujet. Il permet d'incarner les cartes BeE et d'affiner les composants (socio-techniques) impliqués, générant ainsi des jumeaux numériques. Capacités de modélisation et d'exécution orientées sujet considèrent les systèmes comme des ensembles de sujets en interaction. Les sujets sont définis comme l'encapsulation du comportement. Ils correspondent aux Behavior-encapsulating Entities (BeEs) définies pour l'analyse des flux de valeur.
- **Adaptation dynamique et analyse prédictive** : Après avoir affiné BeEs, l'adaptation dynamique et la prédiction du comportement peuvent être abordées. L'adaptation dynamique est basée sur un déclencheur, tel qu'un résultat de l'exécution d'une fonction ou d'un signal de capteur, qui nécessite une spécification de comportement spéciale, Le déclencheur d'adaptation dynamique indépendant de sa mise en oeuvre peut transporter certaines données comme charge utile.

3.4 Internet du comportement et systèmes d'IA explicables pour influencer le comportement de l'IoT

Elayan et al [19] ont proposé un scénario de cas d'utilisation d'un système de consommation d'énergie électrique domestique qui vise à influencer le comportement de consommation d'électricité d'une maison en contrôlant automatiquement la quantité de puissance active globale. Le flux de travail du système dans la montre que le courant passera par un moniteur d'énergie et un dispositif de contrôle (EMC) pour connaître la quantité d'électricité de chaque appareil IoT en utilisant. Après cela, les données seront transmises de l'EMC à une base de données cloud pour être stockée et normalisée puis utilisée un modèle d'IA pour prédire la puissance active globale qui peut être consommée dans l'heure qui suit. Le résultat de la prédiction sera transmis à un décideur et à un applicateur pour déterminer comment la quantité d'énergie que les appareils doivent utiliser, puis les résultats sera renvoyé au contrôle CEM la consommation dans le l'heure suivante si elle est supérieure à un certain seuil afin de réduire. Ce modèle présente un cadre composé de toutes les technologies nécessaires pour former un modèle IoB/XAI qui réduira la consommation d'énergie. Dans un premier temps, l'algorithme LSTM sera utilisé pour construire un modèle de prédiction de la consommation d'énergie, puis le processus de consommation d'énergie sera contrôlé.

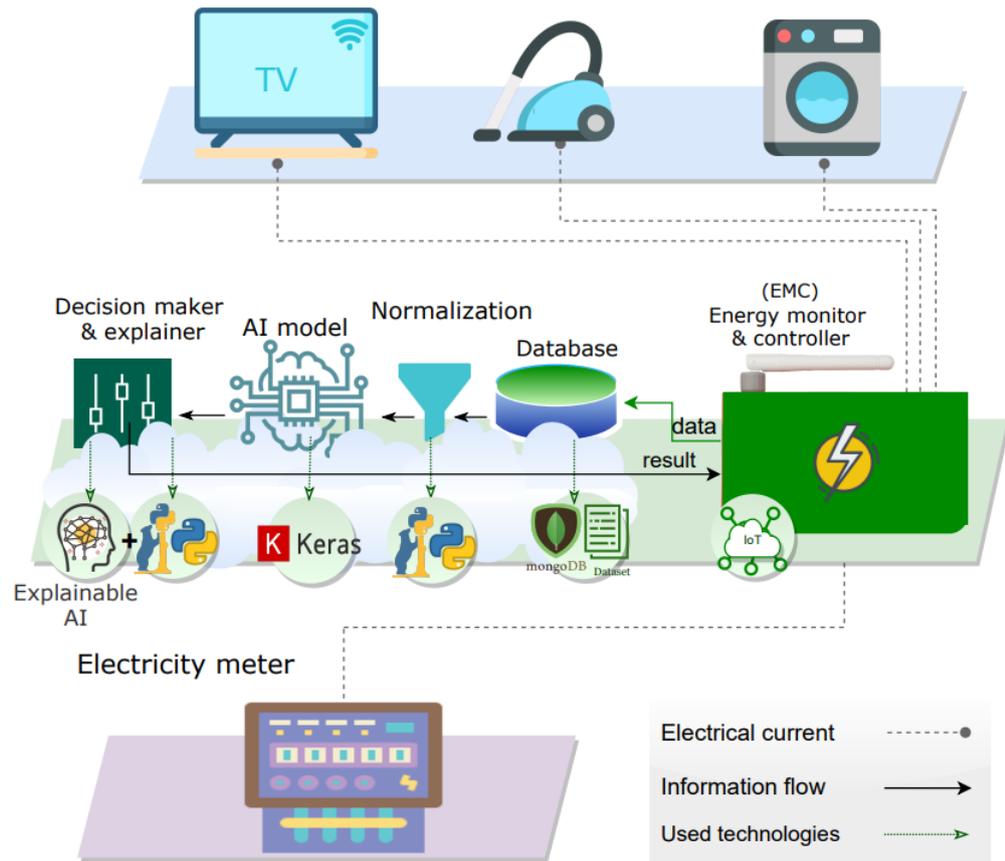


FIGURE 3.1 – Flux de travail du système IoB pour influencer le comportement de l'IoT

3.5 Conception de systèmes Internet des comportements

Moghaddam et al [25] avaient appliqué leur modèle IoB concret au système de gestion des files d'attente des Galeries des Offices à Florence, en Italie. Ils capturent de manière dynamique le comportement de mobilité humaine dans les Offices afin de fournir un ensemble de services qui facilitent la réservation des billets et éliminent les longues files d'attente. En plus du système RFID, ils ont utilisé plus de 500 caméras de vidéo surveillance déjà installées et utilisé des compteurs de personnes pour obtenir des données plus précises. Les données sur la mobilité humaine sont envoyées à un composant de traitement et de stockage, un cloud, où elles sont analysées. Le contrôleur MVC, le service de données et le référentiel. Le contrôleur concerne principalement une interface pour l'acquisition des données. Les données collectées sont stockées dans un SGBD MySQL. Le solveur est développé en Python en appelant le solveur ILP. Le SGBD fournit au composant interne du solveur des statistiques sommaires et en temps réel. Le système IoT possède un noyau algorithmique. Ces données passent par l'application web Spring vers le

solveur pour calculer les créneaux horaires disponibles pour les visiteurs toutes les 15 minutes. Les actionneurs IoT sont les kiosques et les tableaux de bord, qui fournissent une connaissance de la situation aux visiteurs et aux opérateurs. Les créneaux horaires disponibles sont présentés aux visiteurs dans les kiosques où ils peuvent réserver leurs billets. Ce système préliminaire IoB permet une réservation facile et flexible afin d'améliorer la qualité de l'expérience des visiteurs.

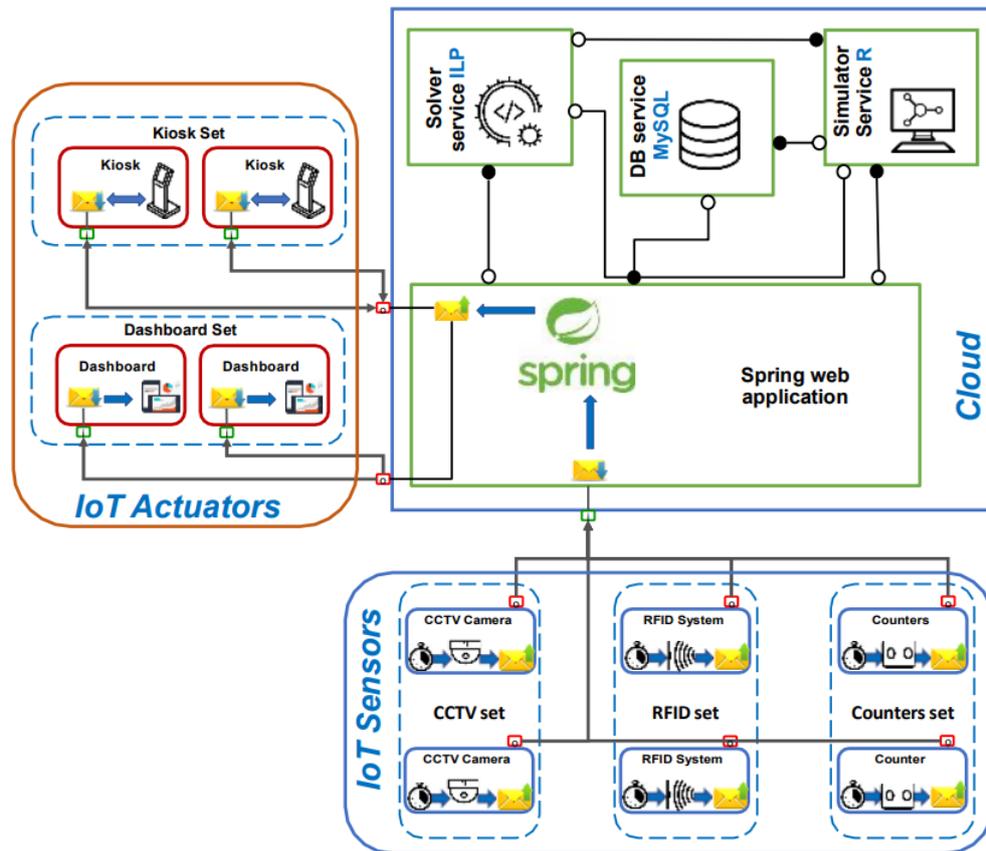


FIGURE 3.2 – Architecture logicielle de gestion des files d'attente des Galeries des Offices à Florenc

3.6 Etude de la pollution de l'air a Pékin affecté ou non sur les activiés de vie quotidienne des personne basé sur l'intenet des comportemnts

