

Université Assane SECK de Ziguinchor

UFR Sciences et technologies

Département Informatique



Mémoire de fin d'études

Pour l'obtention du diplôme de Master

Mention : Informatique

Spécialité : Génie Logiciel

Sujet :

Proposition d'une Architecture Basée sur l'Intelligence Artificielle pour le Contrôle de Congestion dans l'Internet des Objets

Présenté par : Mlle Penda KA

Soutenu le 31 mars 2022

Sous la direction de : Pr Youssou FAYE

Devant le jury composé de :

Ousmane DIALLO	Maitre de conférences (Cames)	Président
Youssou FAYE	Maitre de conférences (Cames)	Encadreur
Ibrahima DIOP	Maitre-assistant (Cames)	Rapporteur
El Hadji Malick NDOYE	Maitre-assistant (Cames)	Rapporteur

Année universitaire 2020-2021

Dédicaces

À ma très chère maman Oumou Salmata BA !

À mon très cher papa Ousmane KA!

À mon très cher oncle Ibrahima KA,

À ma très chère tante Bineta GNING,

À mes adorables frères et sœurs,

À mes très chers cousins et cousines,

À la mémoire de ma grand-mère Fatimata BA,

À la mémoire de mes promotionnaires : Mamadou Petit
Coulibaly, Alpha SANE et Mamour DIOUF

À mes camarades de promotion en témoignage de l'amitié et les bons
moments passés ensemble en vous souhaitant un avenir radieux et plein de réussites.

Je dédie ce travail...

Penda KA

Remerciements

En premier lieu nous remercions DIEU le tout puissant de nous avoir donné la patience, la santé et la volonté pour achever ce travail !

Je souhaite rendre hommage à tous ceux qui ont contribué à la réalisation de ce travail!

Je tiens à remercier en premier lieu et à exprimer toute ma reconnaissance à mon encadreur, le **Pr Youssou FAYE** qui a su partager son savoir faire, ses connaissances son expérience, ses idées et son temps pour me porter aide pendant et hors de ses heures de travail à la fois compréhensible et serviable avec beaucoup de présentations préliminaires. J'aimerais adresser plus qu'un merci pour ses bons conseils, ses bons réflexes, sa disponibilité et surtout pour tout le savoir qu'il m'a transmis. Enfin, veuillez recevoir mes chaleureux remerciements pour tout !

J'adresse mes remerciements aux membres du jury pour m'avoir honoré en acceptant d'évaluer ce travail et de participer à ma soutenance de mémoire :

Je suis très honoré par la présence de *Monsieur Ousmane DIALLO* Maitre de conférence à l'Université Assane Seck de Ziguinchor pour le temps qu'il a bien voulu consacrer à l'évaluation de ce travail ;

Je suis également très honoré par la présence de *Monsieur Ibrahima DIOP* Maitre-assistant à l'Université Assane Seck de Ziguinchor pour le temps qu'il a bien voulu consacrer à l'évaluation de ce travail ;

Je suis également très honoré par la présence de *Monsieur El hadji Malick NDOYE*, Maitre-assistant à l'Université Assane Seck de Ziguinchor pour le temps qu'il a bien voulu consacrer à l'évaluation de ce travail !

Je ne saurais terminer sans remercier l'administration de l'UFR ST, le corps professoral du département informatique d'avoir contribué à ma formation, Dr Gorgoumack SAMB pour ses conseils !

Et mes remerciements vont tout particulièrement à mes parents, pour leur soutien et leur amour !

Nous vous devons ce que nous sommes aujourd'hui, grâce à votre amour, votre patience et leurs innombrables sacrifices. Que ce modeste travail, soit pour vous une petite compensation et reconnaissance pour tout ce que vous avez fait. Que Dieu, vous préserve et vous procure santé et longue vie afin que nous puissions à notre tour vous combler !

À Fatoumata Mbéndary, à Mr Mor NGOM toute la famille NGOM à Bambey de m'avoir accueilli et soutenu !

Mes sincères remerciements à ma famille d'accueil à Ziguinchor, à Modiéré Dicko ainsi que toute la famille Malou pour leur soutien inconditionnel et leur amour inestimable qu'ils ont su m'apporter !

À mon Cousin Baye Dieng KA et ma cousine Khardiata KA pour leurs soutiens et encouragements !

À mes frères et sœurs Mamadou KA, Abibatou KA, Fatimata KA, Ibrahima KA pour leurs soutiens sans faille et leurs encouragements continus !

Mes chaleureux remerciements à mon ami Diadia William MANGA, à Dioncounda Yock pour tout!

Mes chaleureux remerciements à Mr Alioune Badara DIENG, Mame Diarra SY, Mor MBOUP pour tous les efforts fournis !

À tous ceux qui m'ont soutenu de près ou de loin MERCI !

Résumé

De nos jours, nous assistons de plus en plus à un développement des objets connectés sous plusieurs formes. Pour une variété d'applications, une grande masse de données est générée lors des communications entre objets. Il devient alors difficile de mettre en place des protocoles de communications résistants à la congestion.

L'intelligence artificielle vient en application pour relever ces défis de congestion dans l'IoT par le biais de techniques d'algorithmes d'apprentissage supervisé et non supervisé qui peuvent permettre d'extraire, de classifier, de prédire afin de prendre de bonnes décisions notamment la réduction du temps de transfert des données.

Avec l'expansion anarchique de certaines villes les rendant incapables de faire face à l'augmentation du nombre des véhicules, il devient alors difficile de circuler sans embouteillage et les congestions deviennent inévitables tout comme dans l'IoT.

Aujourd'hui, des travaux de recherche basés sur l'IoT se développent afin d'apporter des réponses face au phénomène de congestion dans la circulation urbaine. Dans la littérature scientifique, plusieurs travaux de recherche utilisent des techniques basées sur un apprentissage supervisé, non supervisé ou même profond. KMeans est une technique de classification des données en cluster souvent utilisée pour la résolution de congestion.

Dans ce mémoire, nous l'utilisons dans un contexte d'une architecture de cinq couche en vue de pouvoir la combiner avec les réseaux de neurones centroïdes pour une classification plus efficace. C'est ainsi que nous mettons en œuvre KMeans dans notre architecture à cinq niveaux pour prédire des congestions liées aux voies de circulation à travers des sources de données mises à disposition.

Nous avons fait une implémentation avec un distributeur de python d'anaconda et les résultats obtenus pourraient permettre aux décideurs d'une région, d'une ville ou de n'importe quelle agglomération de prendre des décisions idoines face à une problématique de congestion.

Mots clés : internet des objets, intelligence artificielle, apprentissage non supervisé, KMeans, réseau de neurone centroïdes, congestion

Abstract

Nowadays, we are increasingly witnessing the development of connected objects in many forms. For a variety of applications, a large amount of data is generated during communications between objects. It becomes difficult to implement congestion-proof communication protocols.

Artificial intelligence is being applied to address these congestion challenges in the IoT through supervised and unsupervised learning algorithm techniques that can extract, classify, and predict data in order to make good decisions, including reducing data transfer time.

With the anarchic expansion of some cities making them unable to cope with the increase in the number of vehicles, it becomes difficult to circulate without traffic jams and congestion becomes unavoidable just like in the IoT.

Today, IoT-based research is being developed to provide answers to the phenomenon of congestion in urban traffic. In the scientific literature, several research works use techniques based on supervised, unsupervised or even deep learning. KMeans is a clustered data classification technique often used for congestion resolution.

In this thesis, we use it in the context of a five-layer architecture in order to combine it with centroid neural networks for a more efficient classification. Thus, we implement KMeans in our five-layer architecture to predict traffic lane congestion through available data sources.

We have made an implementation with anaconda python dispatcher and the results obtained could allow decision makers of a region, a city or any agglomeration to take appropriate decisions in front of a congestion problem.

Key words: internet of things, artificial intelligence, unsupervised learning, KMeans, centroid neural network, congestion

Sommaire

Dédicaces.....	ii
Remerciements	iii
Résumé.....	v
Abstract	vi
Sommaire	vii
Liste des Figures	xii
Liste des Tableaux	xiii
Glossaire.....	xiv
Introduction Générale.....	1
PARTIE 1: Etat de l'art sur l'internet des objets et l'intelligence artificielle.....	4
Chapitre 1 : Généralités sur l'internet des objets	5
I. Contexte	6
II. Quelques définitions	7
III. Histoire et évolution de l'internet des objets	7
IV. Technologies de communication dans l'internet des objets.....	9
1. Technologies de courtes portées	9
a. Bluetooth.....	9
b. Zigbee	9
c. NFC	9
2. Technologies de moyennes portées.....	9
a. Z-wave	9
b. Wi-Fi.....	10
c. Bluetooth Low Energy	10
3. Technologies de Longues portées	10
a. Réseaux de types LPWAN.....	10
i. Sigfox	10
ii. Lora Wan	10
b. Les Réseaux Cellulaires.....	10
c. Les Réseaux propriétaires	10
d. Fibre Optique.....	11
V. Architecture dans l'internet des objets.....	11
1. L'objet.....	11
2. Le réseau	11
3. Logiciel et infrastructure	11

4.	Applications et services	12
VI.	Domaines d'applications	14
1.	Industrie 4.0.....	14
2.	Santé connectée	14
3.	Transport connecté	14
4.	Domotique.....	15
5.	Smart City	15
6.	Transport logistique connecté.....	15
7.	Agriculture de précision	16
8.	Élevage intelligent	16
9.	Limites dans les domaines d'applications	16
VII.	Plateformes dans l'internet des objets	17
1.	Définition.....	17
1.	Types de plateformes dans l'internet des objets.....	17
a.	Plateformes horizontales	17
VIII.	Composants de l'internet des objets.....	19
1.	Fabricants de capteurs	19
2.	Objet connecté	19
3.	Connectivité.....	19
4.	Plateformes	20
5.	Développeurs d'applications	20
6.	Technologies pour une gestion optimale de données IoT	20
IX.	Procédures de normalisations.....	21
1.	Initiatives d'organisations internationales	21
2.	Initiatives de l'union Européenne	21
3.	Consortiums pour une normalisation spécifique de l'IOT.....	21
X.	Obstacles qui ralentissent le développement de l'internet des objets et perspectives	22
1.	Besoin d'une norme internationale	22
2.	Problèmes d'alimentation des capteurs.....	22
3.	Confidentialité des utilisateurs.....	22
4.	Compétences décisionnelles insatisfaites	22
Chapitre 2 : Intelligence artificielle et l'internet des objets : enjeux et applications.....		24
I.	Qu'est-ce que l'intelligence artificielle.....	25
1.	Définitions	25
2.	Histoire et évolution de l'intelligence artificielle	26
3.	Fonctionnement	27

II.	Le machine Learning.....	27
1.	Apprentissage supervisé.....	27
2.	Apprentissage par renforcement	28
3.	Apprentissage non supervisé	28
4.	L'apprentissage profond.....	28
III.	Domaines d'applications	28
1.	L'application de l'IA dans le domaine de la santé	28
2.	L'industrie du futur.....	28
3.	L'intelligence artificielle au cœur du transport	29
4.	L'intelligence artificielle pour la sécurité des personnes et des biens.....	29
5.	L'agriculture intelligente.....	30
6.	L'Intelligence Artificielle au service de la domotique	30
IV.	Défis-Enjeux et perspectives de l'Intelligence artificielle.....	30
1.	Défis de l'intelligence artificielle	30
2.	Enjeux de l'intelligence artificielle.....	31
a.	Enjeux économiques	31
b.	Enjeux Sociétales	32
3.	Perspectives de l'intelligence artificielle	32
V.	L'application de l'intelligence artificielle dans l'internet des objets.....	33
1.	Exemples d'applications d'AIOT	33
2.	La cognitive dans l'internet des objets.....	33
3.	L'IA dans les réseaux de capteurs sans fil (WSN)	34
4.	L'IA dans la gestion d'énergie dans l'IOT.....	34
5.	L'IA au service de l'exploitation des données générées par l'loT.....	34
	Conclusion	35
	PARTIE 2: Etat de l'art de l'apprentissage machine et contribution	36
	Chapitre 3 : L'apprentissage machine, méthodes et algorithmes.....	37
I.	Principe de fonctionnement des algorithmes d'apprentissage machine :.....	39
1.	Les méthodes d'apprentissages supervisées	39
a.	Classification.....	39
b.	Régression	40
2.	Méthodes d'apprentissages non supervisées	41
II.	Méthodes hiérarchiques	41
a.	Principe de fonctionnement des méthodes hiérarchiques	41
1.	Clustering Hiérarchique ascendante (agglomérative ou bottom up en anglais) :.....	42
2.	Clustering Hiérarchique descendante (divisive ou top down)	43

III.	Principe de fonctionnement des méthodes non hiérarchiques.....	44
IV.	Algorithmes basés sur la classification non supervisée (clustering)	45
1.	Méthodes par partitionnement	45
a.	KMeans.....	45
2.	Les méthodes à base de densité	48
a.	DBSCAN.....	49
b.	DENCLUE.....	50
3.	Les méthodes basées sur un modèle	51
a.	L'algorithme Self-Organizing Maps (SOM)	51
b.	Fonctionnement	51
4.	Méthodes basées sur association	52
a.	Apriori.....	52
b.	Illustration de l'algorithme Apriori.....	53
5.	Méthodes à base de grille	54
a.	STING	54
b.	WaveCluster	55
V.	Les réseaux de neurones artificiels	55
1.	Historique et évolution.....	55
2.	Types de réseaux de neurones	56
a.	Les réseaux de neurones feed-forwarded (traitement de la donnée).....	56
b.	Les réseaux de neurones récurrents	57
c.	Les réseaux de neurones à résonance.....	57
d.	Les réseaux de neurones auto-organisés	57
e.	Les réseaux de neurones convolutifs	57
f.	Les réseaux de neurones Capsules.....	58
	Conclusion	65
	Chapitre 4: L'Apprentissage non Supervisée dans l'Internet des Objets	66
I.	Différents problèmes dans l'IOT.....	68
1.	Routage dans l'internet des objets.....	68
2.	Le problème de Normalisation dans l'internet des objets	68
3.	La gestion de l'énergie dans l'internet des objets.....	68
II.	La qualité de service dans l'internet des objets (QoS)	68
1.	L'architecture dans l'internet des objets.....	69
2.	La congestion dans l'internet des objets.....	69
III.	Traitement de la congestion dans l'Internet des objets	69
1.	La méthode basée sur l'architecture.....	69

2. Méthode basée sur le routage	70
3. Méthode basée sur les algorithmes d'apprentissage non supervisé.....	70
Conclusion	72
Chapitre 5: Implémentation de KMeans dans une Architecture à Cinq Niveaux.....	73
I. CONTEXTE.....	75
II. PROBLEMATIQUE	75
III. Présentation du modèle Existant	76
1. Architecture du système	76
a. Unités d'acquisition de données	76
b. Unités de transmission de données	76
c. Unités de traitement de données	76
d. Principe de fonctionnement de KMeans.....	77
IV. Proposition d'une méthode efficace dans les algorithmes de résolution de la congestion	77
V. Description du modèle	79
1. Fonctionnement de l'algorithme des réseaux de neurones centroïdes	80
2. Classification dans KMeans	80
VI. Implémentation du modèle.....	81
3. VI.1. Paramètres de simulation	81
1. VI.2. L'environnement de simulation	82
2. VI.3. Les étapes de la simulation	83
Conclusion	88
Conclusion Générale.....	89
Bibliographie et Webographie	89

Liste des Figures

Figure 1 : Estimation du nombre d'objets connectés en milliards en 2022	8
Figure 2 : Le modèle de référence de l'IOT (Cisco,2014)	12
Figure 3 : l'industrie 4.0	14
Figure 4 : la ville connectée [38]	15
Figure 5 : le transport logistique connectée [28]	16
Figure 6 : agriculture de précision	16
Figure 7 : élevage intelligent	16
Figure 8 : l'intelligence artificielle dans le domaine de la santé	28
Figure 9 : l'IA dans le domaine de l'industrie	29
Figure 10 : l'IA dans le domaine du Transport	29
Figure 11 : saut minimal	41
Figure 12 : saut maximale	42
Figure 13 : critère de la moyenne	42
Figure 14 : illustration du clustering hiérarchique	43
Figure 15 : illustration du clustering divisive	44
Figure 16 : Principe de fonctionnement de K-medoids	47
Figure 17 : structure de DBSCAN	50
Figure 18 : perception simple	56
Figure 19 : structure de la perception multicouche	57
Figure 20: Architecture de chemin plus court pour contrer la congestion	70
Figure 21: routage maximal par poids maximal	70
Figure 22 : Système d'identification du schéma de congestion du trafic	77
Figure 23 : la distance de proximité inter-véhicule	78
Figure 24 : système de gestion de congestion pour un réseau routier	79
Figure 25:Extrait du fichier trafic.csv	82
Figure 26:L'environnement et les composants d'Anaconda Navigator	83
Figure 27: Représentation de la densité, de la vitesse et des paramètres externes	84
Figure 28: Le rapport des données de vitesse densité	85
Figure 29: clustering KMeans sans la méthode Elbow	85
Figure 30: Distance dans Elbow en fonction du nombre de groupes	86
Figure 31: Nombre de clusters avec méthode du Elbow	86
Figure 32: K=1	86
Figure 33:K=2	87
Figure 34: K=3	87
Figure 35 : K=4	87
Figure 36:K=5	87

Liste des Tableaux

Tableau 1 : illustration de l'association	53
Tableau 2 : Etude Comparative des méthodes d'apprentissages	59
Tableau 3 : Étude comparative des algorithmes d'apprentissages machine	61
Tableau 4: les algorithmes d'apprentissages non supervisé et leurs domaines d'applications	71
Tableau 5:Kmeans et les autres algorithmes d'apprentissage pour minimiser la congestion dans un réseau routier	71

Glossaire

IdO : Internet des Objets

IoT: (Internet of Things) Internet des Objets

IA: (Intelligence Artificiale)

IOE: (Internet of Everything)

AIOT: (Artificial Intelligence of Thing)

NB-IOT: (Narrow Band IOT)

IETF: (Internet Engineering Task Force)

AIOTI: (Alliance for internet of things Innovation)

UIT: (Infrared Data Association)

ISO: (International Standardisation Organisation)

ITS: (Intelligent Transport System)

LTE: (Long Term Evolution)

MIT : (Massachusetts Institute of Technology)

NFC : (Near Field Communication)

CENELEC : (Comité Européen de Normalisation Electronique)

RCSF : Réseaux de Capteurs Sans Fil

TIC : Technologies de l'Information et de la Communication

AA: (Apprentissage Automatique)

M2M: (Machine to Machine)

CNN: (Convolutional Neural Network)

BIRCH: (Balanced Reducing and Iterative Clustering using Hierarchies)

DIANA: Divisive Analysis

LAN: Local Area Network

MAN: Metropolitan Area Network

WAN: Wide Area Network

CH: Cluster Head

CM: Cluster Member

4G : Quatrième Génération de réseau

MPL: (Multiple Perception Layer)

RNN: Recurrent Neural Network

V2I : (Vehicle to infrastructure) Communication entre les véhicules et les infrastructures

V2M : (Vehicule to machine) Communication entre les véhicules et les machines

V2V :(Vehicule to vehicule communication) Communication entre les véhicules connectés

Wi-Fi : (Wireless Fidelity)

WLAN : (Wireless Local Area Network)

Z-Wave : Protocole radio conçu pour la domotique (éclairage, chauffage)

Introduction Générale

L'évolution qu'a connu les technologies émergentes ces dernières années va de plus en plus dans le sens de digitaliser le quotidien des hommes qui communiquent de plus en plus avec des objets. Cette interaction entre homme-objet, entre objet-objet est un fruit de plusieurs systèmes interconnectés, dont l'internet des objets. Ce qui peut créer des congestions dans les systèmes. De nos jours, on assiste de plus en plus à un développement anarchique de certaines villes les rendant incapables de faire face à l'augmentation du nombre des véhicules. Il devient alors difficile de circuler sans embouteillage et les congestions deviennent inévitables tout comme dans l'IoT où l'effet d'envoi et la réception d'une grande masse de données captées peut connaître un retard de transmission.

L'intelligence artificielle vient en application pour relever ces défis de congestion dans l'IoT par le biais de techniques d'algorithmes d'apprentissage supervisé et non supervisée qui peuvent permettre d'extraire, de classier, de prédire afin de prendre de bonnes de décisions notamment la réduction du temps de transfert des données. Aujourd'hui, dans certains pays comme le Sénégal(Dakar) [1]. La France(Paris), le Brésil(Sao Polo) [2], la Chine(Pékin), l'Inde(Mumbai) [2], des travaux de recherche basés sur l'IoT se développent afin d'apporter des réponses face au phénomène de congestion dans la circulation urbaine.

Dans la littérature scientifique, il existe plusieurs travaux de recherche tels que des méthodes de prédiction basées sur des données historiques et des données en temps réel qui, après traitement, sont envoyées aux utilisateurs via une architecture client-serveur [3]. Dans [4], les auteurs ont comparé les algorithmes KMeans, C-means Fuzy, KMeans Fuzy dans un réseau VANET, ils ont considéré la vitesse, la consommation de carburant et l'émission de CO2.

Dans [5], les auteurs ont combiné l'internet des objets et les réseaux de neurones de type LSTM (Long Short Term Memory), ils considèrent la vitesse recueillie par deux capteurs placés dans deux directions Nord et Sud alors que les auteurs du papier [6] ont classifié des données de type normale et anormale. Les données anormales sont considérées et classifiées par KMeans et les arbres de décisions dans un temps record avec les paramètres météorologiques. Les travaux dans [7] utilisent une architecture IoT à quatre couches puis comparent les algorithmes DBSCAN, KMeans, Affinity propagation basés sur la classification non supervisée puis ont essayé d'apporter des solutions face à ce phénomène de congestion.

Cependant des insuffisances et limites liées à la gestion de la distance inter-véhicule et le paramètre temps lié au trafic ont été notées dans le mécanisme de résolution. Pour combler certains manquements notés, dans [8] les auteurs prennent comme paramètre la vitesse et la densité pour illustrer le degré de la congestion qui est lié à d'autres paramètres.

Pour une meilleure illustration, nous proposons une solution basée sur l'algorithme KMeans et les réseaux de neurones centroides appliquée à une architecture en cinq niveaux : le niveau collecte, le niveau transmission, le niveau stockage et les niveaux traitement et apprentissage.

Dans ce sillage, notre défi est de proposer les méthodes de l'intelligence artificielle applicables à l'internet des objets que nous allons appliquer dans le contexte de villes intelligentes, notamment dans le domaine du transport intelligent. Pour ce faire nous nous penchons sur une architecture internet des objets de réseaux de voitures connectés qui envoi des données de vitesse vers un serveur, ces données vont être exploitées par des algorithmes de l'intelligences artificielles dans le but d'aider les décideurs à gérer la congestion du réseau routier. Nous avons par simulation mis en œuvre les différents niveaux et nous avons utilisé KMeans pour le partitionnement des données issues de la ville de Rennes (France) [9] et leur classification. L'implémentation est faite par le langage python, à l'appui d'anaconda un distributeur de python. Les résultats obtenus permettent aux décideurs d'une région, ville ou n'importe quelle agglomération de prendre des décisions idoines face aux problèmes décelés.

Ce mémoire est composé de cinq chapitres organisés en deux parties. La première partie est composée de deux chapitres et est consacrée aux fondamentaux de l'internet des objets et de l'intelligence artificielle :

- ⊗ **Le chapitre 1** élabore les généralités dans l'internet des objets.
- ⊗ **Le chapitre 2** passe en revue les généralités dans l'intelligence artificielle et ses points de convergence avec l'internet des objets.