Zhang et al [18] ils ont proposé d'étudier le comportement des habitants de Pékin, Si la pollution de l'air affecte-t-elle ou non sur leur comportement quotidien en utilisant l'Internet des comportements, Cet étude prend en compte trois modules : (entrée, traitement et sortie), Tout d'abord ils doivent identifier les sources de données IoB appropriées qui peuvent être

utilisées pour suivre le comportement humain de masse, comme les mouvements humains : enregistrement détaillé des appels CDR et décrire l'heure de l'appel, la station de base la plus proche et les informations de localisation, Il est utile de différencier les points d'intérêt habituels PIH des gens associés à l'activité de la vie quotidienne, ils classent cela comme un lieu spécifique LSAS, Ensuite, ils doivent étudier comment les changements environnementaux, par exemple la qualité de l'air, affectent le comportement des humains dans LSAS. Le flux de travail est divisé en deux parties principales (flux de travail 1 et 2) le premier flux basé sur la première catégorie des activités de la vie quotidienne sont représentées par cinq PIH fixes qui sont des LSAS : (1) Tourisme utiliser les zones touristiques comme point d'intérêt, (2) Restauration utilise les restaurants comme point d'intérêt, (3) Rester à la maison, (4) Voyager en mode de transport (Voyager en bus utilise les arrêts de bus comme points d'intérêt et voyager en métro utilise les stations de métro comme points d'intérêt). Le deuxième flux de travail basé sur la deuxième catégorie des activités de la vie quotidienne autre que le lieu spécifique LSAS, Ce qui représente qu'un seul AVQ Voyage en mode de transport (Voyage en vélo et marche à pied). Pour le premier flux de travail l'unité de traitement est utilisée pour construire des modèles à effet fixe FEM pour détecter la relation entre les AVQ et pollution de l'air, Pour le deuxième flux construire le modèle linéaire régression multivariable prédire combien de personnes ont tendance à faire du vélo ou à marcher et si elles réduisent la distance voyager ou changer de mode de transport. le module de traitement est utilisé pour construire des Modèles à effet fixe FEM pour détecter la relation entre les AVQ et pollution de l'air, où tout comportement impactant les indices (un coefficient dans un FEM pour la densité de population. Ici, Y_{it} est la variable dépendante, qui change en fonction du temps et de l'individu. a_0 est la constante, tandis que i est l'effet individuel qui est invariant dans le temps. De plus, avant d'utiliser les CDR pour extraire les valeurs de densité des personnes, ils analysent si la résolution spatiale de la distribution calculée des personnes est suffisamment précise pour extraire des valeurs représentant les quatre AVQ. Ils concluent qu'à l'exception du restaurant PIH/LPSA AVQ, Les quatre autres types de PIH peuvent être utilisés pour représenter les LPSA AVQ sans la stratégie 1. Ainsi, pour diminuer l'erreur d'estimation lors du traitement de MEF pour le PIH du restaurant, Maintenant nous présentons la deuxième stratégie qui échantillonne les polygones de VORONOÏ qui comprend de vastes zones de restaurants. Enfin, le résultat estimé les changements dans les revenus des restaurants dus à la pollution de l'air sur la base des indices impactant le comportement. Afin

d'analyser de toutes les comportements des habitants de pékin ils ont trouvé que 4 effect : (1) la pollution de l'air a une impact négatif sur les personnes qui font le tourism après midi, (2) la pollution de l'air a une impact positif pour les personnes qui restent chez eux, le matin et en milieu de journée, (3) la pollution de l'air réduit le désir des gens d'aller au restaurant pour le déjeuner mais beaucoup moins de soir et (4) la pollution de l'air à mesure que la qualité de l'air se détériore, les gens ont tendance à moins marcher et à faire du vélo e a se déplacer davantage en autobus ou en métro.

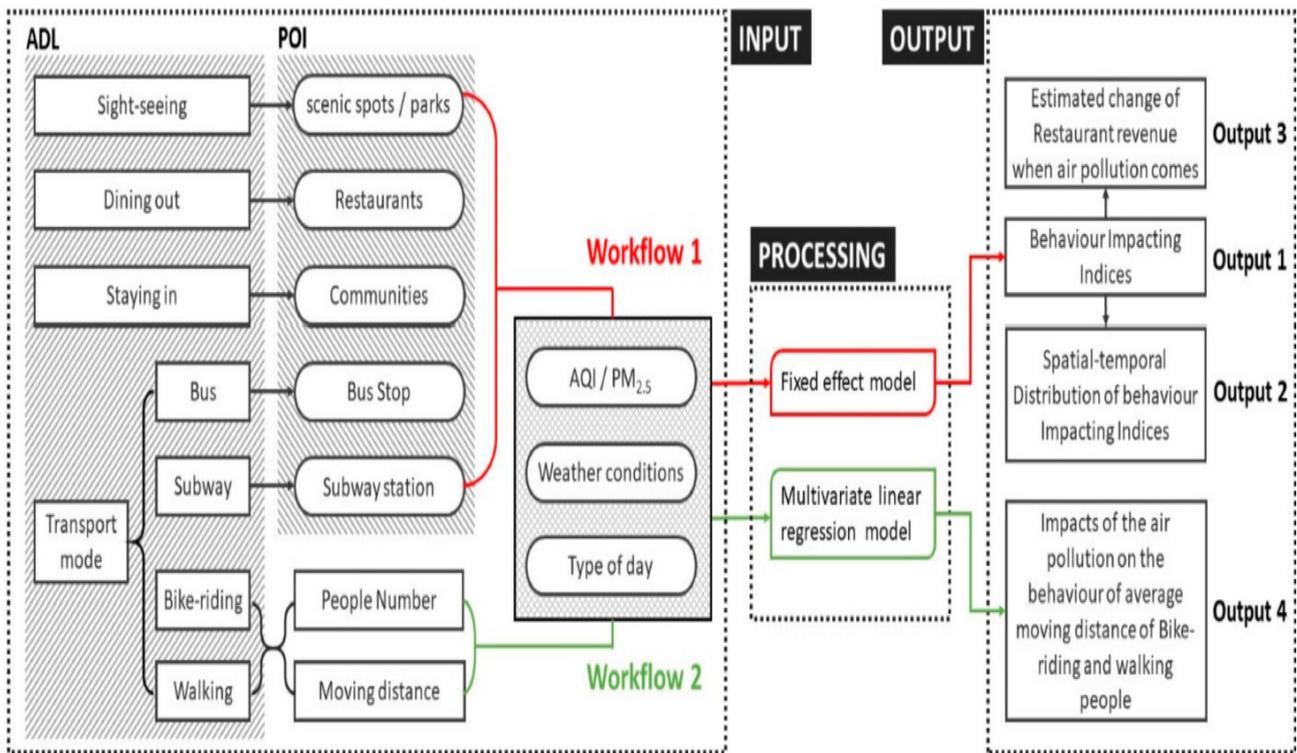


FIGURE 3.3 – Le framwork IoB pour Etude de la pollution de l'air a Pékin

3.7 Analyse comparative

Cette section vise à présenter une étude comparative entre les différents approches présentées. Un tableau récapitulatif est dressé pour comparer les approches selon les critères identifiés, et les compare en termes de de Méthodologie, et Modèle de résolution, est si le est Système adaptatif, et la Difficulté de réalisation.

3.7.1 Critères de comparaison

- **Méthodologie** : Étude systématique, par observation de la pratique scientifique, des principes qui la fondent et des méthodes de recherche utilisées.
- **Modèle de résolution** : Fait référence aux techniques utilisées pour résoudre les problème tel que les méthodes de graphes, les techniques de l'IA, etc.
- **Système adaptatif** : Un système adaptatif est un système qui est complexe en ce sens qu'il s'agit d'un réseau dynamique d'interactions, mais dont le comportement de l'ensemble peut ne pas être prévisible en fonction du comportement des composants.
- **Difficulté de réalisation** : Sa difficultés de réalisation dans le monde réel.

3.7.2 Tableau Comparatif

Le tableau 3.1 résume les approches présentées dans ce chapitre classées selon les critères suivants : méthodologie, le modèle de résolution, Système adaptatif, et l'évolutivité.

Approches	Méthodologie	Modèle de résolution	Système adaptatif	Difficulté de réalisation	Avantages	Inconvénients
Javaid et al [27] 2020	La recherche fondamentale	L'apprentissage automatique	Non	Non	<ul style="list-style-type: none"> l'amélioration de la qualité des services 	<ul style="list-style-type: none"> De nombreux facteurs n'ont pas été pris en compte
Stary et al [38] 2020	Recherche basée sur la conception	l'apprentissage en profondeur	OUI	OUI	<ul style="list-style-type: none"> Un système capable de traiter tous les problèmes, aussi complexes soient-ils 	<ul style="list-style-type: none"> Un système complexe difficile à réaliser
Elayan et al [19] 2021	étude de cas sur la consommation d'électricité	l'apprentissage en profondeur (LSTM) et Explicateur de l'IA	NON	OUI	<ul style="list-style-type: none"> réduire la consommation d'énergie et améliorer le comportement des utilisateurs 	<ul style="list-style-type: none"> Contrôle total du système électrique et peut causer des problèmes à l'utilisateur
Moghaddam et al [25] 2021	Entretien semi structuré	Solveur ilp	OUI	OUI	<ul style="list-style-type: none"> obtenir une caractéristique statistique plus homogène sur le comportement des visiteurs 	<ul style="list-style-type: none"> Le système est difficile à réaliser, et toute erreur peut endommager le système
Zhang et al [18] 2021	étude de cas sur la pollution de l'air a Pékin	linear regression multiple variables et modèle à effet fixe	NON	NON	<ul style="list-style-type: none"> Méthode facile à appliquer qui ne nécessite pas beaucoup 	<ul style="list-style-type: none"> Méthode statique, utilisée des méthodes traditionnelle

TABLE 3.1 – Table de comparaison des approches utilisant IoB

3.8 Conclusion

Ce chapitre était destiné à représenter l'état de l'art sur les différentes techniques et méthodes employées sur l'internet des comportements. Les défis les plus représentatifs qui ont été utilisés pour comparer les approches étudiées telles que la méthodologie, le modèle de résolution et l'évolutivité. Nous tenons également compte des avantages et des limites de chaque approche. Dans le prochain chapitre, Nous exposerons notre travail, qui comporte deux parties, une partie étudier les données consommation l'énergie dans une maison, et une partie est de prévoir la consommation d'électricité.

Chapitre 4

Mise en œuvre et évaluation

4.1 Introduction

Dans la suite de l'ouvrage, nous présenterons les différents outils nécessaires lors de la processus de mise en œuvre du modèle. Ensuite, nous expliquerons toutes les étapes que nous avons suivies à partir en définissant les logiciels et matériels utilisés. Enfin, nous discuterons de notre cadre expérimental et les résultats.

4.2 Environnement de développement

4.2.1 Matériels

Dans notre projet, nous avons utilisé nos propres ordinateurs portables que nous considérons suffisamment puissants pour construire le modèle, nos ordinateurs ont comme spécifications :

Caractéristiques	Machine 1	Machine 2
Processeur	Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz	Intel(R) Core(TM) i7-1065G7 CPU @ 1.30GHz 1.50 GHz
RAM	8,00 Go	16,0 Go
Système d'exploitation	Ubuntu linux 20.04	Windows 11

TABLE 4.1 – Spécifications des ordinateurs utilisés

4.2.2 Logiciels

4.2.2.1 Langages de programmation et packages

Python : Python est un langage de programmation orienté objet. Créé et publié en 1991 par Guido van Rossum. Il fournit un travail plus rapide que d'autres programmes langues. Pour les programmeurs expérimentés et débutants, python leur permet de trouver la syntaxe propre et les structures d'indentation, il est facile à apprendre et à utiliser. Cette simplicité permet aux développeurs de se concentrer sur la résolution de problèmes. Sans prendre trop de temps à comprendre la syntaxe et le comportement du langage de programmation. En écrivant moins de code et en obtenant plus fait[17]. Python s'est avéré être l'un des langages de programmation les plus efficaces, etc. en IA, en particulier dans les solutions ML et DL. Il aide à obtenir de meilleurs résultats[35].

Numpy : Cette bibliothèque, dont le nom signifie Python numérique, constitue le noyau de de nombreuses autres bibliothèques Python qui en sont issues. En effet, numpy est la bibliothèque de base pour le calcul scientifique en Python, car elle fournit des structures de données et une haute effectuer des fonctions que le package de base de Python ne peut pas fournir[32], comme un multidimensionnel objet tableau, divers objets dérivés (tels que des tableaux masqués

et des matrices), et un assortiment de routines pour des opérations rapides sur des tableaux, en plus des mathématiques, logiques, manipulation de forme, tri, sélection, E/S (entrée/sortie), transformées de Fourier discrètes, il aide à l'algèbre linéaire de base, aux opérations statistiques de base, à la simulation aléatoire et bien plus encore Suite[16].

Pandas : Pandas est un package open source Python largement utilisé pour la science des données, l'analyse de données et les tâches d'apprentissage automatique. En tant que l'un des jeux de données les plus populaires packages, Pandas fonctionne bien avec de nombreux autres modules de science des données à l'intérieur de Python écosystème, il propose des structures de données et des opérations de manipulation de tables numériques, et des séries chronologiques. Le concept fondamental de ce package est le DataFrame, une structure bidimensionnelle structure de données tabulaire avec des étiquettes de ligne et de colonne. Les pandas offrent une haute performance propriété de la librairie NumPy pour la manipulation de données dans des tableaux ou en relationnel bases de données (bases de données SQL). En utilisant l'indexation avancée, il sera facile d'effectuer de nombreuses des opérations telles que le découpage en tranches, le remodelage, les agrégations et les sélections de sous-ensembles[16].

Matplotlib : Ce paquet est la bibliothèque Python la plus populaire actuellement pour produire graphiques et autres visualisations de données en 2D. l'analyse des données nécessite des outils de visualisation, c'est donc la bibliothèque qui convient bien à cet usage. matplotlib ressemble à une bibliothèque graphique qui vous permet de gérer par programmation les éléments graphiques afin que le graphique l'affichage peut être contrôlé dans son intégralité. les visualisations sont obtenues par codage avec cette bibliothèque, puis exporté dans les formats graphiques les plus courants (tels que PNG et SVG...) pour être utilisé dans d'autres applications, documentation, pages web, etc[16].

TensorFlow : Est une bibliothèque open-source, développée par Google's Machine Learning Intelligence organisation de recherche, dans le but de mener un apprentissage automatique et une analyse approfondie recherche sur les réseaux de neurones. Il combine l'algèbre computationnelle de l'optimisation de la compilation techniques, facilitant le calcul de nombreuses expressions mathématiques où le problème est le temps nécessaire pour effectuer le calcul. Il utilise Python pour fournir un API frontale pratique pour créer des applications avec le framework, tout en exécutant ces applications en haute performance. TensorFlow prend en charge les prévisions de production à échelle, avec les mêmes modèles utilisés pour l'entraînement[39].

Keras : Keras est une API (interface de programmation d'applications) de Deep Learning sous python qui permet la création de modèles rapidement et de manière simple. Deux API principales sont utilisés pour créer le modèle Keras[29].

Scikit-learn : Une bibliothèque d'apprentissage automatique open source écrite en Python facilement et intégrer rapidement des méthodes d'apprentissage automatique dans le code Python. sci-kit learn contient diverses méthodes de classification, de régression et de covariance. Estimation matricielle, dimensionnalité réduction, prétraitement des données, génération de problèmes de benchmarking. C'est disponible sur la plupart des systèmes d'exploitation et est facile à installer. La bibliothèque est constamment améliorée, étendu et largement utilisé dans de nombreuses applications commerciales[34].

4.2.2.2 Environnement de Développement Intégré IDE

Anaconda : Anaconda est une distribution gratuite et open-source de la programmation Python et R langages pour le calcul scientifique (science des données, applications d'apprentissage automatique, traitement de données à grande échelle, analyse prédictive, etc.)[3]. Il peut être facilement installé sur n'importe quel système d'exploitation tel que Windows, Linux et MAC OS. Il fournit plus de 1500 Python/R paquets. La distribution Anaconda fournit l'installation de Python avec divers IDE tels que comme Jupyter Notebook, Spyder, Anaconda prompt, etc. C'est donc un emballage très pratique solution que vous pouvez facilement télécharger et installer. Il s'installera automatiquement Python, quelques IDE de base et bibliothèques avec. Dans nos thèses, nous avons utilisé VSCode et python Pour construire notre modèle.

Jupyter Notebook : Jupyter Notebook est une application Web IDE gratuite développée en 2014. Elle permet le partage de code en direct, d'équations, de visualisations et de texte explicatif dans une seule extension de document .ipynb[30].

Visual Studio Code : Il s'agit d'un éditeur de code source léger mais puissant qui s'exécute sur un ordinateur de bureau et est disponible pour Windows, macOS et Linux. Il inclut une prise en charge intégrée de JavaScript, TypeScript et Node.js et dispose d'un riche écosystème d'extensions pour d'autres langages (tels que C++, C, Java, Python, PHP, Go) et des runtimes (tels que .NET et Unity)[40].

4.3 Métriques d'évaluation

4.3.1 RMSE (Racine de l'erreur quadratique moyenne)

L'erreur quadratique moyenne (RMSE) (erreurs de prédiction) est définie comme l'écart type des résidus. Les résidus sont une mesure de la distance entre les points de données et la régression ligne. En d'autres termes, il indique la concentration des données autour de la ligne de meilleur ajustement. En ordre pour valider les résultats expérimentaux, l'erreur quadratique moyenne est couramment utilisée en climatologie, prévision et analyse de régression[36]. Sa formule est donnée comme suit :

$$RMSE = \left[\sum_{i=1}^n (y_i - y_{gi})^2 / N \right]^{1/2} \quad (4.1)$$

Les symboles :

i : Variable i.

N : Nombre de points de données non manquants.

y_i : Observations réelles.

y_{gi} : Valeur prédite.

4.3.2 MAE (Erreur Absolue Moyenne)

L'erreur absolue est la quantité d'erreur dans nos mesures. C'est la différence entre les valeur mesurée et la « vraie » valeur. L'erreur absolue moyenne (MAE) est la moyenne de tous erreurs absolues[28]. Sa formule est donnée comme suit :

$$MAE = 1/N \sum_{i=1}^n |y_i - y_{gi}| \quad (4.2)$$

Les symboles :

i : Variable i.

N : Nombre de points de données non manquants.

y_i : Observations réelles.

y_{gi} : Valeur prédite.

4.4 Visualisation et analyse des données

4.4.1 Description des données

Pour pouvoir prédire les valeurs d'électricité consommées, un jeu de données reflétant la réalité est nécessaire. Celui que nous avons choisi est intitulé "Household Power Consumption" (HPC ou consommation d'énergie domestique en français)[13] contient les valeurs de consommation d'une région parisienne durant quatre ans. HPC est composé de sept variables qui sont :

Puissance active totale : Correspond à l'énergie effectivement transférée.

puissance réactive : Correspond à la puissance non utile.

Tension moyenne (volts) : Est la circulation du champ électrique le long d'un circuit électrique mesurée en volts.

Intensité du courant : Représente la vitesse à laquelle les électrons circulent dans un circuit électrique.

Énergie active pour la cuisine : Rlle contient principalement un lave-vaisselle, un four et un micro-onde.

Énergie active pour la Buanderie : Elle contient une machine à laver, un sèche-linge, un réfrigérateur et une lumière.

Énergie active pour les systèmes de climatisation : Elle correspond à un chauffe-eau électrique et un climatiseur.

D'autres composants peuvent avoir lieu dans ce jeu de données. Néanmoins, puisque ils ne sont pas dotés d'intelligence, leur problème réside dans le fait que la somme des variables (Buanderie , cuisine , systèmes de climatisation) est inférieure à la puissance active totale.

1	Date	Time	Global_active	Global_reacti	Voltage	Global_intens	Sub_metering	Sub_metering	Sub_metering_3
2	16/12/2006	17:24:00	4.216	0.418	234.840	18.400	0.000	1.000	17.000
3	16/12/2006	17:25:00	5.360	0.436	233.630	23.000	0.000	1.000	16.000
4	16/12/2006	17:26:00	5.374	0.498	233.290	23.000	0.000	2.000	17.000
5	16/12/2006	17:27:00	5.388	0.502	233.740	23.000	0.000	1.000	17.000
6	16/12/2006	17:28:00	3.666	0.528	235.680	15.800	0.000	1.000	17.000
7	16/12/2006	17:29:00	3.520	0.522	235.020	15.000	0.000	2.000	17.000
8	16/12/2006	17:30:00	3.702	0.520	235.090	15.800	0.000	1.000	17.000

FIGURE 4.1 – Ensemble de données sur la consommation électrique des habitations

4.4.2 pretraitement des données

L'ensemble de données montre la puissance active ainsi qu'une certaine segmentation de la puissance active par circuit primaire dans la maison, notamment la cuisine, la buanderie et la climatisation. Ce ne sont pas les seuls circuits de la maison. Une quatrième variable «sub metering 4 » de sous-comptage peut être créée en soustrayant la somme de trois variables de sous-comptage définies de l'énergie active totale comme suit :

$$submetering4 = (globalactivepower * 1000/60) - (submetering1 + submetering2 + submetering3) \quad (4.3)$$

Les symboles :

globalactivepower : Correspond à l'énergie effectivement transférée.

submetering1 : Rlle contient principalement un lave-vaisselle, un four et un micro-onde.

submetering2 : Elle contient une machine à laver, un sèche-linge, un réfrigérateur et une lumière.

submetering3 : Elle correspond à un chauffe-eau électrique et un climatiseur.

Néanmoins, cet ensemble de données est devenu une référence pour évaluer les approches de prévision de séries chronologiques et d'apprentissage automatique pour la prévision en plusieurs étapes, en particulier la prévision de puissance active.