La deuxième partie est composée de trois chapitres :

- ⊗ **Le chapitre 3** nous présentons les méthodes d'apprentissage supervisé et non supervisé dans l'intelligence artificielle avec leurs algorithmes applicatifs.
- ⊗ **Le chapitre 4** nous présentons et comparons les différentes solutions proposées pour minimiser la congestion dans les réseaux internet des objets.
- ⊗ **Le chapitre 5** est consacré à la mise en œuvre et l'implémentation d'une technique de gestion de congestion sur une architecture en cinq niveaux que nous proposons.

Enfin, nous terminons ce mémoire par une conclusion générale, puis une présentation de différentes perspectives, qui feront certainement l'objet de travaux de recherche ultérieurs.

PARTIE 1: Etat de l'art sur l'internet des objets et l'intelligence artificielle

Chapitre 1 : Généralités sur l'internet des objets

Contexte

Le développement des nouvelles technologies émergentes ouvre plusieurs voies de recherches dans le but de faciliter le quotidien des hommes. C'est ainsi que leur sphère d'influence s'étale et touche de nombreuses domaines. L'internet des objets, un de ces nouvelles technologies émergentes est appliqué dans des domaines à savoir :

La maison intelligente qui offre de nombreux services comme la surveillance de patients à domicile. Cette approche vise à aider les personnes âgées, malades ou dépendantes dans leurs routines quotidiennes. Dans l'optique de fournir de nouveaux services d'assistance dans les maisons intelligentes (« smart home ») afin de surveiller en continue leurs constantes physiologiques à moindre coût pour détecter de possibles dégradations de leur état de santé [10].

L'agriculture autre segment de système connectée, qui connaît un formidable essor. Elle offre, et continuera à offrir, plus de contrôle et de précision aux agriculteurs [11].

L'élevage intelligent, l'utilisation de colliers trackers permet désormais aux éleveurs d'effectuer un suivi plus précis de la santé, de l'alimentation, de la sécurité. En plus d'autres points de surveillance, Ils peuvent aussi géolocaliser leur bétail en temps réel [12].

L'industrie 4.0 est l'une des décisions importantes prises par les entreprises. Cette industrie comprend des groupes technologiques tels que les robots autonomes, l'internet des objets (IoT), l'intégration de systèmes horizontaux et verticaux. Ces systèmes apportent des changements de grande envergure, aux entreprises, non seulement par l'intégration et l'interconnectivité des personnes, des machines. En outre elle touche également aux modèles d'affaires et de création des valeurs, les processus et activités et le management organisationnel des entreprises [13]. Dans l'environnement, à travers des solutions de l'internet des objets la déforestation et les feux de brousse sont contrôlés via des capteurs.

L'internet des Objets et la Smart City sont des sujets nouveaux et innovants capables de créer de nouveaux services. En plus il crée de la valeur ajoutée à la ville de demain à travers des solutions d'avenirs [11]. C'est ainsi que nos travaux se basent sur ces villes intelligentes notamment dans le transport intelligent.

Le problème majeur dans le domaine du transport routier reste la congestion. Ainsi nous proposons un système de transport intelligent pouvant contribuer à la décongestion d'une ville donnée.

L'avènement des technologies microélectroniques a favorisé la naissance d'un nouveau concept nommé internet des objets qui est devenu une réalité et interpellent beaucoup de chercheurs scientifiques ces derniers décennies. Il repose sur l'interconnexion d'objets via des réseaux ayant des adresses uniques dont leurs points de convergence est l'internet [10]. Il a pour ambition de faire dialoguer les objets entre eux mais aussi entre les personnes dans une perspective d'un monde de connexion plus dense qui revêt plusieurs défis, de profondes transformations de la vie des individus, des entreprises et des institutions [11].

L'objet de ce premier chapitre est d'étaler les différents composants de ce système, leur technologie, leurs champs d'interconnexion et d'interaction et leurs domaines d'applications avant de décliner quelques limites pour afin augmenter sa sphère d'influence [12].

Quelques définitions

Pour mieux comprendre notre champ d'investigation, nous donnons, dès le début quelques définitions relatives à l'Internet des Objet. Il faudra alors noter que d'autres définitions existent selon le contexte et l'application.

Définition 1 : L'Internet des objets, ou IoT (Internet of Things), est un scénario dans lequel les objets, les animaux et les personnes se voient attribuer des identifiants uniques, ainsi que la capacité de transférer des données sur un réseau sans nécessiter aucune interaction humain-à-humain ou humain-à-machine [13].

Définition 2 : C'est l'extension d'internet à des choses et à des lieux du monde physique, qui représente des échanges d'informations provenant de dispositifs du monde réelle vers le réseau internet [14].

Définition 3 : C'est un réseau de réseau qui permet via des systèmes d'identification électronique normalisées et sans fil, d'identifier et de communiquer numériquement avec des objets physiques afin de pouvoir mesurer et échangé des données entre les mondes physiques et virtuelles [15].

Histoire et évolution de l'internet des objets

Des objets communicants à l'internet des objets, aujourd'hui on parle de l'internet de tout ou l'angélisme internet of everything (IoE), faisons feedback pour étudier ce qui englobe cette histoire intéressante pour mieux tracer son évolution voire les tendances pour en bénéficier de ces avantages optimaux.

À l'origine, les premiers dispositifs électroniques apparaissent avec le télégraphe, le fax et la radio. Dès 1926, Nikola Tesla prédit déjà un monde connecté, je cite « quand le sans-fil sera parfaitement appliqué, la terre entière sera convertie en un gigantesque cerveau ». En 1989, Tim Berners-Lee inventait le World Wide Web puis en 1990, on assiste à l'apparition des grille-pains connectés et autres machines à cafés reliées à internet. Néanmoins, ce n'est qu'en 2001 que le constructeur géant coréen LG crédibilise l'idée d'un électroménager connecté à Internet ; en l'occurrence, il s'agit d'un réfrigérateur.

En 2003 l'entreprise Violet a mis sur le marché la lampe dall connectée en wifi qui est composé de 9 LED. Grâce à ses 9 LED, le dispositif pouvait s'allumer de différentes couleurs en fonction de différents événements, liés à la météo, la Bourse, la pollution, les alertes Google ou encore des "envois de messages de couleurs" par sms ou email.

En 2005, un rapport de l'ONU sur l'Union Internationale des Télécommunications prévient : "de l'Internet partout nous sommes parvenus à l'Internet pour tout [16]. Les connections vont se multiplier et créer un nouveau réseau dynamique, un réseau de réseaux, l'internet des objets [17].

Nous vivons actuellement l'invasion imminente des objets connectés dans notre vie quotidienne avec notamment des voitures autonomes approche, et les objets connectés tels que les thermostats connectés font leur apparition dans nos maisons [18]. Rapidement, les géants du secteur, et notamment Intel, lancent des initiatives pour ne pas manquer cette opportunité fabuleuse qui semble en toute logique destinée à peser des milliards de dollars. Les livres traitant du sujet se multiplient, les études fleurissent, la sphère scientifique s'intéresse de près aux nouvelles applications de la connexion. Inutile de dire que les entreprises commerciales aussi.

Tandis que cette tendance semble prendre de l'ampleur, Cisco avance le nombre de 80 milliards d'objets connectés pour l'horizon 2025.

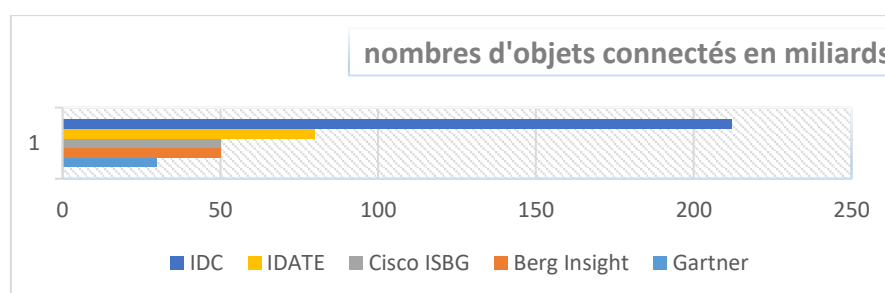


Figure 1 : Estimation du nombre d'objets connectés en milliards en 2022

Technologies de communication dans l'internet des objets

Étant le cœur de l'architecture de l'internet des objets, la communication occupe une place primordiale pour la bonne marche d'un projet IoT. Néanmoins cette communication englobe des catégories de réseaux qui se distinguent selon leur fonctionnalité donnant un choix à l'utilisateur selon son type de projet [19].

En ce sens l'utilisateur doit répondre aux questions concernant la connexion réseau, la distance, la portée, le type de source d'énergie, la topologie et tant d'autres [20] etc. L'étude des différentes technologies de communications va répondre aux questions suivantes : quel milieu ? À quel porté ? Quel volume de données à transporter ? Quelle topologie ? Quelle durée ? Quel type de source d'énergie ? [21] .

1. Technologies de courtes portées

a. Bluetooth

C'est une technologie sans fil d'une portée de 10m maximum dotée d'une faible bande passante lui permettant de transmettre une petite quantité de données. On retrouve cette technologie dans la quasi-totalité des téléphones portables. Sa communication se fait entre téléphones mais aussi avec des objets connectés de natures différentes tels que la montre connectée, l'oreillette de discussion téléphonique, le moniteur de fréquence cardiaque [22], etc.

b. Zigbee

Spécialement développé pour des applications domotiques, Zigbee à une portée moyenne de 10m offrant une faible bande passante pour des volumes minimaux. Elle fonctionne dans certains appareils de détecteurs de fumée [23]. Zigbee n'est pas gourmande en termes d'énergie [24].

c. NFC

Standard de communication radiofréquence sans fil très courte sur un rayon de quelques centimètres, on la retrouve aux niveaux des badges d'entreprises, d'abonnement de transport

2. Technologies de moyennes portées

a. Z-wave

Protocole de communication dédiée à la domotique, la portée de z-wave s'étend sur 30m en interne et inférieur à 100m en plein air. Elle consomme peu d'énergie et fonctionne avec de simples piles pouvant durer des années [23].

b. Wi-Fi

Très connu par les utilisateurs, le Wi-Fi est un protocole de communication sans fil sur une portée de 20m à 100m. Il est très gourmand en énergie et existe sous différentes normes avec des débits variables [20].

c. Bluetooth Low Energy

Cette technologie fonctionne comme le Bluetooth mais à une différence de portée (60m) et d'énergie consommant 10 fois plus que le Bluetooth [20].

3. Technologies de Longues portées

Les dérivés de la 4G : NB-IoT (Narrow Band IoT) de SFR et LT-ME d'Orange, assez similaire. Ils sont nouveaux dans le marché IoT, par rapport à LoRa et Sigfox, en termes de communication. Ils permettent d'envoyer de lourds messages [23, 25, 26].

a. Réseaux de types LPWAN

i. Sigfox

Sigfox est un réseau de communication radio avec une portée moyenne de 10km en ville et de 30 à 50km en milieu rural. Ce réseau fonctionne en basse fréquence et bas débit et peut transporter que de faible volume de données [23].

ii. Lora Wan

C'est un protocole de communication radio très basse consommation d'une portée de 2 à 5km en milieu urbain et peut atteindre jusqu'à 45km en milieu rural. Il est recommandé pour le transfert de petit volume de données comme le cas des capteurs qui émettent périodiquement.

b. Les Réseaux Cellulaires

Ces réseaux sont obtenus grâce à des opérateurs de télécommunication qui sont basé sur la technologie GSM, pouvant transmettre des gigantesques volumes de données sur de longue distance et très énergivore en l'occurrence de la 4G et 5G [23].

c. Les Réseaux propriétaires

Ce sont des réseaux privés conçu spécialement pour les grandes entreprises dotées de chiffre d'affaire conséquent. Ils ont des bandes passantes dédiées pour un échange optimal et sécurisé de leurs données [23].

d. Fibre Optique

C'est une technique de communication qui permet d'utiliser un fil très fin de silice pour transmettre des signaux numériques, très haut débits guidés par la lumière sur des milliers de kilomètres [27].

Architecture dans l'internet des objets

L'internet des objets utilisé dans plusieurs domaines et composé de technicités diverses est constitué par plusieurs couches définies selon de nombreuses structures de recherche. Néanmoins, l'architecture proposée par Cisco reste la référence pour mieux définir l'internet des objets avec plus de précision pour ce paradigme de grand potentiel pour les nouvelles technologies de communications sans fil.

Ce modèle aide à décortiquer les systèmes complexes composant ce dernier pour une compréhension de chaque sous partie, et plus de précision de chaque niveau afin de bien fusionner les produits et services. Par ce modèle nous pouvons identifier tous les types spécifiques de traitements optimisés des composants du système. Cette communication entre produits et systèmes rend le concept « IOT » plus qu'une réalité concrète.

C'est ainsi qu'une architecture orientée services a été proposée par (Xu, He et Li, 2014). Cette architecture se définit sur quatre (4) couches que sont :

1. L'objet

L'objet devenu intelligent avec la capacité de communiqué avec son entourage. En ce sens le capteur placé sur l'objet doit répondre aux exigences du domaine de l'objet qui est souvent une ville, une maison, une entreprise, une usine [28], etc ;

2. Le réseau

L'objet devenu intelligent avec la capacité de communiqué avec son entourage. En ce sens le capteur placé sur l'objet doit répondre aux exigences du domaine de l'objet qui est souvent une ville, une maison, une entreprise, une usine [10], etc ;

3. Logiciel et infrastructure

À ce niveau on retrouve le choix de la technologie utilisé meilleur configuration meilleur compréhension de l'outil, des types qu'offre le fournisseur. C'est à ce stade qu'intervient les plateformes IoT, leurs choix vont jouer sur le fonctionnement de la chaîne notamment de la sécurité des données collectées ;

4. Applications et services

La technologie IoT est laissée en rade au niveau de cette couche. En effet, la quantité de données (data mining) captée est privilégiée au détriment de celle qualitative. Ces données sont utilisées pour des cas de besoins de services. Cette couche freine en quelques sortes le développement de l'IoT, pour échapper à ce blocus des compétences avérés en ce domaine seront souhaiter. A ces fin décideurs et acteurs doivent penser de lourd investissement mais aussi une collaboration des fabricants et informaticiens pour un duo gagnant [28] ;

Cependant, l'architecture proposée par Cisco reste le modèle de référence afin de bien structurer par niveau donc chaque terminologie peut être une tendance.

Les réseaux traditionnels de communication de données ont de multiples fonctions, comme en témoigne le modèle de référence à 7 couches de l'Organisation internationale de normalisation (ISO) [29]. Cependant, un système IoT complet contient de nombreux niveaux en plus du réseau de communication ;

En effet, dans [14] le modèle définit les fonctions nécessaires pour qu'un système IoT soit complet (CISCO, 2014) la figure suivante montre les sept (7) différentes couches de l'internet des objets :

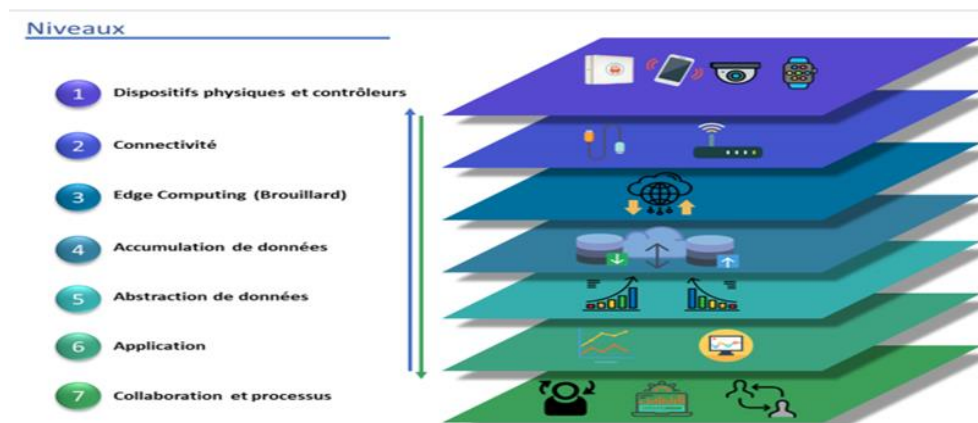


Figure 2 : Le modèle de référence de l'IOT (Cisco,2014)

Niveau 1

Couche 1 : périphériques physiques et contrôleurs pouvant contrôler plusieurs périphériques. Ce sont les « objets » de l'Internet des objets, et ils comprennent une large gamme d'appareils de terminal qui envoient et reçoivent des informations ;

Niveau 2

Les communications et la connectivité sont concentrées au niveau 2. La fonction la plus importante du niveau 2 est la transmission fiable et opportune de l'information ;

Niveau 3

Edge Computing (Brouillard), les fonctions de niveau 3 sont dictées par la nécessité de convertir les flux de données du réseau en informations appropriées pour le stockage et le traitement de niveau 4 (hiérarchisé) ;

Niveau 4

Le niveau 4 convertit les données basées sur des événements en traitement basé sur une requête. C'est une étape cruciale pour surmonter les différences entre le monde des réseaux en temps réel et le monde des applications en temps non réel ;

Niveau 5

Abstraction de données, les systèmes IoT devront évoluer au niveau de l'entreprise ou même au niveau mondial et nécessiteront plusieurs systèmes de stockage pour prendre en charge les données IoT provenant des systèmes ERP (« Entreprise Resource Planning »), HRMS (« Human Resource Management System »), CRM (« Customer Relationship management ») et d'autres systèmes d'entreprise traditionnels.

Niveau 6

Application, le niveau 6 est l'endroit où l'interprétation de l'information se produit. Le logiciel à ce niveau interagit avec le niveau 5 et les données au repos.

Les applications varient en fonction des marchés verticaux, de la nature des données de l'appareil et des besoins de l'entreprise.

Niveau 7

Collaboration et processus l'une des principales distinctions entre l'IoT et le système IoT est que l'IoT inclut les personnes et les processus. Cette différence devient particulièrement claire au niveau 7 (Collaboration et processus).

Le système IoT et les informations qu'il crée ont peu de valeur à moins de produire une action qui nécessite souvent des personnes et des processus.

Domaines d'applications

L'application de la technologie de l'internet des objets touche de plus en plus de nombreux domaines, ces usages changent le quotidien du particulier, de l'entreprise. Ses retombés positifs vont élargir le champ d'application ou mettre à jour ceux existant. La sphère d'application est assez large mais citons quelques-uns [30].

1. Industrie 4.0

L'industrie dans l'internet des objets est l'ensemble de moyens mis en œuvre dans une chaîne de production en communication permanente sans l'intervention humaine. L'objectif est d'aboutir à une meilleure industrie plus compétitive par une synchronisation de tous les éléments avec l'optimisation en temps et en énergie [31] [32].



Figure 3 : l'industrie 4.0

2. Santé connectée

L'internet des objets touche largement le domaine de la santé, le système de télémédecine est en pleine mutation suivant des identités pour capter les données de patients en temps réelle.

Cette technologie participe au désert médical, on assiste à une consultation des malades dans les zones les plus reculés qui crée ainsi un vide entre patient et soignant [33]. La sécurité doit être également évoquée pour la gestion des données clients [30].

3. Transport connecté

À l'échelle d'une ville on peut contrôler le trafic du transport, identifié les voies libres, gérer la pollution mais aussi piloter les voitures automatiques. Du côté de l'aviation, la maintenance

prédictive et l'optimisation des opérations des avions connectés optimisent leurs voies ainsi du fuel est économisé de (1,5%). Le repérage de parking connecté à proximité est un plus pour économiser l'énergie dans le transport routier [33].

4. Domotique

La domotique peut être défini comme étant l'ensemble des techniques permettant de centraliser tous nos appareils de ménage via un smartphone [34].

L'éclairage est géré grâce à un smartphone, les heures de repos sont synchronisées avec du son, parfois de la lumière ou bien même un verrou connecté en veille [35]. Un client va désormais payer ce qu'il consomme ce qui est devenu une réalité avec les compteurs d'eau et d'électricité connectés [33, 36].

5. Smart City

La ville est désormais connectée, la consommation d'énergie est bien gérée par les institutions connectées. Les rues sont bien surveillées en cas d'incident ou d'accident, dans le cas où l'auteur prend la fuite [37].

Le trafic est obtenu en temps réelle notamment on a une vision claire des espaces publics et parkings, l'éclairage de la ville se fait en fonction des déplacements des citoyens, les réseaux d'eau, de gaz sont surveillés.



Figure 4 : la ville connectée [38]

6. Transport logistique connecté

La logistique représente 90% du commerce mondiale, ce secteur est en plein évolution avec l'internet des objets qui se charge de la gestion des stocks à la livraison des colis. Ces données recueillis doivent être à jour, fiables et authentique. Ce travail va partir d'une position de localisation GPS passant par le réseau, puis le cloud pour livrer les données au client concerné [37]. À cela s'ajoute le respect des réglementations, des standards de qualité, la satisfaction du client, respect des délais et la sécurité.



Figure 5 : le transport logistique connectée [28]

7. Agriculture de précision

Le développement de véhicules connectés va toucher en premier l'agriculture intelligente. Ces tracteurs, moissonneuses et batteuses autonomes vont servir d'application, de l'agriculture dans l'internet des objets [33]. De surcroît, des capteurs positionnés dans ses serres offrent au cultivateur la possibilité de contrôler les conditions de température et d'humidité à distance ainsi, d'améliorer sa prise de décision [37].



Figure 6 : agriculture de précision

8. Élevage intelligent

La surveillance de bétail est désormais facile, chaque animal d'un troupeau peut être équipé de capteurs mesurant fréquences cardiaque et respiratoire, taux de pression sanguine ainsi que température, les déplacements, en plus de détecter une éventuelle maladie [37].



Figure 7 : élevage intelligent

9. Limites dans les domaines d'applications

Les applications décrites précédemment sont conçues sous des formes diverses qui laissent prendre en charge plusieurs domaines. Cette prise en charge large ne laisse pas apparaître beaucoup de choix pour les utilisateurs. Ainsi s'ils veulent spécifier le domaine d'application, ou bien sécuriser à leur manière; ils sont bloqués pour personnaliser à leur guise.

Plateformes dans l'internet des objets

Une liste de plateformes existant dans l'internet des objets montre les différents critères d'analyse des applications de celui-ci. Beaucoup de questions se posent par rapport à l'hétérogénéité des objets intégrant ces plateformes ceci implique la façon dont ils sont traités, stockés et analysés pour y récolter des résultats. La question est comment gérer les données issues de ces applications. Ils existent deux types de plateformes : verticales et horizontales [39].

Définition

Considéré comme un ensemble de services, la plateforme permet de collecter, de stocker, et de corréler les données recueillies.

Ces données participent à la prise de décision, au développement, à la génération de nouvelles opportunités business de l'entreprise, au gain de temps personnel. Ces failles de sécurité risquent de bouleverser la vie d'un particulier, d'une structure, d'un espace public pour ne citer que cela ...

Ainsi les bases de données de services de traitement, web, visualisation de données des tableaux de bord et d'exploitation seront incluses pour la bonne marche des plateformes.

1. Types de plateformes dans l'internet des objets

Les plateformes IoT existent sous deux formes qui ont des caractéristiques généralistes et spécialisées mais comment se penser d'un côté ? Faisons une étude des deux plateformes citées ci-dessous :

a. Plateformes horizontales

Ce sont des types qui sont développés sur mesure, qui permettent d'agréger des applications IoT. Apparus le premier sur le marché les plateformes horizontales sont constituées de quatre segments :

-les plateformes pour le développement de périphériques : c'est un type de plateformes qui émerge dans la catégorie horizontale avec des produits innovants.

Ces produits IoT vont communiquer via la technologie LPWAN tel que Sigfox, LoRa, NB-IoT des mises à jour sécurisé pour un ensemble d'objets situant sur un espace bien défini avec les fournisseurs Microsoft et AWS.

-Les plateformes pour la gestion des périphéries : assurent les fonctionnalités critiques des parcs d'objets connectés à grande échelle. Elles gèrent la mise à jour des logiciels centralisés, leurs configurations et le contrôle à distance.