Nous pouvons marquer toutes les valeurs manquantes indiquées par un caractère '?' avec une valeur NaN, qui est un flottant. Cela nous permettra de travailler avec les données comme un tableau de valeurs à virgule flottante plutôt que des types mixtes, ce qui est moins efficace. Nous pouvons maintenant créer une nouvelle colonne «sub metering 4 » contenant le reste du sous-comptage, en utilisant le calcul de la section précédente. Vous trouverez cidessous les cinq premières lignes de données (et l'en-tête) du fichier de données de consommation d'électricité du propriétaire.

datetime	Global_active_power	Global_reactive_power	Voltage	Global_intensity	Sub_metering_1	Sub_metering_2	Sub_metering_3	Sub_metering_4
2006-12-16 17:24:00	4.216	0.418	234.84	18.4	0.0	1.0	17.0	52.266670
2006-12-16 17:25:00	5.360	0.436	233.63	23.0	0.0	1.0	16.0	72.333336
2006-12-16 17:26:00	5.374	0.498	233.29	23.0	0.0	2.0	17.0	70.566666
2006-12-16 17:27:00	5.388	0.502	233.74	23.0	0.0	1.0	17.0	71.800000
2006-12-16 17:28:00	3.666	0.528	235.68	15.8	0.0	1.0	17.0	43.100000

FIGURE 4.2 – Ensemble de données sur la consommation électrique des ménages

4.4.3 Observation de la consommation d'électricité au fil du temps

La meilleure façon de comprendre une série chronologique multivariée (données) est de tracer des graphes linéaires. Nous pouvons commencer par en créer un séparé pour chacune des sept variables pendant quatre ans de janvier 2007 à janvier 2011. L'exemple complet est répertorié ci-dessous :

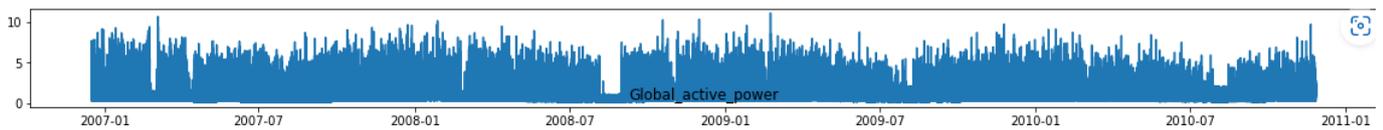


FIGURE 4.3 – La puissance active totale

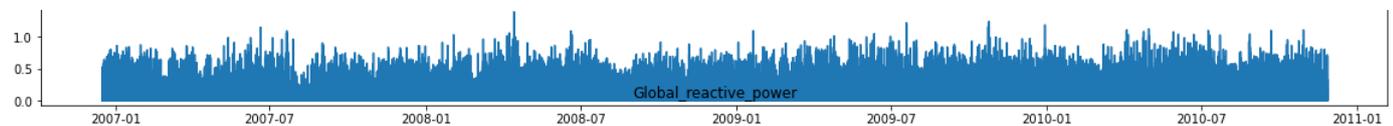


FIGURE 4.4 – La puissance réactive totale

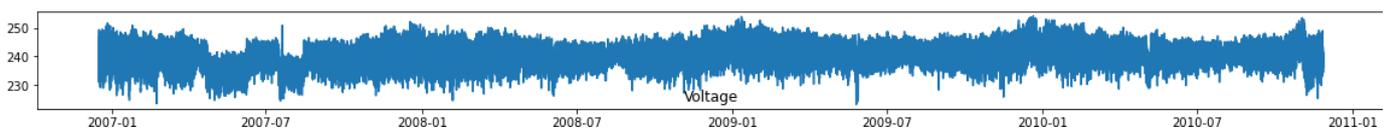


FIGURE 4.5 – Tension moyenne

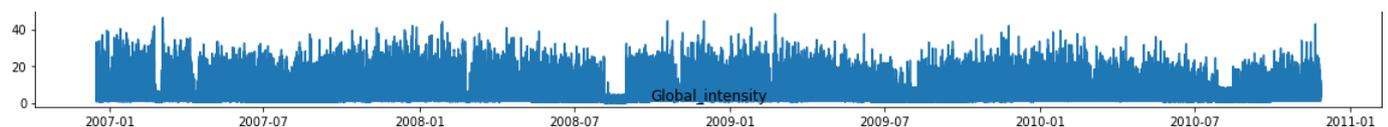


FIGURE 4.6 – Intensité moyenne du courant

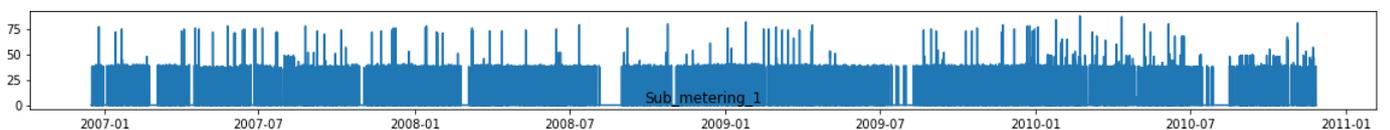


FIGURE 4.7 – Énergie active pour la cuisine



FIGURE 4.8 – Énergie active pour la Buanderie

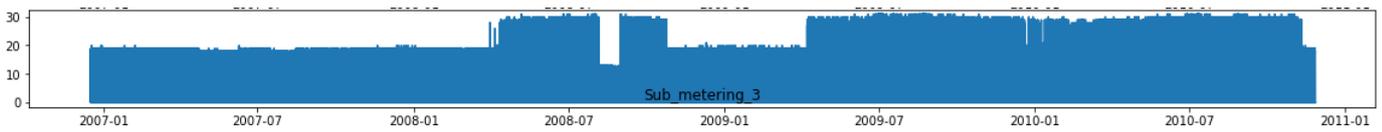


FIGURE 4.9 – Énergie active pour les systèmes de climatisation

Ces graphiques schématisant les métriques d'énergie d'un endroit particulier d'un ménage montrent différentes valeurs, entre évolution et dégradation, de la consommation d'électricité. Analysons comme premier cas les résultats obtenus par le sub metering 1 dans la cuisine : cet endroit qui dispose de divers appareils électroménagers a comme fonction principale la préparation des aliments et des plats. Cette fonction invariable est synonyme de dire que les accessoires de cuisine dont nous avons et nous aurons besoin sont toujours les mêmes, ce qui fait que les valeurs d'électricité consommée reste dans un niveau analogue.

Passons à un autre exemple, celui des mesures de la métrique "sub metering 3" pour les systèmes de climatisation que nous pouvons expliquer en 4 phases principales :

[2007-07, 2008-07] : De l'été 2007 à celui de 2008, la consommation d'électricité est trop faible vu qu'il n'est pas encore installé un système de climatisation.

[2008-07, 2009-01] : La consommation a évolué à ce stade puisqu'ils ont installé un nouveau système.

[2009-01, 2010-01] : En hiver, nous retombons à nouveau dans la dégradation de ces valeurs à cause de non emménagement du système de climatisation (une chaudière à titre d'illustration).

[2010-01, 2010-07] : De nouveaux appareils ont été installés impliquant l'augmentation de la consommation d'énergie dans le ménage.

4.4.3.1 Visualisation annuelle pour la variable la puissance active totale

Dans cette section, nous discuterons de la variable “La puissance active totale”. On commencera par la visualisation annuelle, on tracera un graphique pour chaque année, puis on présentera quelques observations que l’on peut déduire de ce dernier, la Fig 4.10 ci-dessous reprend les graphiques obtenus.

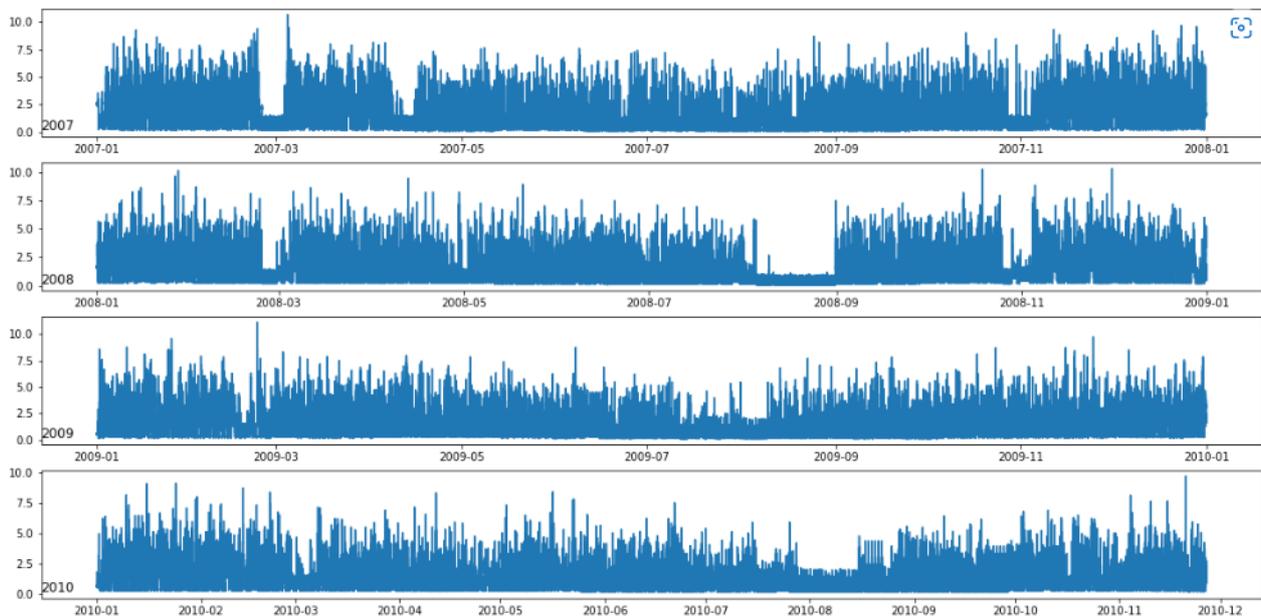


FIGURE 4.10 – Visualisation annuelle de l'énergie active

On peut commencer par regarder les similitudes sur la période de quatre ans, et les graphiques sont assez proches. Par exemple, nous avons remarqué que la consommation pendant les mois d'hiver était très élevée par rapport aux autres mois, de plus, nous avons remarqué que la consommation a diminué pendant les mois d'été (août) de 2008, 2009 et 2010, ce qui est logique car La plupart des gens partent en vacances pendant ces mois, et contrairement à l'année 2007, la consommation en août est normale, ce qui signifie qu'ils sont toujours à la maison. On peut remarquer une baisse de la consommation a la fin février et début mars en 2007 et 2008, qui peut être liée aux autres appareils qui n'ont pas été suivis, ce que l'on peut remarquer dans la Fig 4.9.

4.4.3.2 Visualisation mensuelle pour la variable la puissance active totale

Dans cette partie, nous tracerons des graphiques de la consommation mensuelle de l'année 2008 la Fig 4.11 ci-dessous reprend les graphiques obtenus :

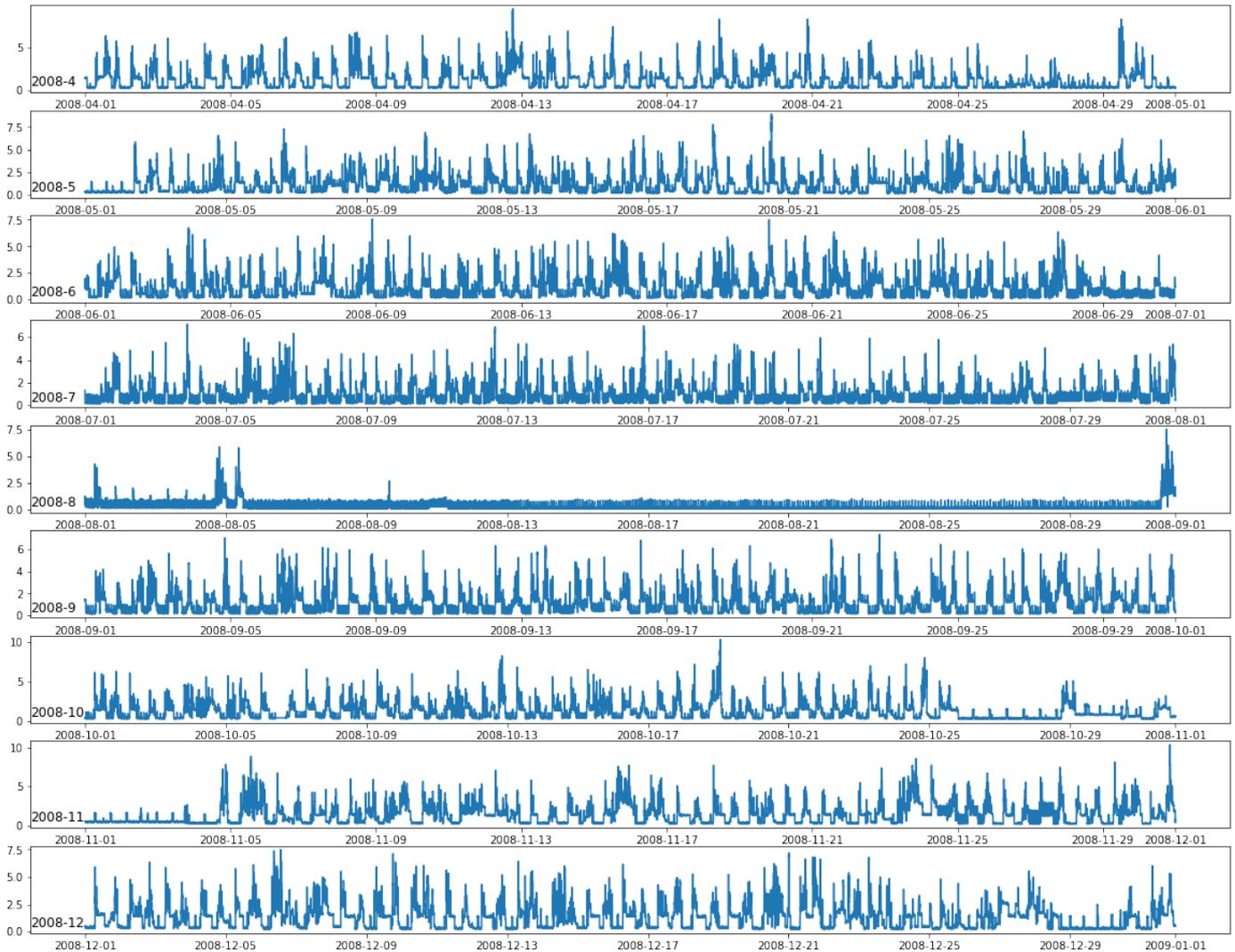


FIGURE 4.11 – Visualisation mensuelle de l'année 2008

On peut remarquer une baisse de la consommation en août, ce qui conforte notre opinion que c'était la période des vacances, et on remarque une augmentation directe de la consommation en fin de mois.

4.4.3.3 Visualisation quotidienne pour la variable la puissance active totale

Concentrons-nous maintenant sur l'analyse quotidienne. Dans cette partie, nous tracerons des graphiques de la consommation quotidienne d'électricité, à partir de janvier pendant 15 jours en 2008, comme indiqué la figure ci-dessous :

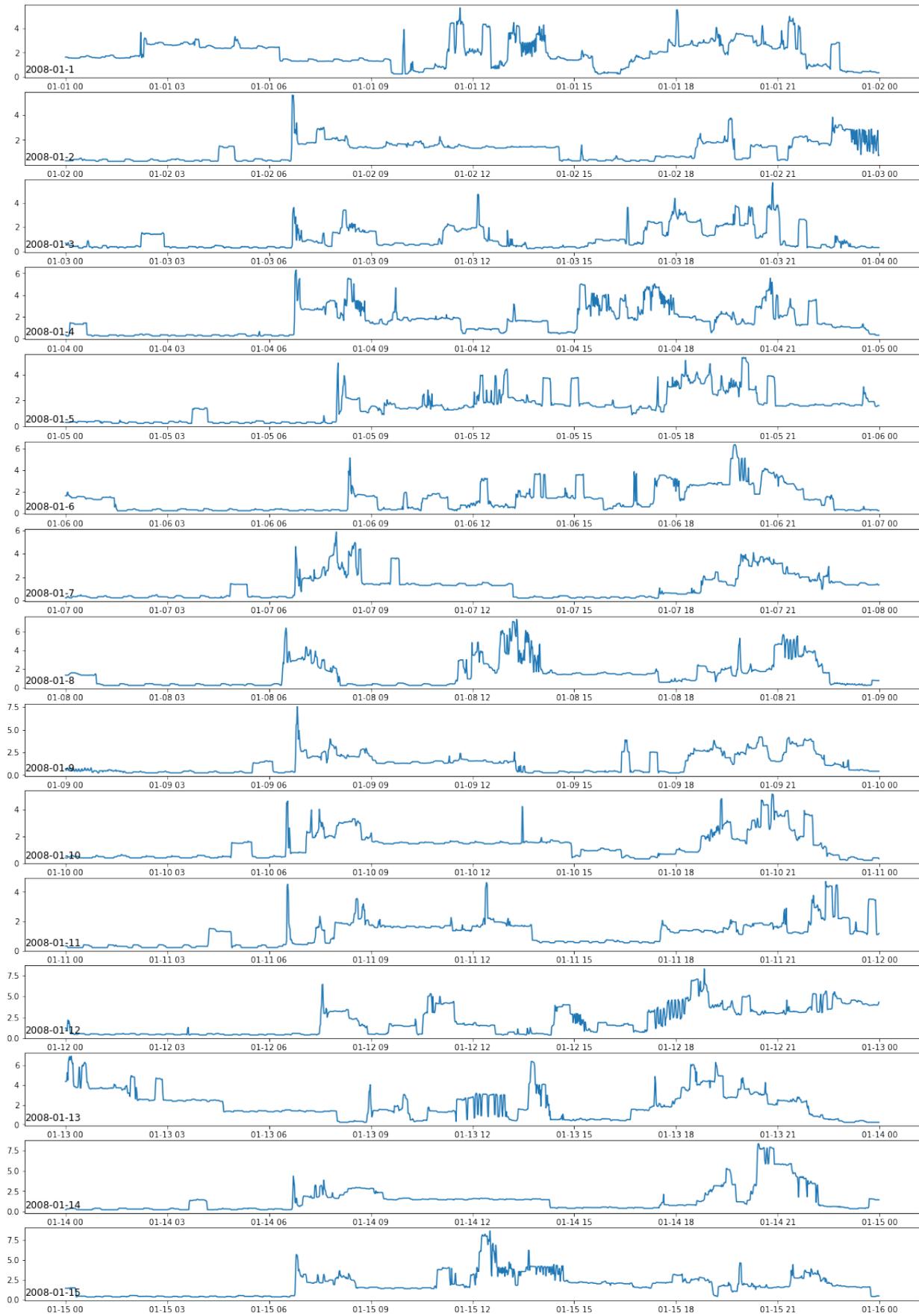


FIGURE 4.12 – Visualisation quotidienne de les 15 premiers jours de l'année 2008

D'après l'observation de ces graphes nous pouvons partager les comportements des locataires à 4 phases principales :

La première phase [22h - 06h] : On remarque la diminution de la consommation d'électricité pendant cette période ce qui nous amène à déduire que les habitants sont endormis.

La deuxième phase [06h - 10h] : On remarque l'augmentation de la consommation d'électricité, ça signifie le commencement des activités ménagers.

La troisième phase [10h - 18h] : On remarque la baisse de la consommation d'électricité, ce qui est logique certaines activités de ménage qui exploitent l'électricité est terminée, et certaines personnes quittent la maison.

La quatrième phase [18h - 22h] : On remarque de nouveau l'augmentation de la consommation d'électricité dû au retour des habitants chez soi, et vu que c'est la nuit donc il y a l'illumination des ampoules, la télévision etc.

mais à part ça il existe des exceptions comme les journées de week-end (Dimanche 13-01-2008) où la consommation d'électricité a tout le temps augmenté.

4.4.3.4 Distribution annuelle pour la variable la puissance active totale

Pour observer et comprendre mieux les données pour la variable puissance active totale est de tracer l'histogramme de distribution totale pour chaque année [2007 - 2010] la figure 4.13 ci-dessous reprend les histogrammes obtenus.