-Les plateformes pour les applications analytiques : Elles permettent de développer des applications complexes pour des usages, de maintenance prédictive. Elles gèrent l'analyse des données, l'intégration d'application, l'intelligence artificielle.

-Les plateformes pour un déploiement rapide : facilitent l'interaction utilisateur et les données issu de l'objet connecté qui sont visualisés via un Dashboard ou un glisser-déposer. On la retrouve souvent chez les nouveaux clients de l'industrie IoT.

- Plateformes verticales

Ces plateformes sont développées pour des solutions de types spécifiques elles proposent un ensemble de solutions pour divers cas avec des fonctionnalités (interface adapté, notification en temps réelle, automatisations des rapports. On les retrouve dans les secteurs de villes intelligents, l'industrie connecté, de véhicules connectés.

Ils sont composés de trois types de segments :

-Le choix d'une plateforme pour la ville connectée : exploites les données des infrastructures d'une ville connectée, ces données sont souvent délivrées aux collectivités locales. Ils peuvent servir d'aide à la prise de décision pour une meilleure condition de vie des citoyens. Les fournisseurs principaux sont FIWARE, Itron, Urban Software Institute.

-Une plateforme pour l'industrie connectée : joue le rôle de suivi des opérations de productions, des travailleurs, des processus en tant réelle pour plus d'efficacité de la chaine de production. Ces fournisseurs sont : PTC, GE Digital.

- Une plateforme pour les véhicules : offre une solution axée sur trois angles, mise en place de services connectés pour le conducteur, les services liés au fonctionnement du véhicule, et la communication des données du véhicule au back end pour des besoins ultérieurs. Ces services sont fournis par Bosch, Arman IBM et Microsoft. Les plateformes de services jouent un rôle fondamental pour la création et gestion des applications de l'internet des objets.

Elles ont la capacité de bien coopérer les objets, le réseau, les technologies afin de réaliser une architecture fonctionnelle de ses composants.

Discussion

Le constat est que toutes les plateformes traitent les mêmes cas c'est-à-dire rendre l'objet un peu intelligent en répondant aux commandes à distances, sécuriser les communications appareil-appareil ou appareil-services cloud, l'envoi et la réception des données.

Ces objets dits intelligents sont connectés à un service cloud qui détient les masses de données collectées puis analysés pour des décisions futures.

Le cloud héberge ces plateformes ainsi un couplage de l'internet des objets et du cloud pour des applications plus spécifiques. Ce dernier garanti un certain nombre d'avantages tels que l'optimisation de l'utilisation des ressources matérielles et logiciels, flexibilité, élasticité. D'autre part il peut bénéficier de la capacité de stockage et de traitement du cloud.

Composants de l'internet des objets

L'écosystème de l'internet des objets couvre un ensemble d'acteurs dont le lien est irrévocable. Les composants sont constitués des fabricants de puces et de modules électroniques, des consommateurs en passant par la connectivité, des plateformes et des développeurs d'applications.

1. Fabricants de capteurs

Fabricant de puces et de modules électroniques, ils font la base de l'IoT sans ce maillon il n'y aurait pas de capteur, de réseaux de capteurs pour capter les données mais aussi de puce RFID pour l'identification par radiofréquence. Texas Instrument, Semtech, font partie de la liste des fabricants.

2. Objet connecté

Deuxième maillon de l'écosystème à ce stade l'objet est muni de capteurs et de transmetteurs lui permettant de dialoguer avec son entourage. Cet assemblage de puces, d'antennes, d'émetteurs-récepteurs, de capteurs rend cet objet « intelligent ».

3. Connectivité

Troisième maillon de la chaîne, pour qu'il est d'internet des objets il faut que la donnée captée transite via un réseau internet de type choisi selon la localisation des données LAN, MAN, WAN.

Des opérateurs nouveaux apparaissent dans ce domaine comme Sigfox, Lora (technologie LoraWan), Qowisio, Bouygues Télécoms (technologie Lora) et Orange (4G, 5G) dans l'ambition de couvrir la transmission de données de la source à la destination.

4. Plateformes

Quatrième position de la chaîne, les plateformes IoT assurent le stockage des données transmises par le réseau, les traitent et les rendent exploitables par les applications dédiées à l'utilisateur final [40]. Les fournisseurs comme Microsoft, Amazon offrent des services spécifiques aux clients souscrivant pour plus d'opportunités, l'utilisateur peut acheter et l'installer sur son système d'exploitation ou chez un hébergeur. En plus d'une carte de tableau de bord, traitement et paramétrage des données dans le but de faciliter la tâche aux développeurs d'applications.

5. Développeurs d'applications

Enfin les applications qui nous livrent directement les données exploitables par le client final en vue du phénomène étudié (zone à surveiller, pollution, climat). Ces applications mettent en valeur les données issues des objets connectés en gérant des données inestimables.

Force est de constater que les acteurs de l'écosystème IoT se penchent vers d'autres alternances, c'est à ce stade qu'intervient l'intelligence artificielle qui s'ajoute au cloud pour une meilleure gestion des données.

6. Technologies pour une gestion optimale de données IoT

Pour une étude plus approfondie des données générées par l'internet des objets, il existe des technologies tels que :

-L'intelligence Artificielle : La masse de données récoltée par l'IoT nécessite une analyse approfondie de ces données hyper dynamiques. Une réponse efficace apportée par l'intelligence Artificielle permet d'identifier, classer et restituer en temps réel les données reçues au format compréhensible pour les décideurs ;

-Le data science, avec sa capacité d'étudier une importante masse de données brutes, il en ressort des informations pertinentes pour le marché notamment des stratégies marketing et commerciales

-Le Cloud, offre une possibilité de stockage de données de grandes quantités. Ces data centers redondants dispersés à travers le monde permettent de restituer des données perdues en cas de catastrophe naturelle ou de problèmes liés à la sécurité.

Procédures de normalisations

L'interopérabilité entre les objets connecté reste un problème majeur ainsi pour éviter la défragmentation de l'écosystème internet des objets, plusieurs organismes et sociétés scientifique dite International Standardisation Organization (ISO) se sont déterminées, bien qu'une norme internationale s'avère imprévisible, dès lors naissent des normes régionaux et nationaux et industriels se mette en place.

1. Initiatives d'organisations internationales

De nombreuses initiatives se multiplient à l'échelle internationale pour une architecture globale de l'internet des objets.

L'objectif est de parvenir à l'élaboration de standards intersectoriels pour des outils utilisés par tous les secteurs. Plusieurs initiatives sont à noter, comme celle de l'Union internationale des télécommunications (UIT), qui cherche à élaborer des lignes directrices qui serviraient de référence commune aux autres organismes de standardisation. D'autres initiatives sont en cours, comme les travaux de l'Institut des ingénieurs électriciens et électroniciens (IIEE), qui travaille sur la standardisation des réseaux de communication, sur des applications plus sectorisées (smart grid, industrie, agriculture et secteur minier) et sur un projet de norme (P2314) visant à définir un cadre architectural pour l'Internet des objets. Enfin, l'Internet Engineering Task

Force (IETF) élabore des standards pour les systèmes de communication, notamment pour l'IPv6.

2. Initiatives de l'union Européenne

En Europe, l'Alliance pour l'innovation de l'Internet des objets (AIOTI, Alliance for Internet of Things Innovation), est lancée en 2015. Des structures telles que ETSI (European Telecommunications Standards Institute), CEN (Comité européen de normalisation) travaillent pour le même marché. En plus des normalisations sectorisées (M/490 smart grid, M/441 compteurs intelligents et M/468 véhicules électriques) apparaissent.

3. Consortiums pour une normalisation spécifique de l'IOT

Pour répondre aux besoins spécifiques de chaque secteur, plusieurs groupes dirigés par des consortiums d'acteurs industriels se sont formés. Le groupe One M2M a développé une suite de standards destinée à l'usage des objets M2M, créa un ensemble de solutions de sécurité.

L'Industrial Internet Consortium (IBM, AT&T, Intel...) développe des cas d'usage, des cadres et des architectures de référence pour le secteur industriel de l'IoT.

L'alliance AllSeen (Qualcomm, Microsoft, Cisco, LG...) planche elle aussi sur des normes destinées à l'industrie et développe en parallèle un logiciel open source (AllJoyn). Enfin, le groupe Thread (Samsung, ARM, Freescale, Huawei, Nest / Google...) s'intéresse à la standardisation du secteur de la domotique.

Obstacles qui ralentissent le développement de l'internet des objets et perspectives

L'internet des objets devrait dépasser son stade actuel vu ces débuts de réalisations, la définition de normes, problème d'alimentation des capteurs, la confidentialité des données utilisateurs, des compétences décisionnelles. Sur ceux plusieurs acteurs et chercheurs tentent d'apporter le meilleur au bon fonctionnement de l'IOT [41].

1. Besoin d'une norme internationale

L'internet des objets est formé de plusieurs systèmes jusqu'à l'heure actuel on peine à avoir une norme internationale. Les grandes entreprises utilisent des normes internes, des organismes régionaux en utilisent d'autres ainsi que les particuliers.

2. Problèmes d'alimentation des capteurs

L'alimentation des capteurs se fait principalement via des batteries, leur fonctionnement nécessite une forte consommation en énergie. Vu que ces piles demeurent irremplaçables, et leurs durées de vie s'étendent à des mois voire des années ; il urge de proposer des moyens techniques pour économiser l'énergie.

3. Confidentialité des utilisateurs

De peur que leurs données atterrissent sur les mains des pirates, bon nombre d'utilisateurs ont peur d'utiliser l'internet des objets vu que la sécurité reste loin d'être atteint.

4. Compétences décisionnelles insatisfaites

Les compétences de l'IOT se résume à identifier, capter, transférer de manière autonome, il doit dépasser ce stade à l'heure où l'on parle de donnée « or », l'extraction de cette dernière devient

prioritaire d'où l'application de l'intelligence artificielle pour plus de compétences décisionnelles.

On assiste à de nouvel apprentissage à travers chaque processus via des infrastructures de calcul et des technologies algorithmiques : ce qui sera l'objet d'un prochain chapitre.

Internet de tout (IoE)

L'internet de toute chose ou l'acronyme internet of everything en anglais (IoE) présente une vision plus large de l'internet des objets. Le réseau est distribué et décentraliser, doubler d'une présence permanente de l'Intelligence Artificielle à tous les niveaux, protégeant ainsi les réseaux avec des données personnalisable et utile à la prise de décision comme Cognitive IOT, Machine Learning, Deep Learning [42]...

Conclusion

Bref l'internet des objets justifie sa réalité dans plusieurs domaines via des plateformes avec des enjeux et perspectives prometteuses. Cet élément de concept fruit d'une histoire mouvementée en ressort des avantages et limites. La sécurité et la flexibilité de l'IoT seraient un atout pour ces adeptes. Notre objectif est d'apporter plus de flexibilité et d'intelligence sur son fonctionnement. Ce qui conduit à l'étude des différents aspects relatifs à l'intelligence artificielle afin de voir son application dans l'IoT.

Chapitre 2 : Intelligence artificielle et l'internet des objets : enjeux et applications

De l'internet des objets (IoT), on passe à l'informatique cognitive c'est-à-dire la capacité d'une machine à imiter la pensée humaine [43] parfois même dépassé celle-ci grâce à l'incorporation de cette intelligence dite artificielle. Ainsi on note la présence de l'intelligence artificielle dans l'exploitation des données générés par l'internet des objets. Cette collaboration est fruit de l'exploitation instantanée de données par l'intelligence artificielle (IA).

L'objectif de ce chapitre va d'abord décrire l'intelligence artificielle dans toutes ces facettes et ensuite analyser l'étude de l'existant de l'application de l'intelligence artificielle sur l'internet des objets.

I. Qu'est-ce que l'intelligence artificielle

L'intelligence artificielle (IA) est la simulation de processus d'intelligence humaine par des machines [44]. Notons qu'en quelques décennies l'IA est passée du concept à la réalité.

Ainsi des ordinateurs sont programmés pour « imiter » le comportement humain. En outre, ces machines exploitent des quantités massives de données pour déterminer des comportements similaires. Cela peut être la reconnaissance des différences entre un chat et un oiseau, la reconnaissance d'images d'une scène de crime ou l'exécution d'activités complexes dans une usine de fabrication [45]. Bien vrai que les visions diverses d'un pays à un autre, de nombreuses questions se posent par rapport à ce qu'est l'intelligence artificielle mais faisons un tour d'horizon pour voir quelques définitions données à cette technologie.

1. Définitions

Le terme d'intelligence artificielle s'avère un langage commun dont l'utilisation est devenue populaire à travers ses diverses applications et définitions variés :

Définition 1

L'intelligence artificielle (IA) est la science dont le but est de faire faire par une machine des tâches que l'homme accomplit en utilisant son intelligence [46] ;

Définition 2

L'intelligence artificielle (IA, ou AI en anglais pour Artificial Intelligence) consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle [47] ;

Définition 3

L'intelligence artificielle peut être définie comme un écosystème formé d'un ensemble de technologies, de sciences théoriques et techniques qui font appel à la statistique, la logique mathématique, probabilités, neurobiologie computationnelle et informatique dont l'objectif est de fournir la machine les capacités cognitives d'un être humain [48].

2. Histoire et évolution de l'intelligence artificielle

L'histoire de l'intelligence artificielle s'avère peut longtemps, mais retenons quelques dates récentes pour tracer une chronologie. L'intelligence artificielle se trouve probablement dans l'article d'Alan Turing « *Computing Machinery and Intelligence* » (Mind, octobre 1950), où Turing explore le problème et propose une expérience maintenant connue sous le nom de test de Turing [49].

1956 - Une première conférence est organisée par les pionniers de l'IA il s'agissait de John MCCARTHY [50], Marvin MINSKY, Allen NEWELL, Herbert SIMON [51].

1960 - Des experts du Massachusetts Institute of Technology (MIT) émettent de sérieuses réserves concernant l'approche neuronale en intelligence artificielle qui désintéressent les chercheurs du domaine [52].

Dans les années 1970-1980, Geoffrey Hinton, Yann LeCun et Kunihiko Fukushima conçoivent des réseaux de neurones artificiels à plusieurs couches, dont l'arrangement est inspiré du cortex visuel, permettant l'apprentissage de tâches plus complexes. Toutefois, la quantité de données disponibles et la puissance de calcul demeurent insuffisantes. Ces concepts sont remisés pendant des années

1990 - Des avancées sont faites dans la reconnaissance d'images, de l'écriture manuscrite et du langage humain et dans la production artificielle de la parole, grâce notamment aux réseaux de neurones artificiels.

En 1997, Gary Kasparov, le maître incontesté des échecs, est battu par Deep Blue, le superordinateur d'IBM. Son secret : sa puissance de calcul. Chaque seconde, Deep Blue analyse 200 millions de possibilités. Son adversaire, seulement 3. Cet affrontement marque l'imaginaire. Désormais, l'ordinateur rivalise d'intelligence avec l'humain.

En 2012, en Juin : Google parvient à entraîner un ordinateur à reconnaître par lui-même un chat, en lui faisant ingurgiter des millions d'images obtenues sur Internet. Une autre victoire pour les partisans de l'apprentissage profond. En Août, Google commence à utiliser les réseaux de

neurones artificiels pour améliorer ses outils de reconnaissance vocale. En Octobre, Geoffrey Hinton, et son équipe remportent la prestigieuse compétition ImageNet, consacrée à la reconnaissance d'images. L'événement est confirmé le potentiel de l'apprentissage profond. En 2014, les investissements privés en apprentissage profond explosent.

En 2015, en Mai Google lance un outil qui permet de classer automatiquement des photos en fonction des personnes et des objets présents, ou encore des lieux où elles ont été prises.

Puis en Décembre Microsoft fait mieux que les humains dans la reconnaissance visuelle (concours ImageNet).

En 2016, Mars, lors des jeux Alpha GO, le logiciel de Google Deep Mind, bat le champion du monde sud-coréen au jeu de go, une victoire qui n'était pas attendue avant une quinzaine d'années. En Septembre l'outil de Google est grandement améliorée grâce à l'apprentissage profond. En Octobre l'ordinateur prouve ses compétences dans le domaine de la reconnaissance vocale par Microsoft. En 2017, l'IA Libratus bat les quatre meilleurs joueurs mondiaux de poker.

3. Fonctionnement

La réussite de la révolution technologique est facilitée par une combinaison des données et de l'informatique. Selon Harry Shum : « une vaste quantité de data ; une puissance informatique extraordinaire ; notamment le Cloud ; les algorithmes révolutionnaires, basés sur le deep-learning ». Les systèmes d'IA identifier l'importance d'un problème par [53] :

- une étude de toutes les variables possibles liées au problème à étudier ;
- une prédiction du résultat de ce problème toujours à partir de données connues [54].

Le machine Learning

La machine learning est l'apprentissage automatique des machines par des données. Ainsi la machine va s'entraîner ce qui va permettre d'approfondir sa compréhension, son autonome sans l'intervention humaine via des types d'apprentissage basés sur des algorithmes [55].

1. Apprentissage supervisé

Ce type d'apprentissage consiste à donner à la machine des données en entrées qui devra fournir les résultats attendus. L'étude se fait suivant un modèle défini par l'algorithme choisi.

2. Apprentissage par renforcement

Ce type d'apprentissage est l'étude d'un comportement d'un agent artificiel. C'est ainsi que beaucoup de chercheurs se penchent sur l'idée d'apprendre par renforcement. Cette technique consiste à subdiviser un problème complexes en plusieurs sous- problèmes l'agent sera récompensé ou punie en fonction des actions réalisé [56].

3. Apprentissage non supervisé

C'est une méthode qui consiste à fournir à une machine un ensemble de données non étiquetés à partir desquels la machine va apprendre à développer son algorithme de corrélation.

4. L'apprentissage profond

C'est le Deep Learning qui est un ensemble de réseau de neurones imitant le fonctionnement du cerveau humain. C'est ce réseau qui crée une machine virtuelle composés de milliers d'unités dont chacun se charge de simples petits calculs.

Domaines d'applications

Depuis quelques décennies l'intelligence artificielle est devenue active et attractive au cœur des services de l'homme. Cette omniprésence visant à représenter la capacité de réaction de ce dernier dans plusieurs domaines n'est plus à démontrer.

1. L'application de l'IA dans le domaine de la santé

Les préoccupations majeures des professionnels de la santé restent la gestion des maladies et des patients. On assiste à un renforcement des liens, des patients sont désormais acteurs de leurs propres santés. Ils posent des diagnostics rapides et précis, des traitements optimisés.



Figure 8 : l'intelligence artificielle dans le domaine de la santé

2. L'industrie du futur

La quatrième révolution industrielle est marquée par l'utilisation de capteurs communicants, le secteur industriel devient intelligent, on assiste à un nouveau système de productivité, l'entreprise numérique est devenue une réalité [57]. Les données sont générées, traitées, analysées en continu, la maintenance machine est désormais prédictive ceci va booster la production en plus des assistants virtuels qui alertent en cas de présence de personne n'ayant pas accès au système. En plus les produits en sortant de l'usine sont étiquetés, bien sécuriser.



Figure 9 : l'IA dans le domaine de l'industrie

3. L'intelligence artificielle au cœur du transport

L'IA occupe une place importante dans le secteur du transport avec l'avènement des voitures autonomes, le transport a subi d'énormes bouleversements. Dès lors qu'on assiste à une réduction drastique des accidents de la route et de la diminution de la congestion des parkings avec plus d'économie de consommation d'énergie.



Figure 10 : l'IA dans le domaine du Transport

4. L'intelligence artificielle pour la sécurité des personnes et des biens

Face à la complexité et au volume croissants des cyberattaques, l'intelligence artificielle (IA) aide les analystes des opérations de sécurité à garder une longueur d'avance sur les menaces [58] [59].

C'est dans ce sens que les acteurs du numérique adoptent l'application des technologies de l'IA pour de nouvelles formes de sécurité [60]. C'est l'exemple de la France qui fusionne le commissariat à l'énergie atomique aux énergies alternatives en partenariat avec Thales utilisent le Deep Learning pour une analyse et une interprétation automatique des vidéos d'applications de sécurité [61].

5. L'agriculture intelligente

La naissance d'une nouvelle forme d'agriculture facilitée par l'IA a donné naissance à de nouvelles machines qui ont tendance à supplanter la présence humaine dans les champs [62].

Avec l'amélioration de la précision des technologies informatique cognitive, qui via la reconnaissance d'images permet de distinguer une plante mur va révolutionner l'agriculture ancienne.

La tendance est que plusieurs startups fournissent des solutions innovantes pour tirer le meilleur de l'IA dans l'agriculture qui est utilisé dans de nombreux pays [63]. Plusieurs des exemples sont à noter :

Le développement d'algorithmes d'apprentissage profond capable de détecter l'état du sol et d'offrir plus de rendements ;

L'arrivée des « robots cueilleurs » dotés des systèmes d'IA avec des caméras dotées de capteurs fournissent des images en temps réel pour un champ donné avec plus de tâches réussies que les humains [64] [65] ;

Des modèles d'apprentissage automatique sont utilisés pour contrer la main-d'œuvre analyse prédictive pour la suivie de l'évolution des rendements [66].

6. L'Intelligence Artificielle au service de la domotique

La présence d'assistants virtuels a un impact majeur du quotidien des utilisateurs des domaines [67]. Directement reliés aux plateformes d'intelligence artificielle, ces assistants allument ou éteignent le chauffage de la maison connectée, gère l'agenda, commande les courses, font le choix de la musique en plus des recherches sur l'internet donc ils deviennent nos fidèles compagnons [68] [69] [70].

Défis-Enjeux et perspectives de l'Intelligence artificielle

1. Défis de l'intelligence artificielle

Défis Sociétale

Les acteurs du développement doivent donc s'engager activement dans un dialogue avec les autorités chargées de la protection des données et les responsables du traitement de ces données et la société civile [71].

Il est crucial d'analyser la portée et l'impact des applications d'IA et les défis qui y sont associés. Toutefois, les recherches précédentes ont examiné les applications de l'IA que de manière isolée et fragmentaire. Dans [72], les chercheurs ont synthétisé les publications scientifiques sur l'IA en vue de dresser un portrait d'ensemble des utilisations de l'IA dans le secteur public

Défis Technologiques :

L'intelligence artificiel promet une demande croissante de solutions. En plus du prestige de la marque, les entreprises doivent apporter des résultats réels, une expérience singulière et plus de valeur ajoutée seront privilégiées [73] .

À une époque où les chercheurs, les développeurs et les scientifiques effectuent des percées dans le domaine de l'intelligence artificielle à une vitesse incroyable et ce, dans divers domaines. Certains concepts juridiques devront inévitablement être adaptés pour faire face aux défis que ces percées amèneront. Ainsi l'essentiel est d'être conscient des risques juridiques associés aux importantes percées dans le domaine de l'intelligence artificielle et de prendre des décisions éclairées dans le cadre de la gestion du développement et de l'utilisation de l'intelligence artificielle. [74].

2. Enjeux de l'intelligence artificielle

Les principaux enjeux de l'IA restent se retrouve dans l'économie et de la société qui seront accompagner par un changement de mentalité de la masse sociale afin qu'elle comprend le meilleur de l'IA.

a. Enjeux économiques

De nos jours, le modèle économique est en plein changement suite à l'évolution du numérique et des technologies qui en découlent, à savoir l'intelligence artificielle (IA), le big data, les biotechnologies, la robotique ou encore l'internet des objets. Depuis lors, les progrès dans ce domaine ont été spectaculaires et incessants. La disponibilité croissante de grands ensembles de données générés dans tous les domaines (agriculture, éducation, santé humaine, commerce, communications) et portés par les progrès continus de la puissance de calcul et du développement d'algorithmes mais aussi des techniques d'apprentissage automatique et améliorées [75] [76].