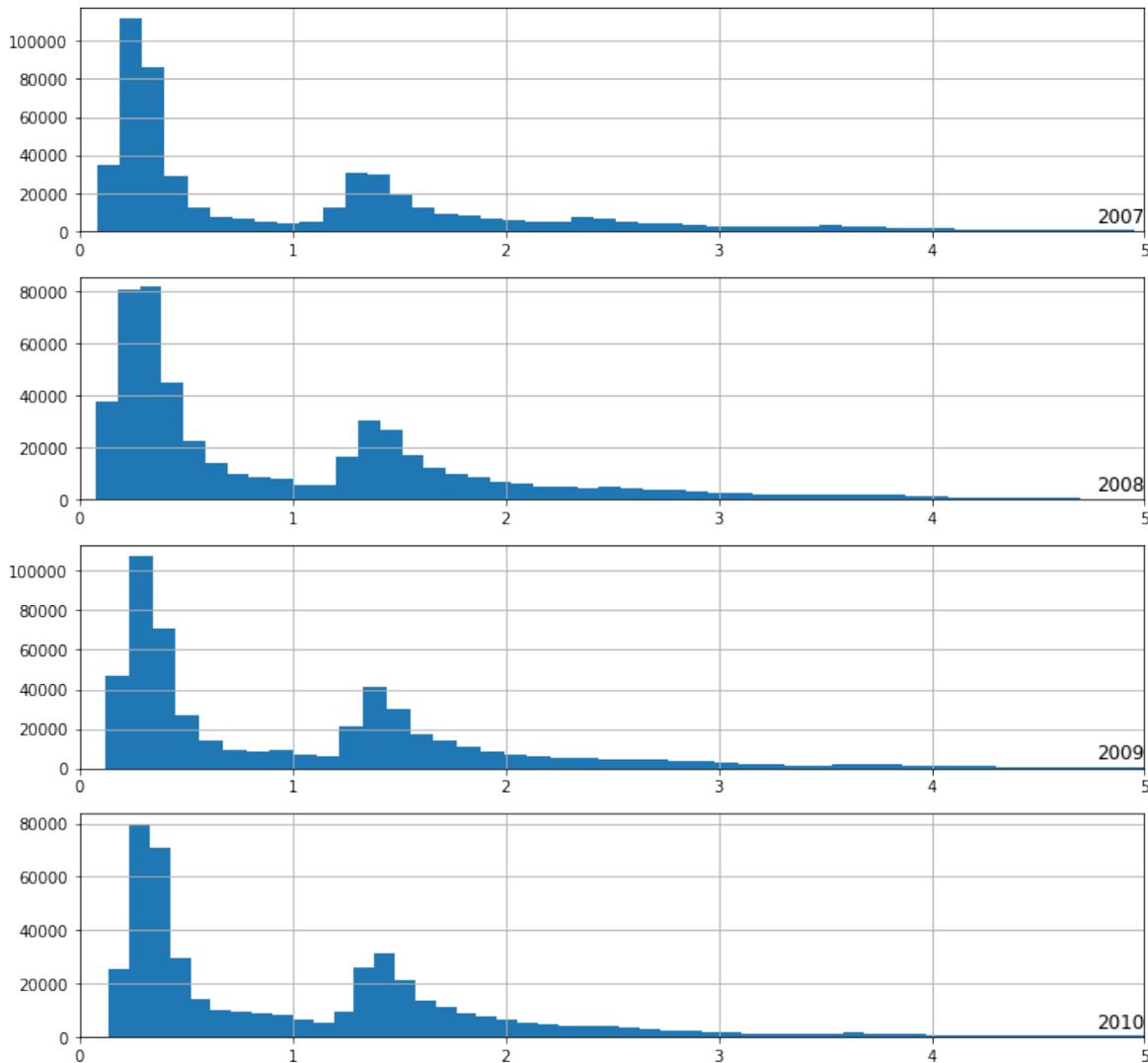


FIGURE 4.13 – Distribution annuelle de la puissance active totale

Au fil des ans, tous les histogrammes sont très similaires et étroitement liés. De plus, nous pouvons voir que la distribution est bimodale ; cela signifie que les données de consommation d'électricité sont discrétisées en deux pics (le pic 1 environ 80 KW, le pic 2 environ 40 KW). Cela suggère que nous pouvons créer des modèles pour prédire et de comprendre le comportement de la consommation d'électricité.

4.5 Modèles de résolution

Dans cette section nous allons présenter les modèles de prédiction utilisés pour la prévision de la consommation d'électricité d'un seul ménage sur quatre ans et une comparaison est faite entre les modèles de séries chronologiques : LSTM long short-term memory, Random forest, XGBRegressor et SVRegressor.

4.5.1 Préparation des données

4.5.1.1 Normalisation des données

La normalisation est une méthode de prétraitement des données qui permet de réduire la complexité des modèles. C'est également un préalable à l'application de certains algorithmes[31]. La normalisation peut- être appliquée par le min-max scaling. Python propose pour cela une classe nommée "MinMaxScaler" dans le package preprocessing. On remarque bien qu'après la normalisation, l'intervalle des valeurs devient limité entre 0 et 1.

4.5.1.2 Données d'apprentissage et données de test

La meilleur façon d'obtenir des meilleur resultats de prédiction c'est de reserve la majorite des données sur le dataset pour l'entraînement. On réserve 75% (3 ans) à des données du dataset pour l'entraînement. Les 25% (1 an) restants seront exploités pour le test.

4.5.2 LSTM long short-term memory

Est un modèle d'apprentissage en profond séquentiel. Le modèle construit est composé d'une couche LSTM avec 300 unités et avec une entrée d'un vecteur de diminution (1,6) et la couche LSTM est connectée avec une couche de dense(couche complètement connecté) avec Trois neurones, cette dernière et connecte avec la couche de sortie à une diminution(couche de dense avec un seul neurone), nous avons ajouté à la couche LSTM la technique Drop-out afin d'éviter le sur-ajustement (Overfitting en anglais), cette méthode de Régularisation permet de désactiver temporairement certains neurones dans le réseau. Le modèle est entraîné pendant 30 époques avec l'algorithme d'optimisation Adam. Nous avons remarqué que la courbe de perte d'entraînement (train loss) est décroissante, ce di-minuit de 0.0180 vers 0.0129 et la même chose

pour la courbe de perte de test, qui diminuait de 0.0088 à 0.0084 comme montre la figure 4.14.

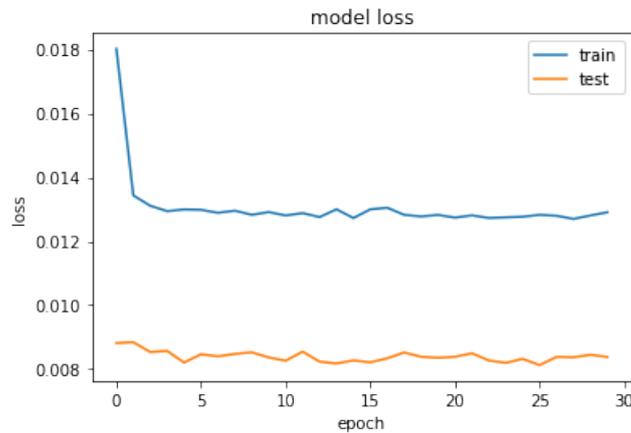


FIGURE 4.14 – evolution du modèle LSTM

La fig 4.15 montre les résultats de perte du modèle pour la puissance active globale et les valeurs réelles pendant 74 heures, lorsque le modèle terminé l'entraînement.

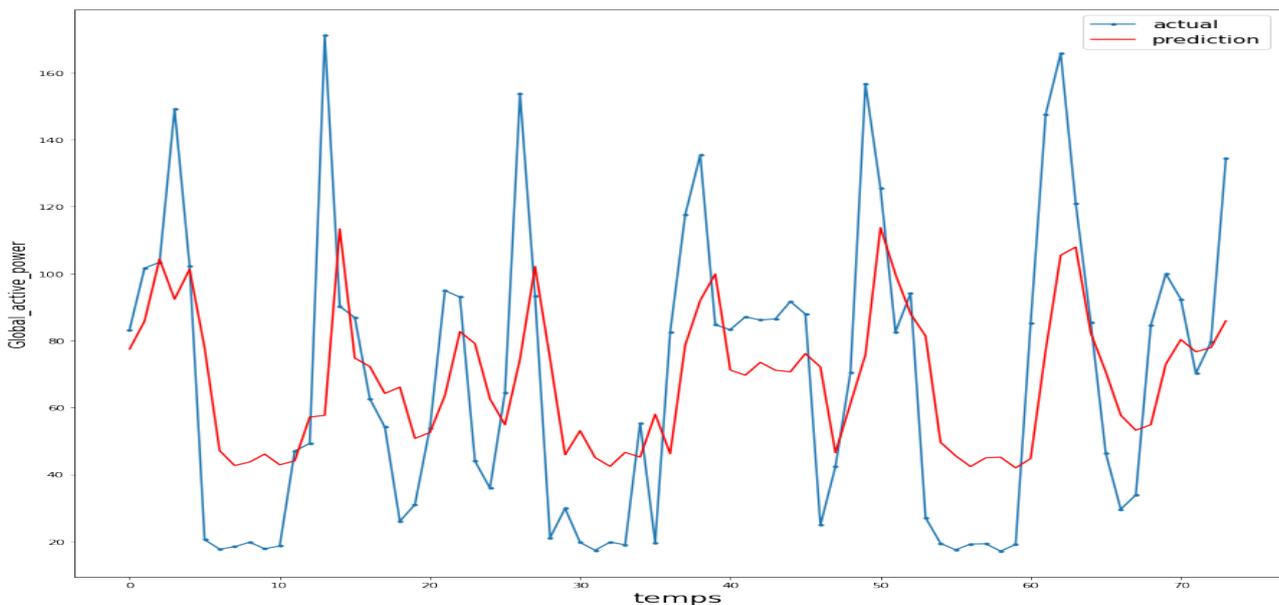


FIGURE 4.15 – Performances du modèle de prédiction LSTM

4.5.2.1 Évaluation du modèle LSTM

Pour évaluer le modèle de LSTM il faut tout d'abord calculer les valeurs de RMSE et MAE. sont les suivant :

Test RMSE : 35.80

Test MAE : 28.71

4.5.3 Random forest

Pour ce modèle, nous allons utiliser l'algorithme des forêts aléatoires qui peut être utilisé pour la prédiction des séries temporelles, et son principe c'est de créer plusieurs arbres de décision. Appliquons ce modèle à nos données supervisées avec ces hyper paramètres :

Nombre d'estimateurs : Initialisé à 1000, donc le modèle crée 1000 arbres décisionnaires.

Random state : Initialisé à 0 pour que les résultats ne changent pas à chaque exécution.

N caractéristiques : Le nombre des caractéristiques considéré par les arbres : La puissance réactive, Tension moyenne, Intensité moyenne du courant, Énergie active pour la cuisine, Énergie active pour la lessive et Énergie active pour les systèmes de climatisation.

La Fig 4.16 montre les résultats de la prédiction du modèle pour la puissance active globale et les valeurs réelles pendant 74 heures, lorsque le modèle a terminé l'entraînement.

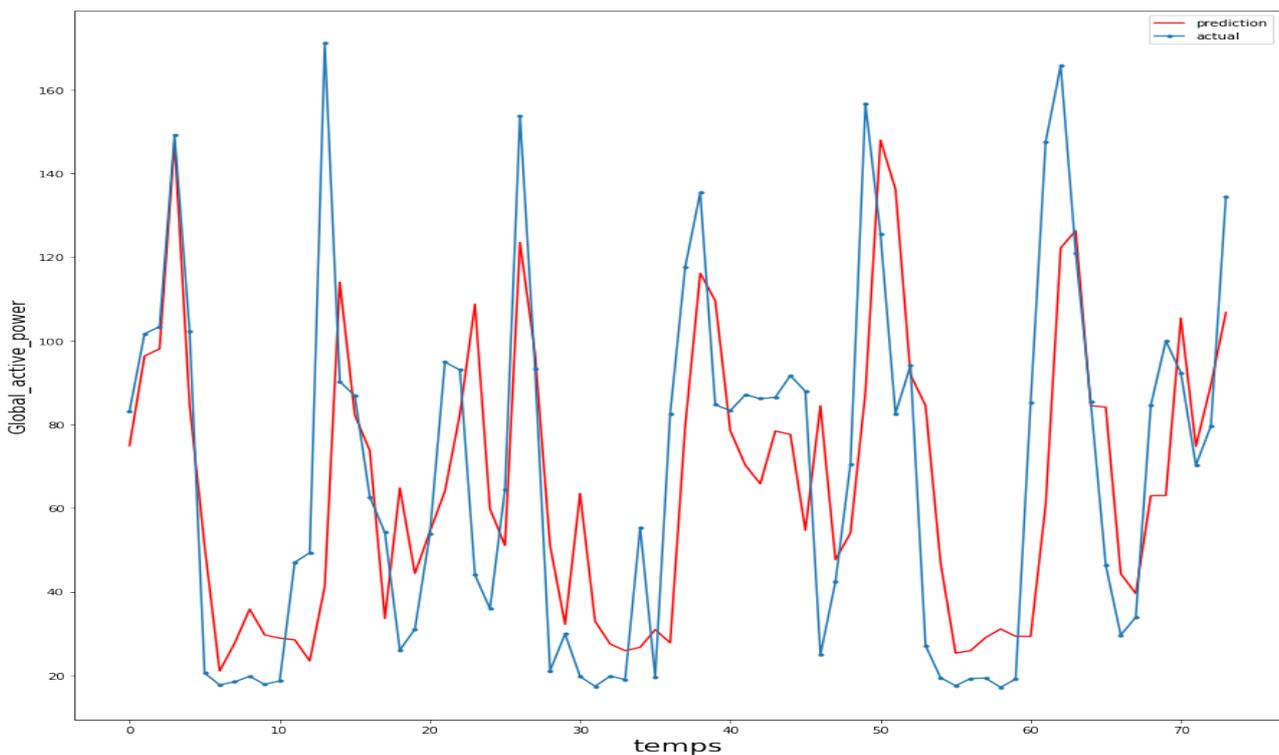


FIGURE 4.16 – Performances du modèle de prédiction Random forest

4.5.3.1 Évaluation du modèle Random Forest

Pour évaluer le modèle de Random Forest il faut tout d'abord calculer les valeurs de RMSE et MAE. sont les suivant :

Test RMSE : 33.10

Test MAE : 23.37

4.5.4 XGBRegressor

Signifie "Extrême gradient Boosting" et c'est une implémentation de l'algorithme d'arbres de amélioration du gradient. Le XGBoost est un modèle d'apprentissage automatique supervisé. Il peut également être utilisé pour prédire des séries temporelles. Appliquons ce modèle sur nos données supervisé avec ces hyper paramètre :

Nombre d'estimations : Initialisé à 1000, le modèle de XGBR créer 1000 arbres décisionnaire.

Objective : Régression erreur carrée.

La profondeur maximale : Initialisé à 6 pour chaque arbre.

Taux d'apprentissage : Initialisé à 0.3.

Sous-échantillon : Initialisé à 1 cela signifie 100% des observations dans l'ensemble des données seront sélectionnées au hasard pour créer chaque arbre individuel dans le modèle et chaque itération aura un échantillon d'observation différent, cela nous aidera à former le modèle rapidement.

Echantillon-bytree : Initialisé à 1 cela signifie 100% des caractéristique seront utilisés à chaque fois sélectionner au hasard qu'un nouvel arbre est intégré dans le modèle.

La figure 4.17 montre des résultats de perte du modèle pour collecter les valeurs d'énergie active et d'énergie active totale réelle pour 74 échantillons, lorsque le modèle terminé l'entraînement.

4.5.4.1 Évaluation du modèle XGBRegressor

Pour évaluer le modèle de XGBRegressor il faut tout d'abord calculer les valeurs de RMSE et MAE. sont les suivant :

Test RMSE : 32.95

Test MAE : 23.35

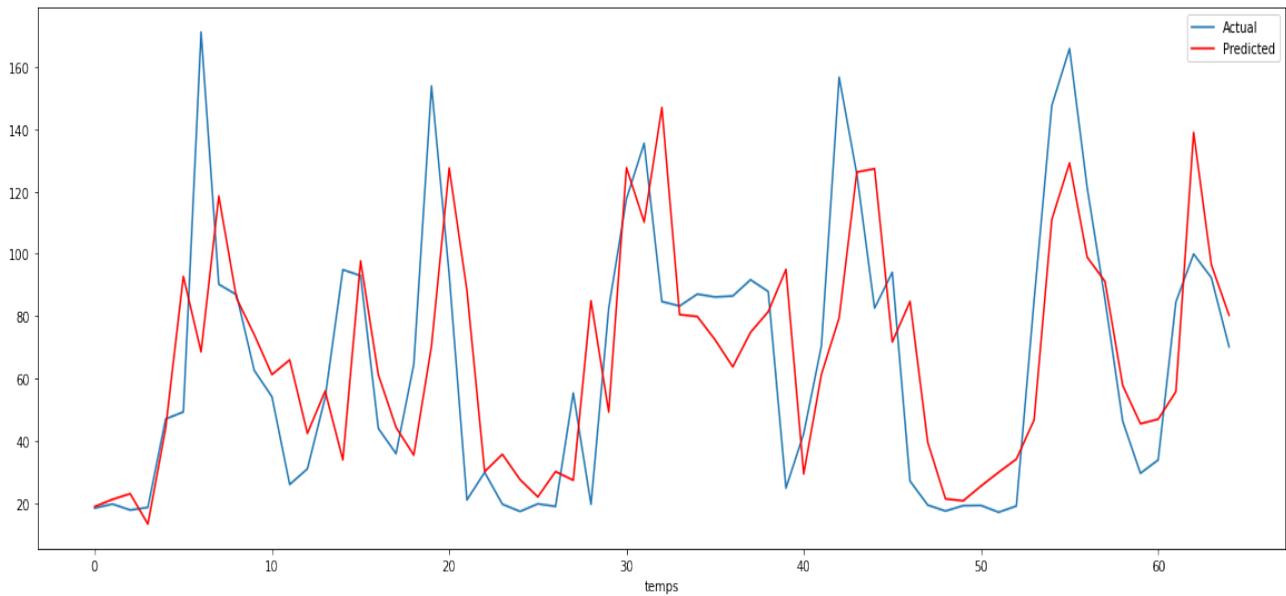


FIGURE 4.17 – Performances du modèle de prédiction XGBRegressor

4.5.5 SVRegressor

SVR peut également être utilisé comme méthode de régression, en conservant toutes les principales caractéristiques qui caractérisent l'algorithme, et il peut aussi prédire les séries temporelles. Appliquons ce modèle sur nos données supervisées avec ces hyper paramètre suivants :

Noyau 'rbf' : Signifie radial basis function.

Gamma : Coefficient de noyau pour 'rbf' initialisé à 0.5.

C : Le paramètre de régularisation initialisé à 10.

Epsilon : Initialisé à 0.05, cette valeur n'est associée dans la fonction de perte d'apprentissage aux points prédits.

La figure 4.18 montre des résultats de perte du modèle Pour collecter les valeurs d'énergie active et d'énergie active totale réelle pour 74 échantillons, lorsque le modèle terminé l'entraînement.

4.5.5.1 Évaluation du modèle SVRegressor

Pour évaluer le modèle de svregressor il faut tout d'abord calculer les valeurs de RMSE et MAE. sont les suivant :

Test RMSE : 51.79

Test MAE : 37.83

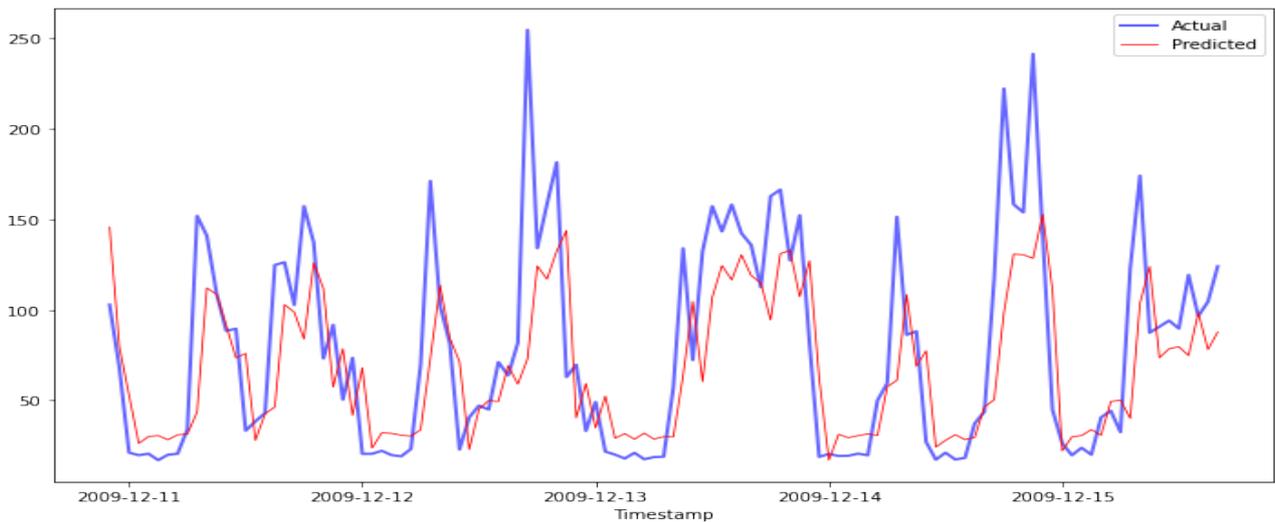


FIGURE 4.18 – Performances du modèle de prédiction SVRegressor

4.5.6 Tableau de comparaison

La tableau 4.2 montre les résultats des tests RMSE et MSE pour tous les modèles LSTM, Random Forest, XGBRegressor et SVRegressor.