Ainsi, l'IA ouvre une nouvelle ère de disruption et de croissance, où l'intelligence humaine est renforcée par la rapidité d'exécution et la précision. Dès à présent tout entreprise désirant intégrer l'intelligence artificielle va devoir relever ses défis pour pouvoir adopter l'IA et décrit

les mesures qu'elle pourra prendre pour poursuivre le déploiement de l'IA, de l'adoption à l'expansion dans sa structure.

b. Enjeux Sociétales

- Voir les compétences : ces informations n'apparaissent pas sur les CV classiques. L'intelligence artificielle permet d'en savoir plus sur la personnalité du candidat, sa trajectoire professionnelle, son potentiel, ses aptitudes à travailler en équipe, à innover, etc. Elle offre en outre une profondeur d'analyse des profils, en quelques secondes;
- Prendre la meilleure décision : l'IA est en mesure de fournir au recruteur plus d'informations pour une plus grande personnalisation dans l'approche. Elle nous fournit toutes les cartes pour rédiger un message impactant. Aussi l'IA par sa profondeur d'analyse, est capable de déterminer si un talent déjà en poste est à l'écoute du marché;
- Créer plus d'engagement : trouver un candidat qualifié, d'accord mais comment être sûr qu'il sera motivé au quotidien dans l'entreprise ? La profondeur d'analyse évoquée précédemment permet d'avoir une idée générale du potentiel du candidat ainsi que de sa réussite future une fois en poste. Et le temps dégagé grâce à l'IA est utilisé en entretien pour approfondir ces questions avec le candidat [77]. [78] [79].

3. Perspectives de l'intelligence artificielle

L'avancée de la technologie de l'IA ouvre une nouvelle phase d'extension et de croissance qui renforce la capacité d'exécution et de précision dans les tâches.

L'intelligence artificielle apporte plus de flexibilité avec l'apprentissage automatique et les réseaux de neurones. Même si ces algorithmes ne sont pas tout à fait nouveaux, la magie ne cesse d'opérer au rythme des promesses. De la reconnaissance d'image à la conduite autonome, de la recherche de fraudes à la lutte contre la Covid-19, l'IA ne cesse d'être convoquée pour résoudre des problèmes toujours plus complexes [80] [81] [82].

Discussion

De plus en plus L'IA intègre le quotidien des hommes, ses avantages on en doute plus notamment dans ces secteurs d'applications: on assiste à un déplacement de métier d'humains à machines ; la sécurité routière et la médecine sont au centre de ces avancées.

À cela s'ajoute l'accroissement de production, de richesse des utilisateurs ce qui doit s'accompagner d'une nouvelle politique de formations sur les métiers de ces nouvelles technologies afin de bien répartir ces tâches.

Cependant comme toute technologie, l'intelligence artificielle revêt des retombées négatives par exemple des conséquences sociétales telles que les problèmes d'éthique sont notés.

L'intelligence artificielle cherche à créer des machines capables d'imiter les différents aspects humains comme le raisonnement, la compréhension des langages naturels, la perception visuelle ou auditive, etc.

L'application de l'intelligence artificielle dans l'internet des objets

Bien que l'internet des objets opte pour la "connexion, de son côté l'intelligence artificielle pour rendre "intelligent" mobilise ses algorithmes qui s'entraînent par des données générées par l'IoT. Ces algorithmes se basant sur différents formes d'apprentissages qui seront développés par la suite dans les lignes qui suivent [83].

1. Exemples d'applications d'AIOT

- Industrie : anticiper l'usure d'une pièce ou machine en vue d'une maintenance prédictive ;
- Santé : connaître les effets secondaires d'un traitement ou prescrire le traitement le plus adapté grâce aux informations du patient ;
- Maison intelligente : qui consiste à interconnecter tout un ensemble de matériels pour le bien être des occupantes. Exemple le projet Google Nest ;
- Villes intelligentes : assurer la sécurité et déclencher l'intervention des services adéquats (police, ambulance, pompiers). Pour le trafic routier, on peut identifier les déviations ou les accidents pour éviter les embouteillages ;
- Véhicules connectés : adapter la température du véhicule selon les conditions météorologiques [84] [85] [86].

2. La cognitive dans l'internet des objets

L'internet des objets a déjà prouvé ses compétences dans plusieurs domaines. Néanmoins, ses limites sont à noter. Ainsi l'infiltration de l'intelligence artificielle dans l'internet des objets offre plus de flexibilité et d'autonomie dans ces domaines d'applications.

La science cognitive vise à reproduire au mieux la « pensée humaine » à travers des machines capables de réagir de la sorte. Ces activités mentales seront reproduites selon la compréhension, la perception, la décision du concept à reproduire. Des apprentissages répétitifs testés à travers plusieurs processus permettent à ces machines d'incorporer l'émotif humain en plus de l'automatisation [87].

Devenu réalité le concept d'IOT passe à d'autres étapes qui en est d'autres que plus d'intelligence dans ses applications, l'intelligence artificielle [88].

3. L'IA dans les réseaux de capteurs sans fil (WSN)

Les réseaux de capteurs sans fils sont confrontés à de réels problèmes notamment dans les systèmes de stockage, de transfert de communication en plus d'une source d'alimentation limitée. Dans l'optique de relever ces défis, plusieurs techniques d'intelligence voit le jour : computationnelle Intelligence, Machine Learning, Deep Learning.

L'usage de ces derniers dans des environnements dynamiques et complexes tels que les réseaux de capteurs sans fils pour ajouter une couche intelligente. Celle-ci fournisse plus d'autonomie et de flexibilité aux cas de changements imprévus : topologie, défaillances de communication...

Malgré ces efforts réalisés, les acteurs des deux domaines reste toujours en déphasage : les développeurs n'adopte ou ne sont pas au courant des compétences d'IC, les chercheurs en IC n'exploitent pas au mieux les réelles problèmes dans les RCSF [89].

4. L'IA dans la gestion d'énergie dans l'IOT

Le gaspillage de ressources noté généralement dans les réseaux de capteurs sans fils et dans l'internet des objets pousse aux concepteurs d'ajouter plus de fonctionnalités. Dès lors, on assiste à différentes formes de contribution pour minimiser l'énergie consommé par les capteurs. Une récente contribution dans la gestion de l'énergie AAIoT dans [90], consiste à allouer le calcul d'inférence à chaque couche réseau à chaque appareil du système IOT. La conception d'un algorithme de programmation dynamique pour minimiser le temps de réponse lors de la pesée du cout du calcul et de la transmission. Les résultats montrent que l'approche apporte des améliorations successives dans le temps de réponses du système.

5. L'IA au service de l'exploitation des données générées par l'IoT

Avec les masses de données générés par l'IoT, les outils traditionnels de business intelligents et l'analyse de données ne sont plus à la hauteur de gérer cette quantité. L'ajout de nouvelles

techniques d'apprentissage comme le clustering, permet au Machine Learning(ML) et Deep Learning l'exploitation optimale de ces données. Les alertes sont envoyées dès lors qu'une anomalie est détectée au niveau des data. Le transfert de ces dernières repose sur des paramètres d'énergie, le retard de transmission, la limite des canaux de transmission, la congestion dans le réseau. Pour cela des techniques comme le réseau neuronal et la logique floue sont utilisés dans l'IA.

Dans [91], les auteurs utilisent les techniques de l'IA plus précisément les réseaux neuronaux et la logique floue appliqué au réseau IOT [92].

Conclusion

L'incorporation de l'intelligence artificielle dans l'internet des objets à donner plus de visibilité sur l'application de ce dernier. Cette collaboration fruit d'une « intelligence interconnectée », appliquée dans plusieurs domaines de l'internet des objets dont chaque cas résultant donne naissance à une technologie complémentaire de machines pour des décisions plus humaines tout en réduisant les couts. L'étude des algorithmes basés sur l'IA sera l'objet du prochain chapitre.

PARTIE 2: Etat de l'art de l'apprentissage machine et contribution

Chapitre 3 : L'apprentissage machine, méthodes et algorithmes

L'intelligence artificielle est un domaine de la science à la croisée de plusieurs domaines notamment de l'informatique, des mathématiques appliquées etc. L'apprentissage automatique (AA) s'est avéré l'approche la plus fructueuse et influente dans le développement de l'IA. Différentes techniques d'apprentissage demeurent intéressantes : la classification supervisée et non supervisée.

La classification supervisée : un ensemble d'objets classés servant d'échantillonnage d'apprentissage tout en associant tout objet à sa classe la plus appropriée.

La classification non supervisée : les classes possibles ne sont pas connues à priori et les exemples disponibles sont non étiquetés [93, 94].

Ces types d'apprentissages sont fondés sur des méthodes exploitées par une diversité d'algorithmes appliqués dans plusieurs domaines.

I. Principe de fonctionnement des algorithmes d'apprentissage machine :

En ce qui concerne l'apprentissage automatique, la partie la plus importante du processus de la formation consiste à acquérir suffisamment de données de qualité pour tester le modèle.

De plus, l'entraînement est un processus continu et itératif, fournissant à chaque étape davantage de données ou affinant le modèle [95]. L'ensemble du cycle d'apprentissage automatique peut être résumé comme suit :

- Acquisition des données, cette première étape consiste à obtenir des données pertinentes pour l'application à développer. Ces données doivent être de grande qualité et détaillées
- Préparation des données, cette étape est également appelée nettoyage de données. Les données doivent être précises, propres et sécurisées.
- Sélection de l'algorithme, l'algorithme le plus approprié pour l'application à développer doit être choisi.
- Entraînement du modèle, un modèle d'apprentissage automatique est une représentation mathématique d'un processus réel. L'algorithme retenu doit être entraîné sur les données pour créer le modèle. Le processus d'entraînement peut être supervisé, non supervisé ou renforcé. La description détaillée de ce processus est fournie dans d'autres sections de la présente monographie.
- Évaluation, le modèle doit être évalué pour s'assurer que l'algorithme retenu est le mieux adapté.

1. Les méthodes d'apprentissages supervisés

Une méthode d'apprentissage supervisée peut être comme une fonction de décision D qui prend en compte les paramètres des données x_i appartenant à un ensemble E . Ces données, en entrées x_i , associées à une variable y_i (valeur étiquetée) qui renvoi des valeurs de prédiction $D(x_i)=y_i$.

a. Classification

Ces tâches consistent à attribuer une classe à des objets. Par exemple, dans le milieu bancaire, on peut identifier si une transaction est frauduleuse ou non et de manière automatique, on parle de détection d'anomalie. Dans l'industrie, on peut déterminer si oui ou non une machine est susceptible de tomber en panne. Ainsi on peut associer une réponse prédéfinie (oui ou non, noir ou rouge) à un objet, avant de demander à l'algorithme de réaliser cette classification [96] [97].

✚ Exemples d'algorithmes basés sur la classification

➤ Les forêts aléatoires

L'algorithme de forêt aléatoire commence par un « arbre de décision » (un graphique de type arborescence ou un modèle de décisions) au sommet duquel des données sont entrées. Les données parcourent ensuite l'arborescence pour être segmentées en ensembles de plus en plus petits, en fonction de variables précises [98]. La forêt aléatoire à une faible classification; cependant, lorsqu'il est combiné à un autre algorithme, il produit d'excellents résultats.

➤ Les plus proches voisins

L'algorithme des plus proches voisins évalue la probabilité selon laquelle une donnée appartient à tel ou tel groupe. Il examine des données afin de déterminer le groupe auquel la donnée en question appartient [98] [99].

b. Régression

La méthode de régression n'attribue pas une classe aux données étudiées mais une valeur mathématique qui peut être un pourcentage ou une valeur absolue. Par exemple, une probabilité pour une machine de tomber en panne (5 %, 10 %, etc.) ou le prix de vente idéal d'un appartement en fonction de critères comme la surface, le quartier, type de construction, etc.

✚ Exemples d'algorithmes basés sur la régression :

➤ La régression logistique (apprentissage supervisé – classification)

La régression logistique consiste à déterminer la probabilité qu'un événement se produise en fonction de données fournies. Elle est utilisée pour les variables dépendantes binaires, c'est-à-dire dans lesquelles seules deux valeurs, 0 et 1, représentent les résultats [99] ;

La régression linéaire

La régression linéaire est la régression de base. La régression linéaire simple nous permet de comprendre les relations entre deux variables continues ;

Les arbres de décisions

Un arbre de décisions est une structure arborescente de type organigramme qui illustre tous les résultats possibles d'une décision à l'aide d'une méthode de branchement. Chaque nœud de l'arborescence représente un test lié à une variable spécifique, et chaque branche en est le résultat.

2. Méthodes d'apprentissages non supervisées

La classification automatique peut être définie comme la division des données en groupes ou sous-ensembles d'objets similaires et dissimilaires avec les objets d'autres groupes. On assiste à l'émergence d'assez puissantes méthodes de classification largement applicables dans plusieurs domaines.

Méthodes hiérarchiques

Elles permettent de former des clusters, de construire une hiérarchie ou arbre de clusters connu sous le nom de dendrogramme. Pour cela il existe deux approches basées sur cette méthode.

a. Principe de fonctionnement des méthodes hiérarchiques

Le clustering hiérarchique est initialement formé de plusieurs clusters dans le cas du clustering hiérarchique, de l'agglomératif ou dans le cas divisive, l'architecture dessine un cluster d'ensemble contenant l'ensemble des données. C'est ainsi que le processus de la fusion ou de l'éclatement se base sur plusieurs critères basés sur divers choix qui différencient les techniques de clustering hiérarchique.

Plusieurs critères d'agrégation ont été étudiés et proposés c'est ce qui met en évidence le calcul de distance dans cet ensemble on peut citer : (single link en anglais) définit comme la distance la plus petit séparant un objet du cluster A et un autre du cluster B :

$$D(A, B) = \min (d(x, y)) , x \in A , y \in B \quad (a)$$

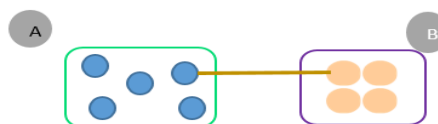


Figure 11 : saut minimal

- Le critère de saut maximale (complete link en anglais), là on va utiliser la distance maximale séparant un objet d'un cluster A à celui d'un cluster B, $D(A,B) = \max(d(x,y))$, $x \in A$, $y \in B$ (b)

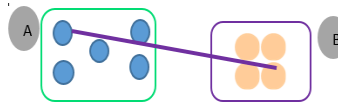


Figure 12 : saut maximale

- Le critère de la moyenne (average Link) pour ce cas, on utilise le calcul moyenne des distances qui incluent les objets du cluster A et B.

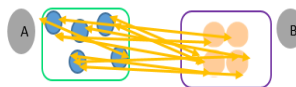


Figure 13 : critère de la moyenne

1. Clustering Hiérarchique ascendante (agglomérative ou bottom up en anglais) :

Elle construit l'arbre du bas vers le haut en démarrant avec autant de cluster que d'objets. Après, on fusionne les clusters considérés comme similaires jusqu'à l'obtention d'un cluster racine.

Dans la suite, nous décrivons quelques exemples d'algorithmes:

BIRCH (Balanced Reducing and iterative clustering using hierarchies)

BIRCH, réduction équilibrée et clustering itératif à l'aide des hiérarchies, est considéré comme le meilleur algorithme pouvant traiter les gros jeux de données. Sa classification est effectuée sur des données compactées puis organisées en une structure d'arbres équilibrés appelé CF tree (clustering Feature), fonction de regroupement arborescente de taille limitée [100].

Il fonctionne suivant trois étapes :

Etape A : l'algorithme scanne linéairement l'ensemble des points et construit l'arbre d'une manière incrémentielle

- Les points sont affectés aux feuilles suivant une similarité définie sans violation du seuil d'absorption SA.
- Si le seuil est atteint une nouvelle entrée dans la feuille est créée

- Si le seuil de E entré dans la feuille est atteint, la feuille est alors décomposée ;

Etape B : l'arbre CF_tree peut être réduit jusqu'à ce qu'un nombre de feuilles soit atteint

Etape C : un algorithme simple de clustering, comme KMeans, peut être appliqué sur les CF des feuilles pour améliorer la qualité des classes obtenues.

➤ **Avantages :**

- Traiter les gros volumes de données ;
- Retrouve rapidement le bon clustering ;
- La structure CF_tree peut être ajuster à l'espace mémoire disponible ;
- Des études ont montré que sa combinaison avec l'algorithmes CLARANS fournit une très bonne qualité de clustering ;

➤ **Inconvénients :**

- L'algorithme considère que des données numériques ;
- Très forte sensibilité à l'ordre des données ;

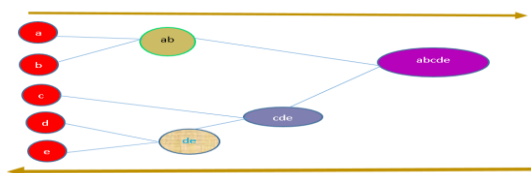


Figure 14 : illustration du clustering hiérarchique

2. Clustering Hiérarchique descendante (divisive ou top down)

Elle construit l'arbre suivant un cluster unique de base contenant tous les éléments de base commençant par le sommet vers le bas. Ensuite, il va diviser successivement les clusters afin d'avoir plus de différence sur les clusters résultants [97, 97].

➤ Exemples d'algorithmes

DIANA (Divisive analysis)

Cet algorithme utilise une méthode de division pour segmenter en plusieurs clusters le cluster initial [100].

➤ **Avantages :**

- Bonne résistance aux données erronées
- Une facilité de traitement des clusters de formes diverses, de similarité ou la distance entre les objets,

Inconvénients :

- Difficulté pour un champ d'étude en évolution, possibilité de fusionner des clusters proches mais distincts s'il y a une chaîne d'objet qui les relie : problème de définition unique des critères.

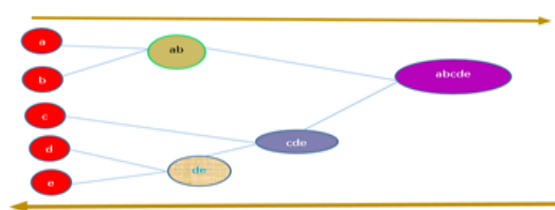


Figure 15 illustration du clustering divisive

Le clustering hiérarchique offre des avantages mais aussi des limites :

Avantages :

- Une facilité de traitement des clusters de formes diverses de similarité ou la distance entre les objets ;
- Applications aux différents types d'objet.

Inconvénients :

- Difficile à appliquer dans les bases de données grandioses ;
- Problème de calcul des distances entre paires d'objets ;
- Le choix du critère n'est pas bien défini, difficulté pour un champ d'étude en évolution ;
- Possibilité de fusionner des clusters proches mais distincts s'il y a une chaîne d'objet qui les relie.

Principe de fonctionnement des méthodes non hiérarchiques

Les méthodes non hiérarchiques offrent plus de flexibilité par rapport aux méthodes hiérarchiques, leurs algorithmes fournissent des résultats qui dépendent du phénomène à étudier.

Les méthodes de Clustering de base

Le Clustering aussi connu sous nom (Segmentation) est un regroupement en classes homogènes consistant à représenter un nuage de points dans un espace quelconque en un ensemble de groupes appelé cluster [97]. L'objectif principal est de classer de façon automatique des données initialement hétérogènes. Il existe plusieurs grandes familles en clustering. Un algorithme de clustering doit suivre l'évolution des clusters, traiter les différents types d'attributs et découvrir les formes de clusters.

Algorithmes basés sur la classification non supervisée (clustering)

Les algorithmes de classification non supervisé existent sous plusieurs formes, nous allons citer quelques-uns. Ensuite nous allons une étude comparative afin de choisir le plus approprié pour notre étude de cas.

1. Méthodes par partitionnement

Dans Les méthodes de partitionnement la formation de cluster se fait à partir d'un ensemble k objet initialiser qui ensuite va itérer afin de réorganiser les clusters pour mieux étaler les dissemblances entre objet [101].

Principe de fonctionnement

Le principe de fonctionnement obéit à ce que chaque cluster contient au moins un objet et que cet objet doit appartenir à un unique cluster. D'abord la formation de partition initiale, ensuite on cherche à améliorer les partitions en attribuant chaque cluster des objets ayant une description plus proche [101] [102].

- ✓ Exemples d'algorithmes basés sur ce modèle

a. KMeans

KMeans, l'un des algorithmes les plus utilisés en partitionnement consiste à diviser l'ensemble des données en un nombre prédéfini de clusters, appelé k [103, 104]. Tous les k points sont ensuite positionnés semi-aléatoirement (1). Ces points, les centroïdes deviennent les centres des nouveaux clusters créés. Chaque élément est par la suite rattaché à l'un de ces nouveaux groupes selon la méthode de distance utilisée (2).

On remplace ensuite le centre de gravité pour qu'il se retrouve au centre des individus de son groupe (3). On répète l'étape 2 et 3 jusqu'à ce que l'algorithme converge et que les centres ne se déplacent plus ou presque.

➤ Avantages

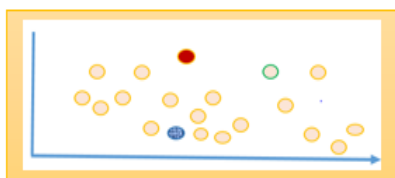
- L'avantage de ces algorithmes est avant tout leur grande simplicité.
- Ils sont appliqués à des données de grandes tailles

➤ Inconvénients

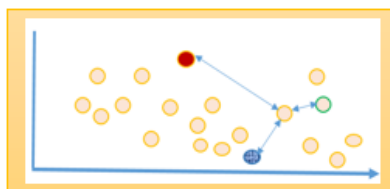
- Le nombre de classe doit être fixé au départ
- Ne détecte pas les données bruitées
- Le résultat dépend de tirage initial des centres des classes.

✓ Structure de KMeans

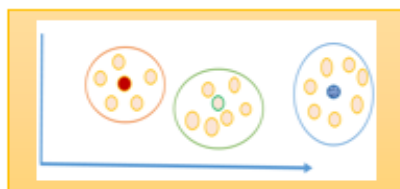
- A l'initialisation : on tire de façon aléatoire les clusters centre ; ici, on a trois centres de classes



Etape 1 : on calcul la distance entre un cluster à un centre



Etape 2 : affectation des objets à un centre le plus proche



Etape 3 : on calcul les nouveaux centres de gravité, on itère encore jusqu' à ce qu'on parcourt tous les objets en leurs affectant à leur centre.

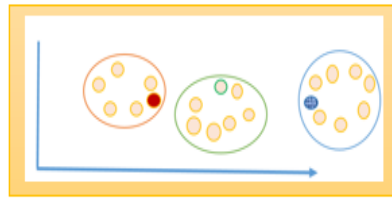


Figure15 : principe de KMeans

b. K-medoids

À l'opposé de KMeans cluster, représenté par une valeur moyenne nommée centroïde, K-medoids lui utilise les objets les plus dominants appelés medoids.

Cette structure offre deux avantages, l'adaptation à n'importe quel type de données, elles ne sont pas sensibles aux objets isolés [103], [104].

Il fonctionne selon le principe suivant :

- 1) Choix des centres initiaux puis affectation de chaque objet à un cluster.
- 2) Recalcul des medoids pour chaque cluster et redistribution des objets selon les nouveaux centres.
- 3) Processus est répété jusqu'à la stabilisation des clusters.

➤ Avantages :

- Bonne résistance aux données erronées.
- Flexibilité avec tout type de distance.

➤ Inconvénients :

Nécessité de spécifier le nombre de clusters k, Figure 16

- Complexité de chaque itération.