Modèle	Test RMSE	Test MAE
LSTM	35.80	28.71
Random Forest	33,10	23.37
XGBRegressor	32,95	23.35
SVRegressor	51.79	37.83

TABLE 4.2 – Tableau de comparaison des tests d'erreur des modèles

4.5.7 Comparaison des résultats des modèles

Après l'analyse du tableau 4.2 nous remarquons ce qui suit :

Le modèle LSTM a donné un résultat de 35,80 pour le test RMSE, et 28,70 pour le test MAE.

Le modèle Random forest a donné un résultat de 33,10 pour le test RMSE, et 23,37 pour le test MAE.

Le modèle XGBRegressor a donné un résultat de 32,95 pour le test RMSE, et 23,35 pour le test MAE.

Le modèle SVRegressor a donné un résultat de 51,79 pour le test RMSE, et 37,83 pour le test MAE.

En comparant les différents résultats obtenus pour chacun des modèles de prédiction d'électricité, on déduit que les meilleur modèles de prédiction sont : le "XGBRegressor" et le "Random Forest", parce que leurs valeurs de RMSE et MAE sont petites par rapport aux autre modèles. Et le pire modèle et le "SVRegressor", parce que leur valeurs de RMSE et MAE sont très grand, quant au modèle restant "LSTM", on peut dire qu'il est au milieu moyenne.

4.6 Conclusion

Dans ce chapitre, nous avons présenté nos travaux sur l'internet des comportement, où nous avons d'abord présenté l'environnement de travail que nous avons utilisé, puis nous avons fait une analyse complète des données en les affichant et enfin, nous avons présenté les différents modèles utilisés pour résoudre le problème de prédiction et nous avons fait comparer ces modèles à l'aide de métriques.

Conclusion générale

Nous avons pris conscience que l'internet des comportements (Internet of behaviors en anglais) est l'une des technologies récentes les plus importantes utilisées dans le monde réel, citons par exemple : e-commerce, domaine de la santé, analyse de sentiments, etc. Pour cela, en choisissant l'électricité comme paramètre d'étude, l'objectif de notre travail était la prédiction quotidienne de la consommation d'électricité dans chaque maison.

Tout au long de ces quatre chapitres traités, nous avons pu voir qu'est-ce que l'internet des comportements, son historique et qu'est-ce qui nous a motivé à travailler sur cette thématique. Nous avons ensuite passé aux différentes techniques reliées à l'IoB. Après, nous avons cité quelques travaux existants dans la littérature portant sur ce processus mené d'un tableau comparatif des divers approches.

Puis, dans le dernier chapitre, il s'agit de spécifier un environnement de travail, d'analyser les données de consommation d'électricité, et enfin de créer un modèle pour prédire et comparer le comportement de consommation de chaque ménage.

Comme perspective, nous souhaitons trouver une manière pour influencer le comportement du consommateur dans le but d'économiser le coût et préserver l'énergie.

Bibliographie

- [1] <http://news.aouaga.com/h/138952.html/>, (consulté le 02/03/2022).
- [2] <https://www.techopedia.com/definition/34552/internet-of-behaviors-iob/>, (consulté le 24/05/2022).
- [3] [https://fr.wikipedia.org/wiki/Anaconda_\(distribution_Python\)/](https://fr.wikipedia.org/wiki/Anaconda_(distribution_Python)/), (consulté le 15/07/2022).
- [4] <https://blog.orsys.fr/les-carnets/index.php/2021/07/06/apres-liot-liob-on-vous-dit-tout-sur-linternet-of-behavior/>, (consulté le 08/04/2022).
- [5] <https://www.techfunnel.com/information-technology/internet-of-behaviors/>, (consulté le 22/05/2022).
- [6] <https://www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-201.html>, (consulté le 2/06/2022).
- [7] <https://research.aimultiple.com/internet-of-behaviors/>, (consulté le 20/05/2022).
- [8] <https://www.futura-sciences.com/tech/definitions/informatique-big-data-15028/>, (consulté le 10/06/2022).
- [9] <https://citrodigital.com/blog/june-2021/internet-of-behavior-iob>, (consulté le 10/06/2022).
- [10] <https://ryax.tech/fr/iot-et-big-data-comprendre-la-relation-entre-ces-deux-technolo> (consulté le 2/06/2022).
- [11] C. Zhang, T. Tan. The impact of big data analysis on consumer behavior. *Journal of Physics : Conference Series*, 2020.

-
- [12] <https://www.analyticssteps.com/blogs/introduction-internet-behaviour-iob/>, (consulté le 20/05/2022).
- [13] <https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption>, (consulté le 08/04/2022).
- [14] https://www.objetconnecte.com/iob-tout-savoir-internet-of-behavior/#IoB_les_avantages_de_lInternet_du_comportement, (consulté le 24/05/2022).
- [15] <https://www.techopedia.com/definition/34552/internet-of-behaviors-iob/>, (consulté le 14/05/2022).
- [16] Fabio Nelli. Python data analytics. *Apress Media, California*, 1(1) :213, 2018.
- [17] Giancarlo Zaccone. Getting started with tensorflow. *Packt Publishing Birmingham*, 1(1) :180, 2016.
- [18] Xiaoping Rui Guangxia Yu Yonglei Fan Xianfeng Song Guangyuan Zhang, Stefan Poslad and Runkui Li. Using an internet of behaviours to study how air pollution can affect people’s activities of daily living : A case study of beijing, china. *mdpi sensors*, 2(1) :37, 2021.
- [19] Mohsen Guizani Haya Elayan, Moayad Aloqaily. Internet of behavior (iob) and explainable ai systems for influencing iot behavior. *arXiv*, 1(1) :7, 2021.
- [20] <https://www.oracle.com/dz/internet-of-things/what-is-iot/>, (consulté le 28/05/2022).
- [21] <https://www.techtarget.com/iotagenda/definition/Internet-of-Things-IoT>, (consulté le 28/05/2022).
- [22] <https://www.oracle.com/dz/internet-of-things/what-is-iot/>, (consulté le 27/06/2022).
- [23] <https://iotindustriel.com/tendances-de-liot-industriel/machine-learning-dans-liot-quels-impacts/>, (consulté le 12/06/2022).
- [24] <https://www.oracle.com/dz/data-science/machine-learning/what-is-machine-learning/>, (consulté le 10/06/2022).
- [25] Julie Dugdale Mikkel Baun Kjærgaard Mahyar T. Moghaddam, Henry Muccini. Designing internet of behaviors systems. *Springer*, 1(1) :7, 2021.

-
- [26] <https://www.objetconnecte.com/iot-et-la-machine-learning/>, (consulté le 12/06/2022).
- [27] Ravi Pratap Singh Shanay Rab Rajiv Suman Mohd Javaid, Abid Haleem. Internet of behaviours (iob) and its role in customer services. *keaipublishing*, 2(1) :8, 2020.
- [28] <https://www.statisticshowto.com/absolute-error/>, (consulté le 22/07/2022).
- [29] Navin Kumar Manaswi. Understanding and working with keras. *Deep Learning with Applications Using Python. Springer*, 1(1) :176, 2018.
- [30] Nikita Silaparasetty. Introduction to jupyter notebook. *Machine Learning Concepts with Python and the Jupyter Notebook Environment. Springer*, 1(1) :118, 2020.
- [31] <https://dataanalyticspost.com/Lexique/normalisation/#:~:text=Normalisation%20%3A%20La%20normalisation%20est%20une,l%27application%20de%20certains%20algorithmes.>, (consulté le 10/10/2022).
- [32] <https://medium.com/@namanbhandari/extratreesclassifier-8e7fc0502c7/>, (consulté le 15/07/2022).
- [33] <https://businessnow.fr/solutions/linternet-des-objets-iot-quel-avenir-dans-le-monde> (consulté le 10/06/2022).
- [34] Oliver Kramer. Scikit-learn. *Machine learning for evolution strategies. Springer*, 20(1) :228, 2016.
- [35] <https://techvidvan.com/tutorials/python-advantages-anddisadvantages/>, (consulté le 15/07/2022).
- [36] <https://www.statisticshowto.com/probability-andstatistics/regression-analysis/rmse-root-mean-square-error/>, (consulté le 22/07/2022).
- [37] <https://imaginovation.net/blog/8-sectors-benefit-from-iot-development-in-2021>, (consulté le 30/05/2022).
- [38] Christian Stary. The internet-of-behavior as organizational transformation space with choreographic intelligence. *Springer*, 2(1) :20, 2020.
- [39] <https://www.infoworld.com/article/3278008/what-istensorflow-the-machine-learning-library.html/>, (consulté le 16/07/2022).
- [40] <https://code.visualstudio.com/docs/>, (consulté le 22/07/2022).

※ *Résumé* ※

Le mémoire ici présent a pour thème l'Internet des comportements : analyse et prédiction des comportements en tant que solution à la prise de décision quotidienne. L'objectif de ce dernier est de se questionner et donc de poser une réflexion sur la possibilité que dans un futur plus ou moins proche, incorporés à plusieurs niveaux dans notre quotidien pour accompagner les humains dans leur prise de choix quotidien.

Deuxièmement, cela montre à travers des recherches sur comment une maison consomme de l'électricité sur une période de quatre ans, et comment l'Internet des comportements peuvent-ils se comporter en matière des heures précises pour fournir des prévisions sur la quantité de consommation d'électricité dans les heures suivant au consommateur. Enfin, dans la dernière partie de ce mémoire, quatre modèles qui prédisent la quantité d'électricité qu'ils consommeront à la maison, et nous déterminerons le plus approprié.

Il est apparu que de nombreux projets actuels tentent, ou parviennent à intégrer l'Internet des comportements en leur sein, ce qui nous amène à la conclusion que l'internet des comportements peut être une réponse à nos problèmes et à nos besoins du quotidien.

※ *Abstract* ※

The theme of this dissertation is the Internet of behaviors : analysis and prediction of behaviors as a solution to daily decision making. The objective of the latter is to question and therefore to pose a reflection on the possibility that in a more or less near future, incorporated at several levels in our daily life to accompany humans in their daily choices.

Secondly, it shows through research on how a house consumes electricity over a period of four years, and how the Internet of behaviors can behave in the matter of precise hours to provide forecasts on the amount of electricity consumption in the following hours to the consumer. Finally, in the last part of this thesis, four models that predict the amount of electricity they will consume at home, and we will determine the most appropriate.

It appeared that many current projects try, or manage to integrate the Internet of Behaviors within them, which leads us to the conclusion that the Internet of Behaviors can be an answer to our problems and needs of everyday life.