✓ Structure de k-medoids :

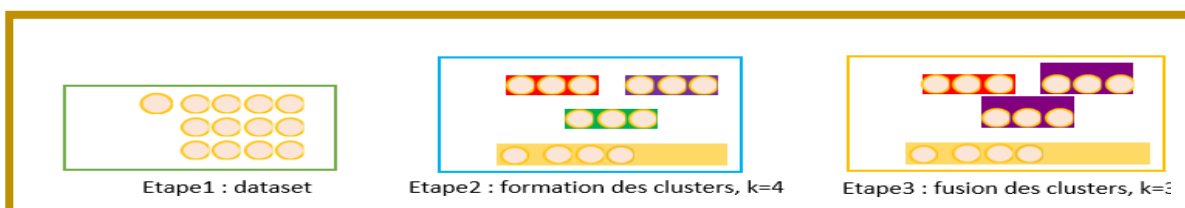


Figure 17 : Principe de fonctionnement de K-medoids

Figure 16 : Principe de fonctionnement de K-medoids

❖ Discussion

Les algorithmes basés sur le partitionnement des données offrent des avantages et inconvénients

Avantages :

- Facilité d'implémentation,
- Application sur des bases de données volumineuse

Inconvénients :

- Les conditions d'arrêt de l'exécution de l'algorithme doivent être fixées par l'utilisateur
- Les paramètres ne sont pas définis de manière fixe mais choisis aléatoirement.
- Difficultés pour l'interprétation et la visualisation des clusters.

2. Les méthodes à base de densité

Les méthodes hiérarchiques basées sur la densité s'appliquent aux zones de densités relativement élevées. C'est-à-dire les zones où beaucoup de points sont proches par rapport à d'autres zones pour former des classes d'objets [105, 97]. En outre, la méthode permet de repérer des objets en dehors de la zone dite bruits ou valeurs aberrantes. La méthode à densité prend deux paramètres en entrées, la distance maximale **D** définissant le voisinage entre deux objets et **P** (le nombre minimal d'objets nécessaire pour former un cluster ou groupe [98]).

Principe de fonctionnement

On aura besoin de stocker deux informations : les clusters successifs et les individus visités au fur et à mesure. La première étape consiste à choisir un point parmi les n disponibles. Grâce au paramètre ϵ , on peut définir le voisinage du point initial, c'est-à-dire l'ensemble des points que l'on peut qualifier comme ses voisins.

Et grâce au paramètre N , on peut dire que si cet ensemble est constitué de moins de N points, alors le point initial correspond à du bruit. On le stocke alors dans les individus visités.

A l'inverse si le voisinage comprend plus de N points, alors on peut initialiser un cluster avec le point de départ. On étudie chaque point de son voisinage initial.

Et pour tout point de ce voisinage, si son propre voisinage comporte plus de N éléments, alors on étend le voisinage initial en le réunissant avec le voisinage du point visité.

Puis on ajoute ce point dans le cluster. Une fois que tous les points du voisinage ont été testés, ceux retenus dans le cluster sont stockés comme individus visités. Et le cluster obtenu est stocké dans la liste des clusters.

Tant que tous les individus n'ont pas été visités, on réitère cette étape en commençant par choisir un individu parmi ceux qui sont encore *disponibles*. Et on obtient finalement notre liste de groupes d'individus ainsi que les individus correspondant à du bruit [106].

✓ Exemples d'algorithmes basés sur ce modèle

a. DBSCAN

L'algorithme Density-based spatial clustering of applications with noise (DBSCAN) est un algorithme plus récent (Ester et al. 1996) basé sur une estimation locale de la densité comme son acronyme le désigne [103]. Il se base sur deux paramètres (nombre minimum de points et rayon d'une boule, il regroupe itérativement les points par paquet sur la base de leur voisinage (nombre minimum d'individus) à l'intérieur d'une boule de rayon [105].

Il est capable de trouver des clusters de forme arbitraire et des clusters avec du bruit (c'est-à-dire des valeurs aberrantes).

L'idée principale derrière DBSCAN est qu'un point appartient à un cluster s'il est proche de nombreux points de ce cluster.

Il existe deux paramètres clés de DBSCAN :

- **Dist** : la distance qui spécifie les quartiers. Deux points sont considérés comme voisins si la distance entre eux est inférieure ou égale à *dist*.
- **MinPts** : nombre minimum de points de données pour définir un cluster.
- **Point central** : Un point est dit central s'il y a au moins *minPts* nombre de points (y compris le point lui-même) dans sa zone environnante avec un rayon *dist*
- **Point frontière** : Un point est un point frontière s'il est accessible à partir d'un point central et qu'il y a moins de *minPts* nombre de points dans sa zone environnante.

- **Valeur aberrante** : un point est une valeur aberrante s'il ne s'agit pas d'un point central et qu'il n'est accessible à partir d'aucun point central.
- MinPts et dist sont déterminés.

Les minPts sont les points voisinant du point central et non marqués comme bruit. Sinon, le point est marqué comme du bruit. Une fois qu'une formation de cluster commence, tous les points dans le voisinage du point initial deviennent une partie du cluster. Si ces nouveaux points sont également des points centraux, les points qui se trouvent à proximité sont également ajoutés à ce cluster [98].

➤ **Avantages :**

- Retrouver des classes de formes arbitraires
- N'est pas sensible à l'ordre d'arriver des objets

➤ **Inconvénients :**

- Complexité pour la définition des paramètres de base, le rayon, voisinage
- N'est pas adapté aux données de dimension très large

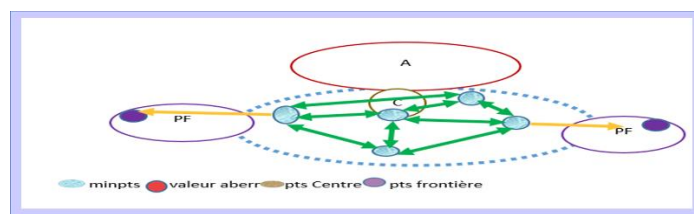


Figure 18 : structure de DBSCAN

b. DENCLUE

Centré sur des maximums locaux de fonctions de densité, l'algorithme DENCLUE utilise une série de techniques descendantes et ascendantes de gradients pour les trouver [105]. En plus de ces clusters définis par rapport au centre, les clusters de formes arbitraires sont définis comme une continuation le long des séquences de points dont les densités locales ne sont pas plus petites que le seuil S prévu. L'algorithme est stable par rapport aux « bruits » [103].

DENCLUE présente un bon passage à l'échelle, puisqu'à sa phase initiale, il construit une carte de cubes hyper rectangles d'une longueur d'arrête $2S$.

Pour cette raison, l'algorithme peut être classé comme une méthode à base de grille. L'explication est que quoique tous les points sont explorés, la majeure partie de l'analyse (dans la phase de clustering) concerne seulement les points des zones densément peuplées. Son fonctionnement obéit deux principales étapes [98] :

Prétraitement des données, découpage de l'espace des objets en cellule de grille

La phase de classement, l'algorithme ne considère que les cellules denses et calcule les fonctions d'influences, ensuite il fusionne les cellules qui peuvent être jointes selon un critère prédéfini [100].

➤ **Avantages :**

- Trouver des classes aux formes arbitraires sans fixer le nombre de classes recherché
- Résiste au bruit

➤ **Inconvénients :**

- Nécessité d'un choix adapter au paramètres seuil **S** et

3. Les méthodes basées sur un modèle

Les méthodes basées sur un modèle sont l'une des méthodes phares pour réaliser le partitionnement des données. Cela permet d'obtenir un point de vue conceptuel du cluster selon un certain modèle dont les paramètres inconnus doivent être estimés. Il s'agit de l'approche, le modèle statistique ou probabiliste et le modèle neuronal. Les algorithmes qui se basent sur ce modèle sont nombreux [105].

a. L'algorithme Self-Organizing Maps (SOM)

L'algorithme des cartes auto-organisatrices de Kohonen, en anglais (SOM), est un algorithme de type non supervisé basé sur les réseaux de neurones artificiels [103].

Les neurones suivent un réseau fixe de l'entrée à la sortie, durant tout le temps d'apprentissage de la couche [97, 100] .

b. Fonctionnement

Chaque neurone possède des coordonnées dans l'espace et représente une classe de données deux neurones reliées s'influence l'un de l'autre au cours de l'apprentissage. L'algorithme

SOM, méthode de quantification vectorielle qui permet de découvrir la topologie des données. Il regroupe les objets dans leurs classes respectives tout en tenant compte de la position dans l'espace d'étude.

✓ **Avantages**

- Rapidité et efficacité pour le traitement des gros dataset
- Pas d'utilisation de modèles complexes mathématiques

✓ **Inconvénients**

- Lenteur d'apprentissage
- Résultat non interprétable

4. Méthodes basées sur association

Elle consiste à analyser des relations entre variables, il crée un lien d'association entre eux cette méthode va fusionner les objets susceptibles d'être ensemble [106].

a. Apriori

L'algorithme Apriori s'utilise dans une base de données transactionnelle pour extraire des ensembles d'éléments fréquents, puis générer des règles d'association. Il est couramment utilisé dans l'analyse des paniers de marché, où l'on vérifie les combinaisons de produits qui apparaissent fréquemment dans la base de données [107]. En général, nous écrivons la règle d'association pour «si une personne achète l'élément X, elle achète ensuite l'élément Y» comme suit : $X \rightarrow Y$ [108].

Exemple : si une personne achète du lait et du sucre, elle achètera probablement de la poudre de café. Cela pourrait être écrit sous la forme d'une règle d'association comme suit : {lait, sucre} \rightarrow café en poudre. Les règles d'association sont générées après le franchissement du seuil de support et de confiance [108].

La mesure de support permet d'éliminer le nombre d'ensembles d'éléments à prendre en compte lors de la génération fréquente d'ensembles [109].

Le principe Apriori stipule que si un ensemble d'éléments est fréquent, tous ses sous-ensembles doivent également l'être [108].

➤ **Avantages :**

- La possibilité de découverte des connaissances utiles, cachées dans les bases de données

- Leurs facilités de compréhension, efficacité et simplicité.
- Le forage des règles d'association est un grand succès dans divers domaines que ce soit dans des activités commerciales, sociales ou humaines [108].
- Inconvénients :
 - La découverte d'un nombre important de règles d'association dont la plupart ne sont pas intéressantes
 - Le temps de recherche des itemsets fréquents est énorme
 - Les algorithmes utilisés ont trop de paramètres, par conséquent l'extraction de données, pour les non experts, devient compliquée.
 - Un problème de sécurité pourrait être posé : des renseignements confidentiels peuvent être facilement divulgués, en utilisant cette technique [108].

b. Illustration de l'algorithme Apriori

Considérons les items fréquents et les numéros de transaction du tableau ci-dessous. Ce tableau montre une forte relation de l'itemsets (pain, sucre, lait) dans l'ensemble **I** des items et **NT** les numéros de transaction enregistré dans la base de donnée, par conséquent ce qui ont acheté du sucre et du lait ont généralement acheté du pain [108, 106].

Le calcul de confiance montre :

$$C(\text{sucre, lait} \rightarrow \text{pain}) = \text{Card}(\text{sucre, lait}) / \text{card}(\text{sucre, lait, pain})$$

$$(2/5) / (3/5) = 0.66 \rightarrow \text{pourcentage de confiance, } C=66\%$$

Tableau 1 : illustration de l'association

Numéros de transaction	Items fréquents
1	Beurre, Pain, sucre, lait,
2	Fruit, boisson, pain, sucre, lait, pain
3	Sucre, lait, café, eau minérale
4	Légumes, Riz, Pain, sucre, lait, huile
5	Sucre, chocolat, riz, huile, bouillon

❖ Discussion

- Avantages :

- Leur application dans plusieurs domaines de la vie quotidienne, comme l'analyse du panier de la ménagère
- La découverte de connaissances utiles, cachées dans les grandes bases des données
- Leur simplicité, efficacité et facilité de compréhension
- Leur formalisme non supervisé et général
- Leurs résultats clairs et faciles à interpréter.
 - Inconvénients :
 - Le temps énorme consacré à la recherche des ItemSets fréquents.
 - La grande quantité des règles d'association générées
 - La difficulté d'évaluer la qualité des règles d'associations par des indices statiques ou par l'expert du domaine.

5. Méthodes à base de grille

L'idée de ces méthodes est qu'on divise l'espace de données en un nombre fini de cellules formant une grille.

Ce type d'algorithme est conçu, pour des données spatiales. Une cellule peut être un cube, une région, un hyper rectangle.

En fait, avec une telle représentation des données, au lieu de faire la classification dans l'espace de données, on la fait dans l'espace spatial en utilisant des informations statistiques des points dans la cellule [97].

Les méthodes de ce type sont hiérarchiques ou de partitionnement. Les algorithmes les plus connus sont STING, CLIQUE, WaveCluster.

- Exemples d'algorithmes basé sur ces modèles :

a. STING

L'algorithme STING (STatistical INformation Grid) utilise ainsi une grille multi résolution. Les données dans chaque cellule sont résumées par le nombre d'objets dans la grille et par leur moyenne.

Chaque cellule est récursivement découpée en 4 sous-cellules. On commence en considérant l'espace entier comme une cellule ancêtre de la hiérarchie et on s'arrête si l'on atteint un critère de terminaison : (le nombre des points dans une cellule est inférieur à un seuil) [105].

b. WaveCluster

C'est un algorithme qui est conçu, pour travailler sur des données spatiales. Dans l'algorithme, on considère les données spatiales comme les signaux multidimensionnels et on applique une technique de traitement de signal qui s'appelle la transformation wavelet pour transformer l'espace de données dans un espace de fréquences [97]. L'idée est basée sur le fait que les parties de haute fréquence du signal correspondent aux régions de l'espace spatial de données où il y a un changement brusque dans la distribution des objets [105].

Discussion

➤ **Avantages :**

- Indépendance des données par accumulation via grille
- Fonctionne avec des variables de types divers.

➤ **Inconvénients :** problème de gestion des grilles de grande taille

Les réseaux de neurones artificiels

L'intelligence artificielle, cette technologie qui se développe à grande vitesse, et ses cas d'usage se multiplient dans tous les secteurs. Selon les experts, l'IA est amenée à bouleverser tous les aspects de notre société dans les années à venir.

Dans ce contexte, il semble important de comprendre comment fonctionne réellement l'intelligence artificielle. Dou l'intérêt d'en apprendre davantage sur les réseaux de neurones artificiels sur lesquels reposent de nombreuses IA.

Définition

Un réseau de neurones artificiels, ou Artificiel Neural Network en anglais, est un système informatique matériel et / ou logiciel dont le fonctionnement est calqué sur celui des neurones du cerveau humain [110].

1. Historique et évolution

- 1890 : W. James, célèbre psychologue américain introduit le concept de mémoire associative, et propose ce qui deviendra une loi de fonctionnement pour l'apprentissage sur les réseaux de neurones connue plus tard sous le nom de loi de Hebb.
- 1943 : J. Mc Culloch et W. Pitts laissent leurs noms à une modélisation du neurone biologique (un neurone au comportement binaire).
- 1949 : D. Hebb, physiologiste américain explique le conditionnement chez l'animal par les propriétés des neurones eux-mêmes. Ainsi, un conditionnement de type pavlovien tel que, nourrir tous les jours à la même heure un chien, entraîne chez cet animal la sécrétion de salive.
- Théorie du fonctionnement et des possibilités des réseaux de neurones par J. J. Hopfield [111].

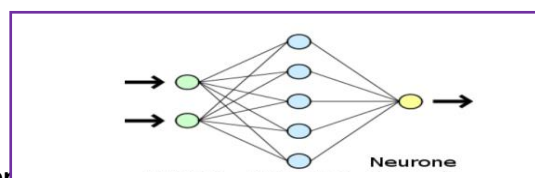
2. Types de réseaux de neurones

a. Les réseaux de neurones feed-forwarded (traitement de la donnée)

C'est la procédure du traitement de la donnée par le réseau neuronal [106]. En effet, feed-forwarded (propagation avant) signifie tout simplement que la donnée traverse le réseau d'entrée à la sortie sans retour en arrière de l'information ; on distingue les réseaux monocouches (perceptron simple) et les réseaux multicouches (perceptron multicouche) [112] [113].

- **Perception simple**

Le traitement de la donnée dans ce réseau se fait entre la couche d'entrée et la couche de sortie qui sont toutes reliées entre elles. Le réseau intégral ne dispose ainsi que d'une matrice de poids. Le fait de disposer d'une seule matrice de poids limite le perceptron simple à un classificateur linéaire permettant de diviser l'ensemble d'informations obtenues en deux catégories distinguées [114].



- **Perceptron multicouche**

Le perceptron multicouche se structure de la même façon. L'information entre par une couche

Figure 19 : perception simple

ontron

multicouche dispose entre la couche en entrée et la couche en sortie une ou plusieurs couches dites « cachées ».

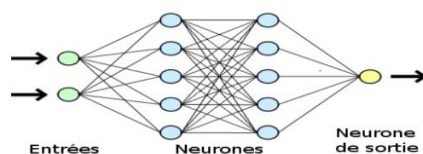


Figure 20 : structure de la perception multicouche

b. Les réseaux de neurones récurrents

Les réseaux de neurones récurrents traitent l'information en cycle. Ces cycles permettent au réseau de traiter l'information plusieurs fois en la renvoyant à chaque fois au sein du réseau.

La force de ces réseaux de neurones récurrents réside dans leur capacité de prendre en compte des informations contextuelles, suite à la récurrence du traitement de la même information ; c'est cette dynamique qui auto-entretient le réseau. Les Réseaux de neurones récurrents se composent d'une ou plusieurs couches. Le modèle de *Hopfield* (réseau temporel) est le réseau de neurones récurrent d'une seule couche le plus connu [115].

c. Les réseaux de neurones à résonance

L'appellation du réseau à résonance fait encore une fois référence à son fonctionnement [106]. En effet, au sein des réseaux de neurones à résonance, l'activation de tout neurone est renvoyée à tous les autres neurones au sein du système. Ce renvoi provoque des oscillations, d'où la raison du terme résonance [116, 117].

d. Les réseaux de neurones auto-organisés

Les Réseaux de neurones auto-organisés sont surtout adaptés pour le traitement de d'informations spatiales. Par des méthodes d'apprentissage non-supervisé, les réseaux neuronaux auto-organisés sont capables d'étudier la répartition de données dans des grands espaces comme par exemple pour des problématiques de clusterisation ou de classifications [116, 117].

e. Les réseaux de neurones convolutifs

Les réseaux de neurones convolutifs sont directement inspirés du cortex visuel des vertébrés. Un réseau de neurones à convolution, appelé aussi convnet (pour « Convolutional Network »), ou encore CNN (pour « Convolutional Neural Network ») [118].

Les CNN Les réseaux de neurones convolutifs ont été créés dans les années 80. L'objectif de cette création était d'analyser une partie d'une image quelconque sans se soucier de la position de l'image [106].

Un CNN est formé de deux parties, une première la partie convolutive du modèle (convolution) et une deuxième classification correspondant aux perceptron multiple (MPL) [114, 117].

f. Les réseaux de neurones Capsules

Ce sont de nouveaux types de réseaux de neurones profond traitant l'information de façon visuelle similaire à l'analyse du cerveau. C'est une alternative améliorant les réseaux de neurones convolutifs [119, 120]

Tableau 2 : Etude Comparative des méthodes d'apprentissages

Classes	Méthodes	avantages	Limites	Domaines d'application
Clustering de base	Hiérarchique	Traitement de différentes formes de similarité	elles sont coûteuses en terme, de temps d'exécution et d'espace mémoire	Médecine Industrie Reconnaissance des formes Marketing, Sécurité
Clustering de base	Densité	Nécessite pas une définition préalable des clusters, les algorithmes sont incrémentaux, peut-être, constitués, d'objets moins similaire	La performance peut diminuer en fonction du nombre d'objets algorithmes se base sur les paramètres de base sensibles au résultat obtenu	Médecine Archéologie Reconnaissance des formes Marketing, Sécurité, Industrie
Clustering de base	Partitionnement	Facile à implémenter Faible exigence en taille Application au gros dataset	l'utilisateur doit fixer le choix des paramètres,aléatoire,visualisation et interprétation des clusters	Médecine, archéologie, Reconnaissance des formes Marketing, sécurité, industrie

Clustering de base	Modèles	Réduction de l'espace multidimensionnelle des données d'entrées l'analyse est insensible aux valeurs manquantes	Condition d'initialisation du modèles, le temps consacré pour la construction d'un réseau.	Médecine Archéologie, industrie Reconnaissance des formes Marketing, sécurité
Clustering de base	Grille	Indépendance des données par accumulation via grille Fonctionne avec des variables de types divers	Problème de gestion des grille de grande taille	Médecine Archéologie Reconnaissance des formes Marketing, sécurité, industrie
Association	Association	Reconnaissance d'objets disponible, création de lien entre les objets	Rupture d'objets associés dans la catégorie une mauvaise association pour un choix mal veillant	E-commerce Paniers ménagères
	Classification	Simplicité d Facilité de calcul Bonne compréhension du problème étudié	Problème avec les types de données	Industrie Banque commerce
	Régression	Insensibilité aux types de données sélection des variables les plus informatives pour l'étude	Sensible aux bruits incompatibilité avec les données incrémentés	Banque Commerce Habitat

Tableau 3 : Étude comparative des algorithmes d'apprentissages machine

Algorithmes	Types d'apprentissages	méthodes	Avantages	Inconvénients
Régression linéaire	Supervisé	régression	Simplicité d'interprétation, facilité de calcul	Très sensible aux valeurs aberrantes
Régression logistique	Supervisé	classification	Des événements passent comme prévu	Echec des événements
Forêts aléatoires	Supervisé	Classification	Meilleure sélection de l'arbre	Erreur sur l'arbre choisi
Arbres de décision	Supervisé	Régression	prise en compte simultanée de variables qualitatives et quantitatives, détermine des seuils discriminants pour ces dernières, sélection des variables les plus informatives, peu d'influence des données erronées, très rapides classification, meilleure compréhension du problème étudié.	Traitement des variables numériques, ne tient pas compte des propriétés de densité, Algorithmes séquentiels sans remise en cause des étapes précédentes, un peu assoupli dans la production de systèmes de règles, instabilité: sensibles aux variations

Les plus proches voisins	Supervisé	Classification	Simple et facile à mettre en œuvre. Aucune hypothèse sur les données (linéaires, affines,), polyvalent	Devient plus difficile le nombre d'exemples augmente. Le choix de la méthode de calcul, complexité
DBSCAN	Non supervisé	Clustering basé sur la densité connecté	Compatible avec tout type de cluster, détection de valeurs aberrantes	Difficulté pour détecter certains cluster selon leurs distances ,pas compatible avec des cluster de grande densité
KMeans clustering	Non supervisé	Clustering basé sur la partition	Simple, rapide et facile à comprendre faible cout de calcul adaptable au gros data set	Problèmes d'atteindre le résultat souhaité, difficiles de retrouver les groupes ayant formes complexes
DENCLUE	Non supervisé	Densité, basé fonction densité	O Trouver des classes au formes arbitraire sans fixer le nombre de classes recherché O Résiste au bruit	Nécessité d'un choix adapter au paramètres Pas compatible avec les clusters flous
BIRCH	Non supervisé	Hiérarchique ascendante	O Traiter les gros volumes de données O Retrouve rapidement le bon clustering La structure CF_tree peut être ajuster à l'espace mémoire disponible	L'algorithme considère que les données numériques, très forte sensibilité à l'ordre des données.

DIANA	Non supervisé	Hiérarchique descendante	Construction de cluster constitués de données plus proche	Problème de de définition de critère de séparation d'un cluster
K-Medoids	Non supervisé	Partitionnement	Bonne résistance aux données erronées, Flexibles avec tout type de distance	Nécessité de spécifier le nombre de clusters k, Complexité de chaque itération.
SOM	Non supervisé	Modèles	Rapidité et efficacité pour le traitement des gros dataset Pas d'utilisation de modèles complexes mathématiques Taux d'erreurs très faible par rapport aux autres méthodes	Lenteur d'apprentissage Résultat non interprétable Pas d'incrémentation des données lors de l'apprentissage
A Priori	Non supervisé	association	La possibilité de découverte des connaissances utiles, cachées dans les bases de données Le forage des règles d'association est un grand succès dans divers domaines que ce soit dans des activités commerciales, sociales ou humaines	La découverte d'un nombre important de règles d'association Le temps de recherche des Itemsets fréquents Les algorithmes utilisés ont trop de paramètres, compliquée. Un problème de sécurité pourrait être posé
STING	Non supervisé	grille	Traitement de grandes base de données, découpage de l'espace des données	Choix des paramètres liés à la taille des grandes cellules

WaveCluster	Non supervisé	grille	Compatibilité avec les données multidimensionnelle numérique	Incompatible avec des données bidimensionnelles
			Découverte de clusters de formes variés	Complexité des données bidimensionnelle suivant la dimension des objets
			Détection de bruit	
			Gestion de la bonne qualité des classes	

Conclusion

L'étude des différentes méthodes d'apprentissage supervisé ou non offre une panoplie de choix suivant divers critères. Des techniques d'associations, de clustering de base et d'ensemble sont exploités par plusieurs algorithmes appliqués à divers domaines d'applications. Les principaux points du prochain chapitre seront autour du choix de l'algorithme qui sera appliqué à l'internet des objets, cette fusion sera appliqué dans le domaine du transport routier.

Chapitre 4: L'Apprentissage non Supervisé dans l'Internet des Objets

Depuis la miniaturisation des appareils de l'internet des objets, les captures des données deviennent en permanence de plus en plus nombreuses, ainsi l'agrégation, l'analyse et le traitement de ces dernières restent des défis pour l'écosystème IoT. En effet l'envoi et la réception de données dans les réseaux IoT connaît un retard de transmission dite congestion.

L'intelligence artificielle vient en application pour relever ces défis. Une combinaison de ces deux systèmes a fait l'objet de plusieurs recherches. L'apprentissage machine issu de l'IA se fait par des techniques d'algorithmes d'apprentissage, par renforcement, supervisé et non supervisé. L'IA peut être appliquée afin de réduire le temps de transfert des données.

I. Différents problèmes dans l'IOT

L'internet des objets est appliqué dans plusieurs domaines qui nécessite un assemblage de moyens électroniques et logiciels. Ce système s'articule autour de plusieurs chaines dont chaque maillon correspond à un niveau : niveau de capture, niveau de transmission, niveau de traitement etc....

Notons que les applications sont confrontées à des limites qui restent les défis majeurs des concepteurs, développeurs et chercheurs. Ainsi plusieurs contributions se font suivant les niveaux matériels et logiciels. Ces derniers sont les points centraux considérés pour apporter de nouveaux défis pour des applications au meilleure usage.

1. Routage dans l'internet des objets

Le routage informatique est un processus de sélection des chemins qui vont transmettre les informations aux destinataires dans un réseau. Le calcul de transfert de ces données entre les appareils de communication se base sur des bandes passantes appropriées, quel que soit la taille des données leurs transmissions reste prioritaires [121].

2. Le problème de Normalisation dans l'internet des objets

Une architecture normalisée n'a pas été définie par l'internet des objets, ce qui occasionne différents types architectures spécifiques définies selon le domaine d'exercice appliqué par des entreprises. Il existe plusieurs plateformes pour la gestion des données : horizontales d'usage multiple, verticales d'applications spécifiques et améliorées qui peuvent parvenir à des fins optimales.

3. La gestion de l'énergie dans l'internet des objets

L'architecture IoT est constituée de capteurs avec une durée vie limitées en énergie puisque le système IoT est très énergivore, il urge de minimiser cette énergie. L'IoT à lui seul n'est pas capable de relever ce défis, d'où l'application de l'intelligence artificielle pour accélérer les processus, minimiser le temps de réponse : ce peut entrainer un économie l'énergie [121].

La qualité de service dans l'internet des objets (QoS)

La qualité de service est la capacité à véhiculer dans de bonnes conditions un type de trafic donné, en termes de disponibilité, de débit, de délais de transmission, de taux de pertes de

paquets, de manque d'inférence dans plusieurs appareils de façon simultanés [121]. Pour pallier à ces problèmes, l'intelligence artificielle peut calculer le temps de réponse équivalent au temps de calcul et de transmission puis alloue plus de bandes passantes aux hautes couches ayant plus de données.

1. L'architecture dans l'internet des objets

L'architecture de réseau est l'organisation d'équipements de transmission, de logiciels, de protocoles de communication et d'infrastructure filaire ou sans fil, de radioélectrique permettant la transmission des données entre les différents composants.

L'architecture définie dans l'internet des objets est multicouche avec plusieurs niveaux. Pour des traitements limpides, les données de captures ont besoin d'agrégation et d'analyses permanentes pour la gestion des flux. Ce défi majeur de l'internet des objets va réduire le flou noté dans l'extraction de données pour une prise de bonne décision [122].

2. La congestion dans l'internet des objets

La congestion correspond à l'état de surcharge, le réseau échange et transporte plus de données qu'il ne peut traiter.

De nombreuses données sont issues de capteurs IoT, ainsi l'envoi et la réception de ce volume de données reposent sur des systèmes qualifiés qui en plus d'être des voies de communication doivent définir l'accès au canal afin de maintenir une transmission qualifiée dans un temps requis. Les transmissions se font via des réseaux de données mais doivent l'être aussi en fonction de paramètres afin d'éviter des données redondantes ou inutiles [122] [123]. Dans la section suivante, nous nous focalisons sur les mécanismes de gestion de la congestion.

Traitement de la congestion dans l'Internet des objets

La congestion reste un problème majeur dans l'internet des objets plusieurs solutions ont été proposées mais tardent à fournir de meilleurs résultats du point de vue architecture, transmission et qualité de service. Ainsi l'intelligence artificielle vient en appoint via des technologies d'apprentissage basées sur des algorithmes pour apporter plus de flexibilité dans l'internet des objets. Nous distinguons ainsi plusieurs méthodes avec des approches différentes.

1. La méthode basée sur l'architecture

Des architectures sont proposées pour mieux accélérer le circuit de transmission, l'intelligence artificielle à travers des données propose des algorithmes de calculs de plus court chemins pour un acheminement rapide des données vers leur destinations [124]. La Figure 20 présente un système de traitement de congestion basé sur l'architecture.

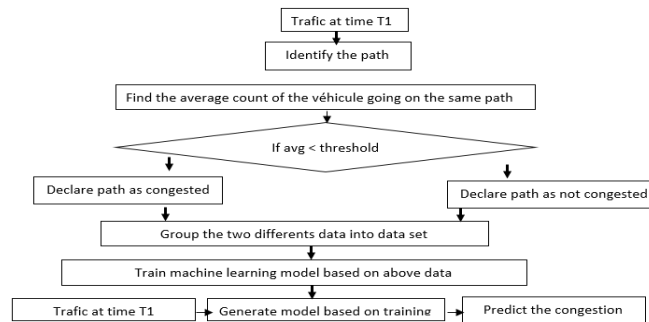


Figure 21: Architecture de chemin plus court pour contrer la congestion

2. Méthode basée sur le routage

Pour résoudre le problème d'acheminement des données dans les réseaux IoT, des protocoles de routages sont définis. L'arrivée des réseaux de transmission définis par logiciels qui, via des calculs de poids maximal résout les poids de routage maximal offrant de meilleurs performances en termes de fluidité d'accès au canal pour restreindre le niveau de congestion et de dépense énergétique dans [125].

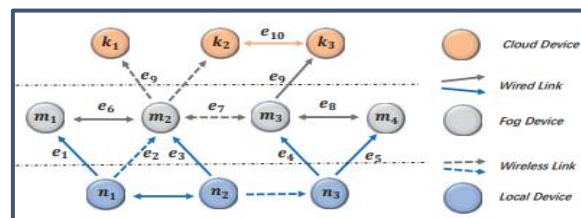


Figure 22: routage maximal par poids maximal

3. Méthode basée sur les algorithmes d'apprentissage non supervisé

La méthode basée sur les algorithmes non supervisés, collecte des données statiques et dynamiques pour gérer la congestion et définit des instants de transmission. L'échange de données entre capteurs favorise la subdivision des éléments similaires. Le traitement des paquets à des intervalles de temps définis pour la génération et la transmission de ces paquets de données qui obéit à des topologies variées.

Dans le tableau 4, nous faisons une comparaison de quelques algorithmes d'apprentissage non supervisés.

Tableau 4: les algorithmes d'apprentissages non supervisé et leurs domaines d'applications

Articles	Solution proposée	Approche	topologie	Algorithmes	paramètres	domaines
Modelling-IOT [124]	Prédire les conditions de circulation avant voyage	Architecture client/serveur	étoile	Apprentissage automatique	Temps, moyenne, seuil	transport
OSUWA [122]	Contrôle de la congestion	Architecture client/serveur	étoile	Réseau de neurone	temps	industrie
WANG [126]	Gestion d'une congestion événementielle	Architecture poste à poste	hiérarchique	Dbscan/K-medoids /AP/Clustering agglomératif	Temps, volume de véhicule, le taux	transport
ZHANG [125]	Résoudre les problèmes de congestion	routage	hiérarchique	Mwmp(Propulsion par vagues à poids maximal)	Paquets, bp,délais	santé

Pour améliorer la congestion à travers un algorithme, une étude comparative est réalisée entre KMeans et d'autres algorithmes d'apprentissages non supervisés. Cette étude se base sur les performances apportées par KMeans seul et celles obtenues en combinant KMeans avec d'autres algorithmes. Le tableau ci-dessous illustre les résultats obtenus.

Tableau 5:Kmeans et les autres algorithmes d'apprentissage pour minimiser la congestion dans un réseau routier

Articles sur la congestion	approches	Paramètres	algorithmes	avantages	inconvénients
Système de reconnaissance des schéma critique [127]	Hiérarchique	Densité, vitesse moyenne	KMeans	Partition des segments critiques	N'apporte pas de solution concrètes

Identification des créneaux horaire critique [128]	Hiérarchique	Densité, vitesse, temps données météorologiques	KMeans /arbres de décision	Partitions des données anormales	Structure des données
Protocole hybride d'acheminement Vannets [129]	Routage multi-niveaux	Densité, vitesse, distance	KMeansModifié /Hopfield	meilleure route, partition, délais livraison	Plusieurs paramètres à fusionner
Détection de la congestion du trafic dans une ville à l'aide des techniques de clustering dans les VANETs [130]	Hiérarchiques	Distance, délais,	K-means, c-means fuzzy, k-means fuzzy	Calcul de distance, délais, Partition	Mauvais structure des données

Conclusion

L'apprentissage non supervisé offre beaucoup de flexibilités dans la science des données. Les algorithmes basés sur ce fonctionnement aident l'intelligence artificielle à participer dans l'évolution de l'architecture de l'internet des objets.

Il est alors possible de gérer la congestion à travers des méthodes algorithmiques basées sur les réseaux de neurones exploitent au mieux les données. Nous avons essayé dans ce chapitre de faire une étude comparative en se basant sur un certain nombre de paramètres.

Au chapitre suivant, nous allons passer à l'implémentation de KMeans dans une architecture à cinq niveaux.

Chapitre 5: Implémentation de KMeans dans une Architecture à Cinq Niveaux

Le XXI^e siècle est marqué par un rythme de croissance de la population urbaine qui augmente de façon considérable, dix fois que la population mondiale. Le développement anarchique des villes les a rendues incapables de faire face à l'augmentation du nombre des véhicules. Avec une prolifération de ces derniers, il est difficile de circuler sans embouteillage, ainsi les congestions deviennent inévitables.

Dès lors on assiste à une spécialisation fonctionnelle de site de logement et d'emploi-service reliés grâce à des systèmes de transports ferroviaires, routier et tant d'autres... Aujourd'hui, dans certains pays comme le Sénégal(Dakar) [1], France(Paris), Brésil(Sao Polo) [2], Chine(Pékin), Inde(Mumbai) [2] des travaux de recherche basés sur l'IoT se développent afin d'apporter des réponses face au phénomène de congestion dans la circulation urbaine.

Ce chapitre traite différents systèmes proposés pour résoudre des congestions qui ouvrent la voie à plusieurs chercheurs afin qu'ils contribuent via des systèmes d'information tels que le big data, l'internet des objets et l'intelligence artificielle.

C'est ainsi que les travaux tels que des méthodes de prédiction basées sur des données historiques et des données en temps réel qui, après traitement, sont envoyés aux utilisateurs via une architecture client-serveur [3], dans [4], les auteurs ont comparé les algorithmes KMeans, C-means Fuzy, KMeans Fuzy dans un réseau VANET, ils ont considéré la vitesse, la consommation de carburant et l'émission de CO₂. Dans [5], les auteurs ont combiné l'internet des objets et les réseaux de neurones de type LSTM (Long Short Term Memory), ils considèrent la vitesse recueilli par deux capteurs placés dans deux directions Nord et Sud alors que les auteurs [6] ont classifié des données de type normal et anormal par WEKA ([131]). Les données anormales sont considérées et classifiées par KMeans et les arbres de décisions dans un temps record avec les paramètres météorologiques. Les travaux dans [7] utilisent une architecture IoT à quatre couches puis comparent les algorithmes DBSCAN, KMeans, Affinity propagation basés sur la classification non supervisée puis ont essayé d'apporter des solutions face à ce phénomène de congestion. Cependant des insuffisances et limites liées à la gestion de la distance inter-véhicule et le paramètre temps lié au trafic ont été notées dans le mécanisme de résolution. Pour combler certains manquements notés, dans [8] les auteurs prennent comme paramètre la vitesse et la densité pour illustrer le degré de la congestion qui est lié à d'autres paramètres. Pour une meilleure illustration, nous proposons une solution basée sur l'algorithme KMeans et les réseaux de neurones centroïdes.

I. CONTEXTE

De plus en plus on assiste à l'inadéquation existant entre la construction des équipements et le rythme de la croissance urbaine.

En effet, les villes poussent en hauteur, s'étendent en surface et exigent une construction de plus en plus rapide de logements et d'édifices publics ce qui rend nécessaire si non indispensable une interrogation sur les moyens comme le surcharge du réseau de transport. Cette congestion devient un phénomène de grande envergure nécessitant l'implication de tous les acteurs confondus que ce soit du côté des gouvernants, des gouvernés et des entreprises ...

Pour mettre en œuvre de telles solutions, le besoin des mises à jour très fréquents des données issues du trafic routier reste un des moyens proposés. Plusieurs solutions ont été proposées pour éradiquer ce fléau touchant plus les pays en voie de développement.

Des architectures différentes ont été proposées et améliorées en fonction des cas d'utilisations de types de schémas variés.

Ces données générées par ces architectures demandent de bons outils afin de les exploiter à des fins de bons usages pour une meilleure réponse à la congestion du réseau routier.

PROBLEMATIQUE

La robustesse de l'internet des objets repose sur le fait que sa capacité des objets à capter, à communiquer et à analyser des données. Cependant pour plus de lucidité, ces données ont besoin des techniques de l'intelligence artificielle pour une meilleure exploration qui peut être appliquée dans le domaine des transports routiers en milieu urbain. Des travaux de recherche tels que [6] avec les algorithmes arbres de décisions et KMeans ont classifié les données du trafic [132], modifié l'algorithme KMeans par les réseaux de neurones Hopfield de types récurrent et ainsi exploré les pistes qui combinent l'IA et l'IoT pour venir à bout à ce problème de congestion. Dans [133], une méthode basée sur la vitesse moyenne et la densité des véhicules qu'ils utilisent comme paramètres mais le problème de distance n'est pas prise en compte.

Plusieurs articles ont été proposés, parmi dont l'algorithme identification du modèle de congestion du trafic à l'aide de la technique de clustering KMeans [8], qui reste dans la même lancée pour régler la congestion via une architecture à trois niveaux en considérant les

paramètres de la vitesse moyenne et la densité. Nous pensons que ces caractéristiques seraient insuffisantes pour déterminer le degré de congestion donc pourraient être améliorées. Sur ce nous proposons une méthode d'architecture qui va garder les avantages de [8] mais plus adaptée aux meilleures capacités d'exploration des données.

Présentation du modèle Existant

Le système existant propose un modèle de reconnaissance des embouteillages, pour cela il envisage une segmentation des routes en fonction des paramètres de trafic à savoir la densité et la vitesse moyenne des véhicules [8]. En fonction de ces paramètres les données de vitesse sont classées en fonction des données issues du trafic sont regroupées et constituées de segments de routes de même types. Chaque cluster ou groupe va contenir des données de vitesse du trafic de même type.

1. Architecture du système

L'architecture proposée est composée de trois unités fondamentales : unités d'acquisition de données, unités de transmission, et unité de traitement de données [133].

a. Unités d'acquisition de données

Pour des stratégies appropriées et de planifications, l'acquisition de données en temps réel est importante. Il existe différentes manières de collecter des informations sur le trafic, ici le système prend en compte des capteurs fixes sur une route pour collecter des données [133].

b. Unités de transmission de données

Cette unité est chargée de transmettre les données acquises par communication sans fil ou filaire depuis le site vers le serveur distant pour un traitement ultérieur.

c. Unités de traitement de données

Les données de trafic qui ont été collectées par des capteurs sont traitées par une unité de traitement pour identifier des segments routiers problématiques dans le routier. Le système utilise la technique de regroupement KMeans pour former un certain nombre de groupes de différents segments de route. Ces connaissances peuvent être utilisées efficacement par les décideurs politiques pour une planification appropriée afin de minimiser les embouteillages.

Dans la section suivante, la technique de clustering KMeans est expliquée, la figure ci-après illustre l'architecture dans [8].

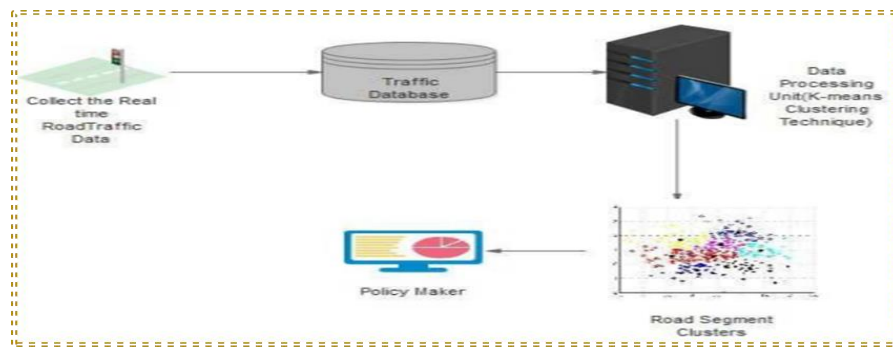


Figure 23 Système d'identification du schéma de congestion du trafic

d. Principe de fonctionnement de KMeans

Algorithme 1: KMeans Clustering

Saisir: base de données $D = d_1, d_2, d_3, \dots, d_n$, contenant m objets,

nombre de cluster k

Sortir: Ensemble de grappes

1 Choisi au hasard k , des données de D comme première centroïdes.

2 Calculer la distance euclidienne entre chaque objet $d_i \in D$ et chaque centroïde et attribuer au cluster le plus proche.

3 Recalculer le centre de gravité au moyen de tous les objets dans un amas nouvellement formé.

4 Répéter étapes 2 à 3, jusqu'à ce que le centre de gravité arrête de bouger.

5 Arrêter

Le système existant forme quatre clusters basés sur la densité et la vitesse moyenne : haute densité-faible vitesse (cluster 1), moyenne densité-faible vitesse (cluster 2), moyenne densité-vitesse modérée (cluster 3) et faible densité-haute vitesse (cluster 4).

Proposition d'une méthode efficace dans les algorithmes de résolution de la congestion

La solution que nous proposons est basée sur KMeans et les réseaux de neurones centroïdes, elle explore tous les avantages de [133] et tente de les améliorer.

D'abord nous avons noté que la concordance de la vitesse moyenne et la densité de véhicule seraient plus valables si on définissait le type de route que ce soit routes à chaussées séparées, autoroute, route à accès réglementé, route nationale, route départementale, route communale. En plus des paramètres de la vitesse et de la densité des véhicules nous rajoutons le paramètre de distance qui permettra de caractériser l'indice de proximité entre deux véhicules successifs.

Cette distance entre deux véhicules peut être calculée par une communication véhicule to véhicule (V2V) ; le non-respect de cette distance est souvent sources de collision, l'un des principaux facteurs de congestion.



Figure 24 la distance de proximité inter-véhicule

Notre méthode est basée sur une architecture à cinq niveaux et sera composée de trois phases.

Dans l'optique d'améliorer l'identification des schémas congestionnés, nous proposons une architecture modifiée basée sur l'architecture existante et composée de cinq niveaux :

Niveau 1 : cette couche est caractérisée par la collecte en temps réel des données véhiculaires qui sont envoyées au point de collecte. La collecte se fait via une unité d'acquisition de données composée de capteurs physiques placés au bord des routes.

Niveau 2 : à ce stade les données sont transmises vers un système de stockage via un réseau de communication sans fil,

Niveau 3 : ce niveau est composé de la base de données pour le stockage

Niveau 4 : c'est une couche primordiale pour de futures décisions, à ce niveau qu'intervient l'exploration des données, dans notre cas, elle est caractérisée par l'utilisation d'un algorithme de partition basé sur l'apprentissage non supervisé combiné à un autre algorithme de réseau de neurone de même type.

Niveau 5: c'est une génération de multiples services d'échanges de données via une architecture client-serveur.

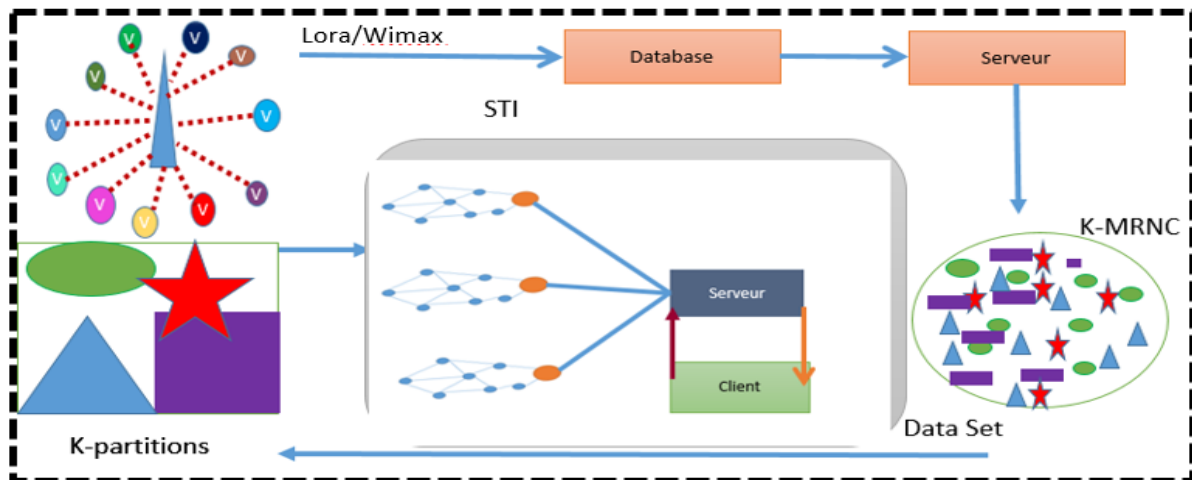


Figure 25 : système de gestion de congestion pour un réseau routier

Description du modèle

Notre point d'intervention se situe dans l'architecture va impliquer la couche de traitement puisqu'elle va gérer l'exploitation des données.

Du point de vue de l'architecture, nous avons défini cinq couches: une couche d'acquisition de données, une couche de transmission, une couche de stockage en plus de la couche de traitement qui intègre KMeans pour le partitionnement des données, s'ajoute une couche application. Cette dernière a pour avantage d'extraire en cas de besoin des données issues du trafic dans une période donnée via une interface client-serveur.

Pour le renouvellement du chef de cluster : la méthode proposée est une mise à jour des centres de cluster basée sur un réseau de neurone centroïdes. Au-delà de la partition statique des données et de leur classement par KMeans, notre méthode va dynamiser les partitions en intégrant la notion de temps proportionnellement à l'entrée ou la sortie de données. Ce qui permettra de porter le choix du CH sur le nœud ayant le plus de réserve énergétique et en conséquence accélérer la transmission sur le canal de communication.

Nous avons par simulation mis en œuvre les différents niveaux et nous avons utilisé KMeans pour le partitionnement et la classification.

Dans la suite nous allons d'abord présenter le fonctionnement des réseaux de neurones centroïdes avant d'aborder la partie simulation.

1. Fonctionnement de l'algorithme des réseaux de neurones centroïdes

Les réseaux de neurones centroïdes utilisent des algorithmes pour mettre à jour des données notamment les CH. Dans le cas de la mise à jour des clusters dans KMeans, ces algorithmes calculent le poids des données puis les comparent. Si une donnée a un poids supérieur à l'autre, elle sera définie comme CH. L'algorithme 1 montre comment choisir le centroïde en fonction des poids :

Le réseau de neurone centroïdes est un algorithme d'apprentissage compétitif non supervisé qui se base sur l'algorithme KMeans.

Le réseau de neurone centroïdes(RNC) associe la date lors de la formation des clusters à chaque entrée ou sortie d'une donnée x à l'instant t .

Les réseaux de neurones centroïdes vont apporter la fonction d'incrémentatation à KMeans par une mise à jour des données.

KMeans va faire la partition initiale puis K-RNC (k va intégrer un réseau de neurone centroïde).

K-RNC se base sur KMeans pour faire une classification plus complète avec l'intégration du calcul du poids en fonction du temps des données reçues ou perdues.

D'abord il calcule le poids de (x) qui prend en entrée une donnée et le temps d'entrée de cette dernière.

Ensuite il compare les poids le cluster Head(ch) avec le poids des données (x) si leurs poids $W(x)$ est inférieur à $W(ch)$ il continue sinon il y a un nouveau neurone gagnant et le CH devient un Cluster Member(CM).

Sur cette base il va classer les données en fonction de leurs appartenances en plus le paramètre temps inclus permettra aux usagers de choisir les meilleurs moments et décideurs bien voire leurs systèmes de prédiction.

2. Classification dans KMeans

C'est ainsi que nous proposons sa combinaison avec KMeans pour une meilleure classification des données afin d'apporter plus d'efficacité lors de l'exécution de KMeans pour la détection et la prédiction des embouteillages dans les zones urbaines.

Dans la suite de nos travaux, avant d'aborder la partie implémentation nous allons donner quelques paramètres essentiels pour le contrôle de congestion en situation réelle. Pour chaque paramètre donné nous indiquons les valeurs utilisées.

La vitesse moyenne : c'est la vitesse moyenne que roule les véhicules au passage du tronçon considéré différente de celle définie par le code de conduite.

La densité : elle est définie comme étant le nombre maximal de véhicule qu'occupe le tronçon qui est égale à 40 unités dans notre cas d'étude.

Le débit : il est défini comme le taux d'évacuation des 40 unités tous les 5 minutes, il prend trois états :

Libre : lorsque l'accès évacue un nombre de véhicules supérieur ou égale à 40 unités tous les 5 minutes.

Moyen : lorsque la capacité d'évacuation est dans l'intervalle de 37 à 39 unités tous les 5 minutes.

Congestionné : il apparait quand les véhicules restent sur le tronçon de plus de 5 minutes et le flux de sorti est inférieur ou égale à 30 unités.

Le temps : c'est un paramètre important, il correspond à la durée du trajet des véhicules entre l'entrée et la sortie du tronçon.

L'identifiant (Id) : il correspond à l'adresse unique attribuer au véhicule dans le réseau VANETs, lors de son à la station de base jusqu' à la portée de cette dernière.

Implémentation du modèle

Dans cette partie, nous allons à partir d'une collection de données (sous forme de fichier) qui reflète la phase d'acquisition des données par les capteurs, procédé à une classification sous forme de clustering via l'algorithme KMeans. Le clustering kmeans est une approche simple pour partitionner un ensemble de données en k clusters ou groupes distincts et non chevauchants.

L'objectif principal étant rendre efficace les algorithmes de contrôle de congestion basés sur KMeans. Pour se faire, nous utilisons une méthode qui définit le nombre de cluster sur lequel va opérer KMeans pour regrouper, classer et définir le chef de cluster.

3. VI.1. Paramètres de simulation

Nous nous sommes d'abord intéressés au paramètre de base qui est la vitesse moyenne des véhicules sur un tronçon considéré. Pour un type de tronçon, celle-ci est la distance entre les deux extrémités divisées par le temps mis. Elle peut être étroitement liée à la densité et au débit.

La densité : 40 unités.

Le débit : 40 unités tous les 5 minutes

- Libre : > 40 unités tous les 5minutes.
 - Moyen : de 37 à 39 unités tous les 5minutes.
 - Congestionné : ≤ 30 unités toutes les 5minutes
- Le temps : c'est un paramètre important, il correspond à la durée du trajet des véhicules entre l'entrée et la sortie du tronçon.

Ci-dessous un extrait du fichier avec l'ensemble des paramètres et données explorés par l'algorithme de classification. D'abord nous avons importé le fichier de donnée via la commande `pd.read_csv("etat-du-traffic1.csv", sep=";")`

	id_Vhcle	datetime	Temps	Densite	Debit	Cat	VitesseV	Param_Extrn
0	213972_D	01/01/2021	7h30mn	46	freeFlow	pl	30	0.0
1	213972_G	01/01/2021	7h35mn	44	freeFlow	pl	60	0.0
2	213973_D	01/01/2021	7h40mn	35	congested	pl	70	0.0
3	213973_G	01/01/2021	7h45mn	38	congested	pl	80	0.0
4	213974_D	01/01/2021	7h50mn	28	congested	pl	40	0.0
5	213974_G	01/01/2021	7h55mn	40	heavy	pl	30	0.0

Figure 26: Extrait du fichier trafic.csv

1. VI.2. L'environnement de simulation

Nous avons effectué nos simulations avec **Anaconda Navigator** qui est distributeur de python, un utilitaire offrant de nombreuses fonctionnalités. Après installation on obtient anaconda prompt option console en outre, anaconda Navigator, un logiciel à usage multiple avec des installations de librairie, de mise à jour en plus d'une grande documentation et de partage communautaire.

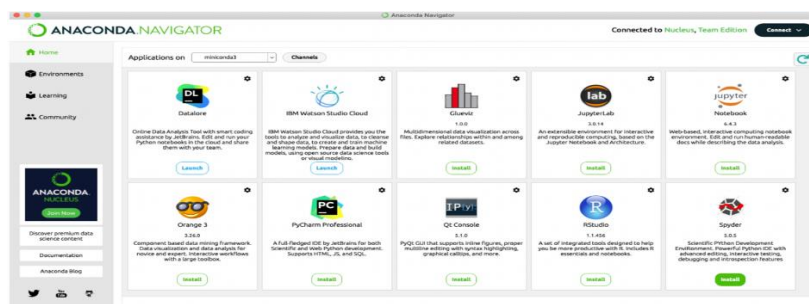


Figure 27: L'environnement et les composants d'Anaconda Navigator

Jupyter notebook est un composant faisant office d'un environnement de développement intégré, permettant d'exécuter le code dans une interface en combinant du code, de ses visualisations sous forme de document interactif. Par ailleurs son succès repose sur son portabilité, le fait de supporter de nombreux langages, c'est un standard du data scientist et du machine learning.

Pour fonctionner, cet environnement a été déployé sur une machine avec un Système d'exploitation Windows 10, d'une Ram de 4Go et d'un processeur Icore5 compatible avec l'environnement de développement.

2. VI.3. Les étapes de la simulation

Il faut noter que nous nous sommes mis juste dans le cadre d'un tronçon et que toutes les étapes de l'architectures physiques n'ont pas été implémentées. Tout d'abord nous devons spécifier le nombre de clusters tout en minimisant le critère de sommes des carrés intra clusters. En situation réelle, comme le montre notre architecture, nous devons commencer par la capture des données via les capteurs depuis les véhicules pour les transmettre via Lora vers le serveur de stockage. A la place, nous avons généré un fichier de données (314 lignes) exporté qui représente des données issues du trafic en temps réel de la ville de Rennes (France) [134] et auquel nous avons fait des ajustements de données pour l'ensemble des scénarios relatifs au débit, mais aussi pour prendre en compte l'existence de paramètres externes tels que la pluies, les incendies, des rafales de vent, le brouillard. Ce fichier regroupe des données issues de tronçons de rues, d'avenues, de route départementale... L'analyse s'est faite sur des données récoltées en temps réel sur un interval de trois minutes

Sur ces données, nous allons ensuite effectuer grâce à l'algorithme de type d'apprentissage non supervisé Kmeans une classification.

Ainsi, en premier lieu nous avons importé les bibliothèques requises tels que :

- Pandas, c'est une bibliothèque permettant de manipuler et d'analyser des données
- Numpy, c'est l'abréviation de « numerical python », bibliothèque open source en langage python pour la programmation en data science. Il comprend des objets multidimensionnels dans " arrays ".

- Matplotlib, la bibliothèque qui gère les graphiques avec l'ajout des éléments comme les lignes, des images ou des textes aux axes des graphiques.
- Kmeans : depuis la librairie Scikit Learn on a besoin de Kmeans et on va le charger via le sous module cluster de sklearn (from sklearn.cluster import Kmeans). Le tableau ci-après illustre l'architecture générale du datasets.

Quelques-unes des données ainsi importées peuvent être représentées graphiquement les unes en fonction des autres et inversement, même si cela ne renseigne en rien sur la classification qui sera faite.

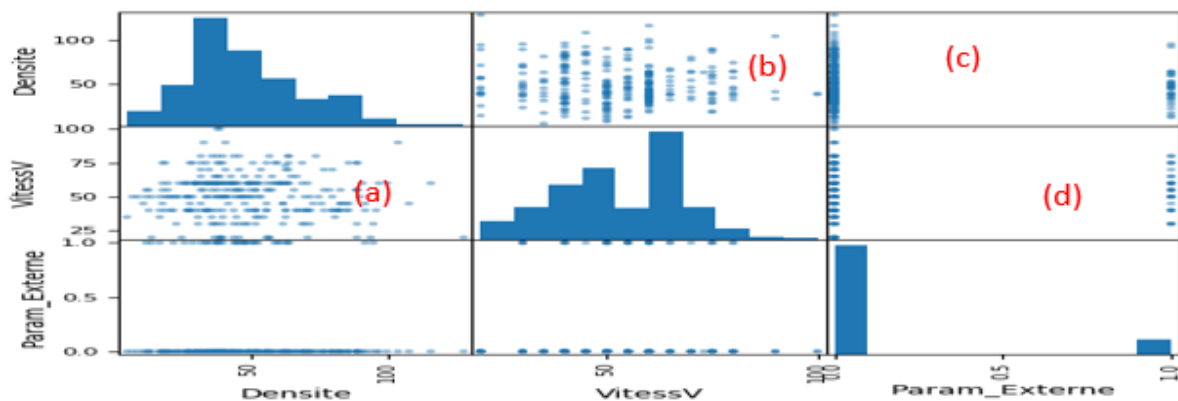


Figure 28: Représentation de la densité, de la vitesse et des paramètres externes

Nous déclinons dans ce qui suit les représentations cohérentes :

- (a) nous avons une représentation de la vitesse en fonction de la densité : quand le nombre de véhicules augmente, la densité augmente et la vitesse diminue
- (b) nous avons une représentation de la densité en fonction de la vitesse : la densité peut varier en fonction de la vitesse au cas où les véhicules qui auraient pu rouler très rapidement sur une voie libre diminuent leurs vitesses.
- (c) nous avons une représentation de la densité en fonction des paramètres externes : les paramètres externes tels que la pluie, la neige et le vent par exemple peuvent entraîner une baisse de la vitesse qui augmente la densité.
- (d) nous avons une représentation de la vitesse en fonction des paramètres externe : la présence de paramètres externes peut entraîner une diminution de la vitesse.

Nous pouvons aussi, comme le montre la figure 28, afficher vitesse en fonction de la densité.

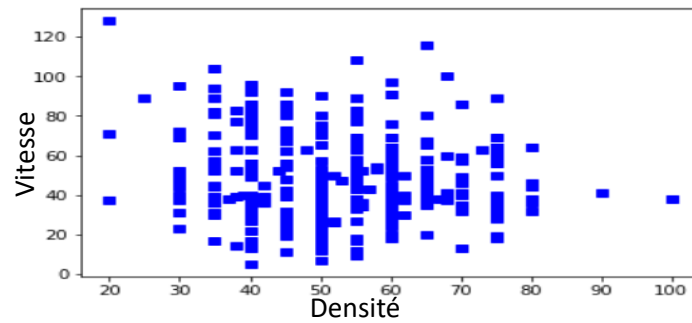


Figure 29: Le rapport des données de vitesse densité

Sur une portion des données (63 lignes) soit sur 20%, sur la base de la vitesse et de la densité, nous avons entraîné ces données via KMeans seul qui produit huit (8) groupes.

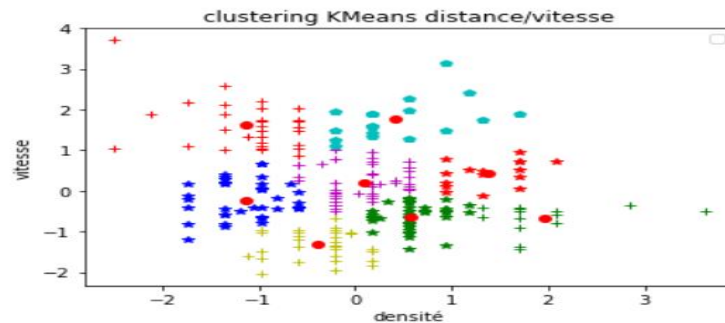


Figure 30: clustering KMeans sans la méthode Elbow

Nous constatons un chevauchement apparent entre les différents groupes. Cela est dû au nombre élevé de groupes. Pour une meilleure clusterisation un compromis doit être fait entre le nombre de groupe et les distances intra-clusters. La méthode Elbow [135] définit le nombre de clusters adéquat sur lesquels KMeans peut être entraîné avec différentes valeurs de k (k étant le nombre d'itérations) et va calculer la variance (distance intra-cluster) pour placer les données vers le cluster le plus proche. Ainsi, nous avons fait appel à la méthode Elbow qui combinée avec KMeans donnent cinq (5) groupes homogènes, et respectant les critères de classification basés sur la densité et la vitesse. Les figures [30] et [31] représentent respectivement la distance produite par Elbow en fonction du nombre de groupes et le regroupement en huit groupes produit par KMeans. En fonction de l'inertie, le nombre de groupes est représenté dans la figure ci-après.

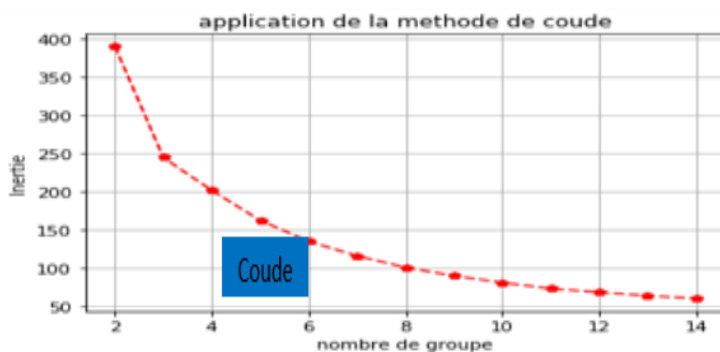


Figure 31: Distance dans Elbow en fonction du nombre de groupes

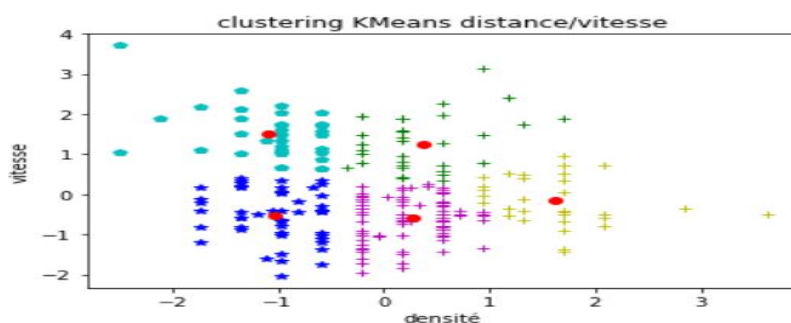


Figure 32: Nombre de clusters avec méthode du Elbow

La figure 31 donne une répartition des clusters avec leurs chefs. Pour plus de détails, nous montrerons dans la suite pour les différentes valeurs de K, la segmentation des données.

-Etape1

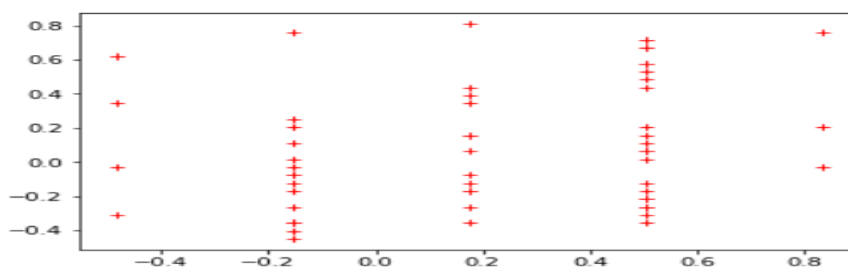


Figure 33: K=1

-Etape2

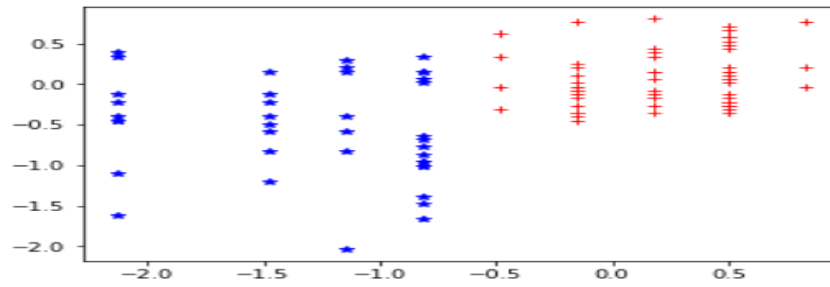


Figure 34:K=2

-Etape3

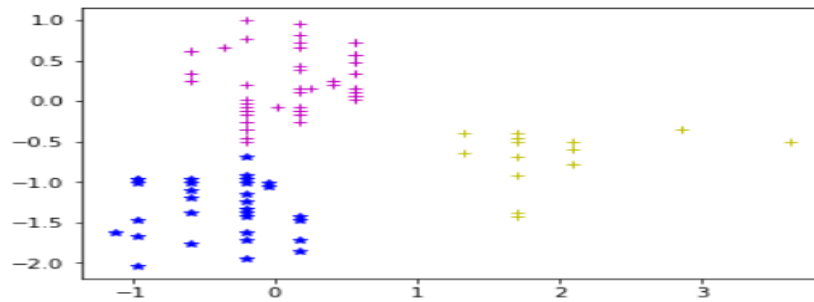


Figure 35: K=3

-Etape4

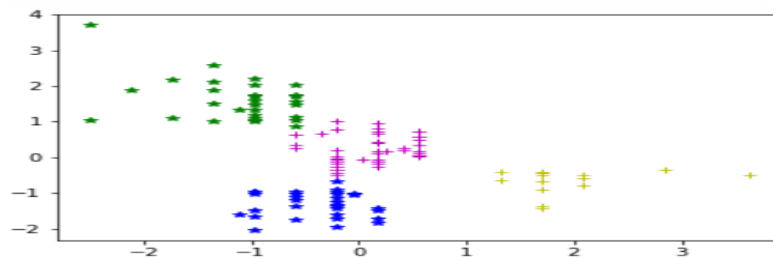


Figure 36 : K=4

-Etape5

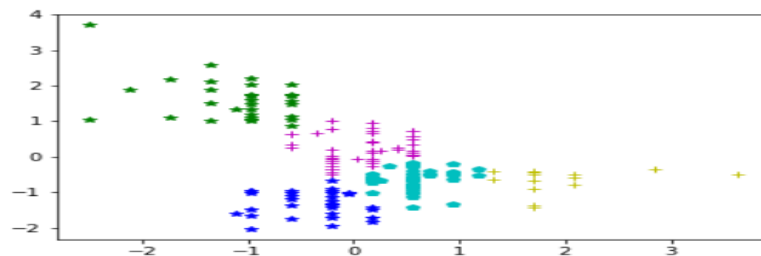


Figure 37:K=5

Comme le montre la figure 31, après la segmentation des données en cinq clusters, avec un chef de cluster pour chaque groupe appelé cluster Head (CH).

- Cluster 1(couleur cyan) : la densité est faible avec une vitesse élevée, ce qui correspond à un fonctionnement normal.
- Cluster 2(couleur jaune) : la densité est élevée et la vitesse faible ce qui correspond avec un fonctionnement normal.
- Cluster 3(couleur verte) : la densité est moyenne et la vitesse peu élevée, ce qui correspond à un fonctionnement normal.
- Cluster 4(couleur rose) : la densité est moyenne avec une vitesse faible, ce qui peut correspondre à un fonctionnement anormal car on s'attend à des vitesses relativement moyennes. Cela peut être dû à un manque de respect des règles de circulation ou aux paramètres externe tels que les intempéries : neige, pluie, incendie...
- Cluster 5(couleur bleu) : la densité est faible et la vitesse faible, ce qui peut être lié à l'état de la route, d'un accident des paramètres externes comme le brouillard, la pluie, la neige...

Pour voire les problèmes liés aux états de trafic notés dans les clusters bleu et rose, les décideurs devraient analyser les paramètres impliqués pour une meilleure prise de décision dans la lutte contre la congestion routière.

Conclusion

L'avènement des réseaux sans fils et des capteurs favorise les systèmes d'informatisation du réseau routier. Aujourd'hui dans la plupart des villes on assiste à des congestions dû à la base à un grand nombre de véhicules. Des recherches en perpétuelle évolution basées sur l'internet des objets et l'intelligences artificielle ont été proposées, et c'est ainsi que nous nous sommes intéressé à la problématique de congestion dont nous avons tenté d'apporter une contribution.

Nous avons créé une architecture à cinq niveau pour y appliquer l'algorithme KMeans combiné avec la méthode Elbow lors de la phase de traitement.

L'expérimentation a été faite sur des données en temps réel de la ville de Rennes. Ce qui permet de connaitre les voies où il ya vraiment un problème de congestion aux instants indiquées. Nous avons eu l'idée d'une fusion avec les réseaux de neurones centroïdes pour un meilleure -choix du chef de cluster et une mise à jour de celui-ci en cas d'arrivage de nouvelles données.

Conclusion Générale

L'écosystème internet des objets offre d'énormes avantages dans plusieurs secteurs, néanmoins il décèle des limites dont la congestion qui reste un sujet préoccupant pour ces adeptes. Ainsi, il urge d'impliquer d'autres technologies émergentes tels que l'intelligence artificielle. C'est dans ce contexte que nous avons proposé d'implémenter sur une architecture que nous proposons, une approche d'apprentissage non supervisé pour une meilleure partition des données dans un réseau d'internet des objets.

Dans ce même sillage nous avons envisagé de combiner l'un des algorithmes les plus populaire en termes de partition non supervisé KMeans avec les réseaux de neurones centroides pour mettre en place un mécanisme de gestion de la congestion qui peut être applicable à des données de trafic d'une quelconque agglomération.

Cette méthode a été implémentée sur la base de la vitesse moyenne des véhicules et de la densité des tronçons, et les résultats de simulation permettent pour un certain nombre de routes, de connaître celles qui présentent un comportement anormal. Ce qui pourra permettre aux décideurs d'identifier les routes à surveiller en un instant en vue de pouvoir déterminer les causes qui sont à l'origine de la congestion.

En perspectives, nous envisageons une phase expérimentale depuis la phase de collecte jusqu'à l'analyse en vue de confronter la situation réelle avec la simulation.

Pour mieux améliorer l'efficacité d'exécution de KMeans nous prévoyons de faire intervenir les réseaux de neurones centroides pour meilleur choix du chef de cluster.

Bibliographie et Webographie

- [1] M. N. F. D. G. a. all, «PROFIL URBAIN NATIONAL DU SÉNÉGAL».
- [2] G. Lesteven, «Les stratégies d'adaptation à la congestion automobile dans les grandes métropoles. Analyse à partir des cas de Paris, São Paulo et Mumbai,» [En ligne]. Available: <https://tel.archives-ouvertes.fr/tel-00981240/document>.
- [3] D. D. M. P. P. N. R. P. Nandkumar Kulkarni, «Averaging Based Predictive Modelling for Traffic Congestion in IoT,» 2018.
- [4] S. M. U. B. Anita Mohanty¹, «Traffic congestion detection in a city using clustering techniques in VANETs,» Indian, 2019.
- [5] B. R. s. s. Sharmila Majumdar, «Sustainable Cities and Society 64 (2021) 102500 Congestion prediction for smart sustainable cities using IoT and machine learning approaches,» Septembre 2020.
- [6] M. V. Muntean, «Identifying Critical Traffic Flow Time Slots Using K-Means and Decision Trees».
- [7] P. T. R. K. T. Ziyue Wang, «A Dynamic Traffic Awareness System for Urban Driving».
- [8] Z. R. Md. Ashifuddin Mondal, «Identification du modèle de congestion du trafic à l'aide Technique de clustering K-means».
- [9] <https://www.data.gouv.fr/fr/datasets/etat-du-traffic-en-temps-reel/>, [En ligne].
- [10] F. LEMOINE, Centre d'Études et de Recherche en Informatique et Communications.
- [11] P. Thebault. [En ligne]. Available: https://pastel.archives-ouvertes.fr/pastel-00869385/file/THEBAULT_-_Pierrick.pdf.
- [12] Jean-Pierre Hauet, [En ligne]. Available: https://www.researchgate.net/publication/311844665_L%27internet_des_objets-Part1-les_reseaux_de_communication.
- [13] LeMagIT, [En ligne]. Available: <https://www.lemagit.fr/definition/Internet-des-objets-IoT>.
- [14] Cisco, [En ligne].
- [15] OpenEditions, [En ligne]. Available: <https://books.openedition.org/editionsms/84?lang=fr>.
- [16] <https://www.itu.int/itu-news/manager/display.asp?lang=fr&year=2005&issue=09&ipage=things&ext=html>, [En ligne].
- [17] [En ligne]. Available: <https://itelligencegroup.com/fr/local-blog/connaissiez-vous-l-histoire-de-l-internet-of-things/>.
- [18] [En ligne]. Available: <http://www.histoire-cigref.org/blog/histoire-des-objets-inanimés-aux-objets-connectés/>.
- [19] matooma, [En ligne]. Available: <https://www.matooma.com/fr/s-informer/actualites-iot-m2m/m2m-comment-connecter-vos-objets>.
- [20] IDNA, [En ligne]. Available: <https://www.idna.fr/2020/10/13/objets-connectés-reseaux/>.
- [21] JDN, [En ligne]. Available: <https://www.journaldunet.fr/web-tech/dictionnaire-de-l-iot/1181267-les-reseaux-iot/>.
- [22] [En ligne]. Available: <https://www.redhat.com/fr/topics/internet-of-things/what-is-iot>.

- [23] M. T. Hammi. [En ligne]. Available: <https://pastel.archives-ouvertes.fr/tel-01997261/document>.
- [24] TACTIS, [En ligne]. Available: <https://www.tactis.fr/iot-revolution-numerique/>.
- [25] Génération Nouvelles Technologies, [En ligne]. Available: <https://www.generation-nt.com/lpwan-definition-reseau-iot-actualite-1983146.html>.
- [26] IoT industriel, [En ligne]. Available: <https://iotindustriel.com/technologies-solutions-iiot/sans-fil/lpwan-un-etat-comparatif-des-differents-reseaux-iiot/>.
- [27] Bouygues ENTREPRISES, [En ligne]. Available: <https://www.bouyguetelecom-entreprises.fr/lexique/fibre-optique/>.
- [28] F. LEMOINE, Centre d'Études et de Recherche en Informatique et Communications.
- [29] R. MARQUES, «SYSTÈME DE PRODUITS ET SERVICES BASÉS SUR L'INTERNET DES OBJETS CONCEPTION ET IMPLANTATION PILOTE DANS UNE STATION-SERVICE,» [En ligne]. Available: https://publications.polymtl.ca/3196/1/2018_RafaelMarques.pdf.
- [30] Oracle France, [En ligne]. Available: <https://www.oracle.com/fr/internet-of-things/what-is-iiot/>.
- [31] «Znet,» [En ligne]. Available: <https://www.zdnet.fr/actualites/industrie-40-l-importance-de-securiser-un-avenir-de-plus-en-plus-connecte-39931463.htm>.
- [32] Oracle France, [En ligne]. Available: <https://www.oracle.com/fr/internet-of-things/what-is-iiot/>.
- [33] Wikimémoires, [En ligne]. Available: <https://wikimemoires.net/2019/09/domaines-d-applications-de-l-iiot/>.
- [34] «LesJeudis,» [En ligne]. Available: <https://blog.lesjeudis.com/10-applications-de-l-internet-des-objets-qui-revolutionnent-la-societe>.
- [35] M. T. a. all, «Initiation à la conception et réalisation d'un réseau de capteurs intelligent : Maison Intelligente».
- [36] S. Triki, «Système Multi-Agents Ambiant pour faciliter l'autonomie et l'accessibilité aux espaces publics des personnes ayant des déficiences cognitives».
- [37] LesJeudi, [En ligne]. Available: <https://blog.lesjeudis.com/10-applications-de-l-internet-des-objets-qui-revolutionnent-la-societe>.
- [38] [En ligne]. Available: <https://www.objetconnecte.com/smart-city-index-2019/>.
- [39] [En ligne]. Available: <https://itsocial.fr/innovation/objets-connectes/plateformes-iiot-europe/>.
- [40] [En ligne]. Available: <https://www.hpe.com/ch/fr/solutions/iiot-platform.html>.
- [41] [En ligne]. Available: <https://www.corning.com/catalog/coc/documents/articles/The-Practical-Internet-of-Things.pdf>.
- [42] [En ligne]. Available: <http://www.i-o-t.org/post/Internet-des-objets-et-perspectives-economiques>.
- [43] oracle-Sénégal, [En ligne]. Available: <https://www.oracle.com/sn/artificial-intelligence/what-is-ai/>.
- [44] actualité informatique, [En ligne]. Available: <https://actualiteinformatique.fr/intelligence-artificielle/qu-est-ce-que-intelligence-artificielle-ia>.
- [45] Hewlett Packard Enterprise, [En ligne]. Available: https://www.hpe.com/emea_africa/fr/what-is/artificial-intelligence.html.
- [46] <https://www.normalesup.org/~pastre/IA.pdf>, [En ligne].
- [47] <https://www.futura-sciences.com/tech/definitions/informatique-intelligence-artificielle-555/>, [En ligne].

- [48] L'USINEDIGITALE, [En ligne].
- [49] Tecno-Science.net, [En ligne]. Available: <https://www.techno-science.net/glossaire-definition/Intelligence-artificielle-page-2.html>.
- [50] LeMagIT, [En ligne]. Available: <https://www.lemagit.fr/definition/AI>.
- [51] Les Echos, [En ligne]. Available: <https://www.lesechos.fr/2017/08/1956-et-lintelligence-artificielle-devint-une-science-181042>.
- [52] ActuaIA, [En ligne]. Available: <https://www.actuia.com/dossiers/quelques-elements-concernant-lhistoire-de-lintelligence-artificielle-12/>.
- [53] Bluumi, [En ligne]. Available: <https://bluumi.net/intelligence-artificielle/>.
- [54] L. CEA. [En ligne]. Available: <https://www.cea.fr/comprendre/Pages/nouvelles-technologies/essentiel-sur-intelligence-artificielle.aspx>.
- [55] lebigdata.fr, [En ligne]. Available: <https://www.lebigdata.fr/machine-learning-et-big-data>.
- [56] E. Ferreira, «Apprentissage automatique en ligne pour un dialogue homme-machine situé».
- [57] SIEMENS, [En ligne]. Available: <https://new.siemens.com/fr/fr/entreprise/stories/industrie-du-futur/l-intelligence-artificielle-dans-l-industrie.html>.
- [58] [En ligne]. Available: <https://www.ibm.com/fr-fr/security/artificial-intelligence>.
- [59] <https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1486693-l-intelligence-artificielle-une-des-cles-de-la-securite-des-reseaux-d-entreprise/>, [En ligne].
- [60] [journaldunet. \[En ligne\]. Available: https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1486693-l-intelligence-artificielle-une-des-cles-de-la-securite-des-reseaux-d-entreprise/](https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1486693-l-intelligence-artificielle-une-des-cles-de-la-securite-des-reseaux-d-entreprise/).
- [61] PIXALARM, [En ligne]. Available: <http://www.pixalarm.net/fr/analyse-et-comptage-via-deep-learning/>.
- [62] SciencesTech, [En ligne]. Available: <https://www.rts.ch/info/sciences-tech/12585675-intelligence-artificielle-et-agriculture-lhumain-ne-sera-pas-remplace-de-sitot.html>.
- [63] «agence ecofine,» [En ligne]. Available: <https://www.agenceecofin.com/entreprendre/1103-86076-au-maroc-l-intelligence-artificielle-pour-optimiser-l-agriculture>.
- [64] CScience, [En ligne]. Available: <https://www.cscience.ca/2021/09/21/sami-4-0-le-robot-cueilleur-au-service-de-la-recolte-au-quebec/>.
- [65] REUSSIR Machinisme, [En ligne]. Available: <https://www.reussir.fr/machinisme/kubota-investit-dans-les-robots-volants-cueilleurs-de-fruits>.
- [66] L'EXPRESS, [En ligne]. Available: <https://www.lexpressmontcalm.com/article/2021/06/15/l-industrie-4-0-s-invite-dans-le-monde-de-l-agriculture>.
- [67] technplay, [En ligne]. Available: <https://technplay.com/comment-lintelligence-artificielle-a-revolutionne-la-domotique/>.
- [68] Selecta, [En ligne]. Available: <https://selectra.info/domotique/actualites/confort/previsions-tendances-2020#1-le-role-cle-de-l-intelligence-artificielle>.
- [69] LEMONITEUR, [En ligne]. Available: <https://www.lemoniteur.fr/article/la-ville-connectee-entre-intelligence-artificielle-et-libertes-publiques.2133344>.
- [70] DOMADOO , [En ligne]. Available: <https://blog.domadoo.fr/81772-domotron-le-hub-domotique-avec-intelligence-artificielle/>.
- [71] ISARTA Infos, [En ligne]. Available: <https://isarta.com/infos/intelligence-artificielle-quels-sont-les-defis-des-organisations/>.

- [72] Université de LAVAL, [En ligne]. Available: <https://www.administration-numerique.chaire.ulaval.ca/recherches/les-applications-et-les-defis-de-lintelligence-artificielle-dans-le-secteur-public>.
- [73] DIGITALWEEK, [En ligne]. Available: <https://mydigitalweek.com/ia-conversationnel-securite-les-tendances-de-lia-pour-2021/>.
- [74] Jean-Sébastien Desroches, [En ligne]. Available: <https://www.lavery.ca/fr/publications/nos-publications/3009-intelligence-artificielle-la-delicate-interaction-entre-les-defis-juridiques-et-technologiques.html>.
- [75] S. MAROUANE, «<https://revues.imist.ma/index.php/DOREG/article/view/21207>».
- [76] A. P. R. Bruno le Maire Cédric O, STRATEGIE NATIONALE POUR L'INTELLIGENCE ARTIFICIELLE PRESENTATION DU VOLET ECONOMIQUE.
- [77] [En ligne]. Available: <https://isarta.com/infos/intelligence-artificielle-quels-sont-les-defis-des-organisations/>.
- [78] LesEchos Start, [En ligne]. Available: <https://start.lesechos.fr/innovations-startups/tech-futur/comprendre-les-enjeux-de-lintelligence-artificielle-en-4-points-1177670>.
- [79] accenture, [En ligne]. Available: <https://www.accenture.com/fr-fr/insight-artificial-intelligence-of-future-growth>.
- [80] [En ligne]. Available: <https://lestempselectriques.net/index.php/2020/12/31/2021-quelles-perspectives-pour-lintelligence-artificielle/>.
- [81] [En ligne]. Available: <https://docs.ifaci.com/wp-content/uploads/2018/03/perspectives-internationales-ia-partie-3.pdf>.
- [82] IREF, [En ligne]. Available: <https://fr.irefeurope.org/Publications/Articles/article/Etat-des-lieux-et-perspectives-de-lintelligence-artificielle-en-France>.
- [83] IPsystemes. [En ligne]. Available: <https://www.ip-systemes.com/details-qu+est+ce+que+l+aiot+definition+applications+et+exemples-820.html>.
- [84] L'intelligence artificielle dans la société, [En ligne]. Available: <https://www.oecd-ilibrary.org/sites/47f9bf3a-fr/index.html?itemId=/content/component/47f9bf3a-fr>.
- [85] GEEKFLARE. [En ligne]. Available: <https://geekflare.com/fr/daily-life-ai-example/>.
- [86] BDM. [En ligne]. Available: <https://www.blogdumoderateur.com/intelligence-artificielle-5-applications-concretes/>.
- [87] A. P.-M. a. all, «Selected methods of artificial intelligence for Internet of Things conception».
- [88] J. Zhou, «AAIoT: Accelerating Artificial Intelligence in IoT Systems».
- [89] A. P.-M. a. all, «Selected methods of artificial intelligence for Internet of Things conception».
- [90] Z. a. all, «Accelerating Artificial Intelligence in IoT Systems».
- [91] A. A. O. e. all, «OSUWA en 2017 9th International Conférence on Computational Intelligence and Communication Networks, Application of Artificial Intelligence in Internet of Things».
- [92] Fadi Al-Turjman Editor, Transactions on Computational Science and Computational Intelligence.
- [93] [En ligne]. Available: <https://datascientest.com/apprentissage-supervise>.
- [94] LeMagIT, [En ligne]. Available: <https://www.lemagit.fr/conseil/Apprentissage-supervise-et-non-supervise-les-differencier-et-les-combiner>.
- [95] [En ligne]. Available: <https://fr.linedata.com/quest-ce-que-lapprentissage-supervise>.
- [96] E. Côme, «APPRENTISSAGE DE MODÈLES GÉNÉRATIFS POUR LE DIAGNOSTIC DE SYSTÈMES COMPLEXES AVEC LABELLISATION DOUCE ET CONTRAINTES SPATIALES».

- [97] G. CABANES, «Classification non supervisée à deux niveaux guidée par le voisinage et la densité».
- [98] W. A. B. ZOUNGRANA, «Application des algorithmes d'apprentissage automatique pour la détection de défauts de roulements sur les machines tournantes dans le cadre de l'Industrie 4.0».
- [99] C. Chesneau, «Elements de classification».
- [100] M. Boubou, «Contribution aux méthodes de classification non supervisée via des approches prétopologiques et d'agrégation d'opinions».
- [101] N. Gueye. [En ligne]. Available: <https://depot-e.uqtr.ca/id/eprint/9416/1/eprint9416.pdf>.
- [102] M. MENAI. [En ligne]. Available: <https://arxiv.org/pdf/1903.03775.pdf>.
- [103] N. ROGOVSCHI, «Classification à base de modèles de mélanges topologiques des données catégorielles et continues».
- [104] O. e. a. Horiya, «Classification non supervisée : Application de k-means».
- [105] K. Nadja, «Les techniques de clustering dédiés au données multimédia».
- [106] L. ALI, «SÉLECTION DES MOTS CLÉS BASÉE SUR LA CLASSIFICATION ET L'EXTRACTION DES RÈGLES D'ASSOCIATION».
- [107] R. RAKOTOMALALA, «MARKET DATA ANALYSIS».
- [108] H. HILALI, «APPLICA TION DE LA CLASSIFICATION TEXTUELLE POUR L'EXTRACTION DES RÈGLES D'ASSOCIATION MAXIMALES».
- [109] A. Ait-Mlouk, «Fouille de données et analyse de qualité des règles d'association dans les bases de données massives : d'association dans les bases de données massives :».
- [110] [En ligne]. Available: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/19540/Ghali_Ramla_2016_these.pdf?sequence=4&isAllowed=y.
- [111] wikiversity, [En ligne]. Available: https://fr.wikiversity.org/wiki/R%C3%A9seaux_de_neurones/Historique.
- [112] par Dong-Hyun Lee, [En ligne]. Available: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/21284/Lee_Dong-Hyun_2018_memoire.pdf?sequence=2&isAllowed=y.
- [113] [En ligne]. Available: <http://cedric.cnam.fr/~thomen/cours/DUI5/codes.html>.
- [114] J. Rynkiewicz, « Introduction au Deep Learning Le perceptron multicouche historique,» 2020. [En ligne]. Available: <http://samm.univ-paris1.fr/IMG/pdf/coursdeep2.pdf>.
- [115] S. D. M. C. B. S. L. C. E. L. D. R. D'ASSOCIATION, «LABIAD ALI».
- [116] O. Mastropietro. [En ligne]. Available: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/20192/Mastropietro_Olivier_2017_memoire.pdf?sequence=4&isAllowed=y.
- [117] E. M. Olasagasti, «Application des réseaux de neurones à l'identification d'un axe de machine-outil,» [En ligne]. Available: <https://tel.archives-ouvertes.fr/tel-00493012/document>.
- [118] [En ligne]. Available: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/25593/Scellier_Benjamin_2020_these.pdf?sequence=6&isAllowed=y.
- [119] sinateTechnologie, [En ligne]. Available: <https://blog.sinatechnologie.com/renforcement-des-reseaux-de-neurones-avec-des-capsules-capsule-networks-capsnet>.

- [120] L. Gloaguen. [En ligne]. Available: <https://www.spiria.com/fr/blogue/intelligence-artificielle/apprentissage-profond-la-revolution-des-reseaux-capsules/>.
- [121] J. Z. a. al, «AAIoT: Accelerating Artificial Intelligence».
- [122] A. A. O. a. al, «Application of Artificial Intelligence in Internet of Things».
- [123] P. K. D. P. a. all, «Beyond Automation: The Cognitive IoT.».
- [124] N. K. a. all, «Averaging Based Predictive Modelling for Traffic».
- [125] C. Zhang, «Accelerate Deep Learning in IoT Human-Interaction Co-Inference Networking System for Edge:».
- [126] Z. W. a. all, «A Dynamic Traffic Awareness System for Urban Driving».
- [127] M. A. M. a. all, «Identification du modèle de congestion du trafic à l'aide Technique de clustering K-means».
- [128] M. V. Muntean, «Identifying Critical Traffic Flow Time Slots Using K-Means and Decision Trees».
- [129] K. K. a. all, «A New Hybrid Routing Protocol Using a Modified K-Means Clustering Algorithm and Continuous Hopfield Network for VANET».
- [130] A. M. a. all, «Traffic congestion detection in a city using clustering techniques in VANETs».
- [131] <https://www.weka-service-public.fr/#:~:text=WEKA%20est%20une%20soci%C3%A9t%C3%A9%20fran%C3%A7aise,du%20domaine%20de%20l'%C3%A9dition.>, [En ligne].
- [132] ,. L. B. A. H. B. KHALID KANDALI 1, «A New Hybrid Routing Protocol Using a Modified K-Means Clustering Algorithm and Continuous Hopfield Network for VANET».
- [133] Z. R. Md. Ashifuddin Mondal, «Identifying Traffic Congestion Pattern using K-means Clustering Technique».
- [134] [www.data.gouv.fr](https://www.data.gouv.fr/fr/datasets/etat-du-traffic-en-temps-reel/), «<https://www.data.gouv.fr/fr/datasets/etat-du-traffic-en-temps-reel/>,» [En ligne].
- [135] <https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501875-k-means-comprendre-la-methode-de-partitionnement-de-donnees/>, [En ligne].
- [136] [En ligne]. Available: <https://www.netapp.com/fr/artificial-intelligence/what-is-artificial-intelligence/>.
- [137] L'USINEDIGITALE, [En ligne]. Available: <https://www.usine-digitale.fr/article/definir-les-mots-de-l-intelligence-artificielle-pour-mieux-la-comprendre.N799370>.
- [138] [En ligne]. Available: <https://iaobs.com/blog/ia-usages-transport/>.
- [139] lesEchos, [En ligne]. Available: <https://www.lesechos.fr/idees-debats/leadership-management/le-defi-de-lia-de-tous-pour-tous-et-par-tous-1244402>.
- [140] ausy by randStad, [En ligne]. Available: <https://www.blogdumoderateur.com/intelligence-artificielle-5-applications-concretes/>.
- [141] [En ligne]. Available: <https://www.iot.nc/solution-transport-et-logistique/>.
- [142] [En ligne]. Available: Source : <http://www.iotboulevard.com/blog/10-objets-connectes-pour-l-agriculture-et-l-elevage.html>.
- [143] [En ligne]. Available: <https://www.shutterstock.com/fr/image-photo>.
- [144] Cisco, [En ligne]. Available: https://cdn.iotwf.com/resources/71/IoT_Reference_Model_White_Paper_June_4_2014.pdf, https://www.cisco.com/web/PH/ciscoconnect/pdf/bigdata/jim_green_cisco_connect.pdf.

- [145] [En ligne]. Available: <https://hal.archives-ouvertes.fr/hal-01166052/document>.
- [146] CScience, [En ligne]. Available: <https://www.cscience.ca/2020/06/08/5-applications-de-lia-en-sante/>.
- [147] [En ligne]. Available: [https://blog.i2s.fr/deep-learning-role-et-usage-dans-la-vision-industrielle/#:~:text=L'apprentissage%20profond%20est%20d,d'exactitude%20des%20processus\)..](https://blog.i2s.fr/deep-learning-role-et-usage-dans-la-vision-industrielle/#:~:text=L'apprentissage%20profond%20est%20d,d'exactitude%20des%20processus)..)
- [148] [En ligne]. Available: Le Deep learning s'appuie sur et concerne des données non-structurées. Il permet, dans l'industrie, de repérer des pièces présentant un défaut et de transformer en action une observation sur une chaîne de montage.
- [149] [En ligne]. Available: <https://www.bing.com/images/search?q=ia+et+transport>.
- [150] journaldunet, [En ligne]. Available: <https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1486693-l-intelligence-artificielle-une-des-cles-de-la-securite-des-reseaux-d-entreprise/>.
- [151] SciencesTech, [En ligne]. Available: <https://www.rts.ch/info/sciences-tech/12585675-intelligence-artificielle-et-agriculture-lhumain-ne-sera-pas-replace-de-sitot.html>.
- [152] [En ligne]. Available: <https://docs.ifaci.com/wp-content/uploads/2018/03/perspectives-internationales-ia-partie-3.pdf>.
- [153] <https://ichi.pro/fr/reseau-de-neurones-centroides-un-algorithme-de-clustering-efficace-et-stable-196789172644391>, [En ligne].