

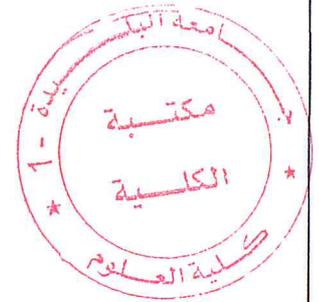
République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saad DAHLEB de Blida
Faculté des sciences
Département d'informatique



Mémoire présenté pour l'obtention
Du diplôme de Master Académique

Filière : Informatique
Spécialité : Ingénierie de logiciel



Thème :

Conception et Réalisation d'une plateforme Cloud intelligente pour la
Prédiction d'une urgence médicale.

Présenté par : Hamadou Meriem

Otmane Zineb

Soutenu devant le jury composé de :

Mr Chikhi nacim fateh

Président

Mme Toubaline nesrine

Examineur

Mme Elghers sabrina

Rapporteur

REMERCIEMENTS

Tout d'abord nous aimerions remercier DIEU de
nous avoir offert le
Courage pour réaliser ce projet et de nous avoir
éclairé notre lanterne

Tout au long des années d'études.

Nous adressons nos remerciements les plus
sincères à notre

Promotrice Mm ELGHÆRS SABRINA pour
leur patience, leur disponibilité,
Leur sacrifice, leur compréhension et leur conseil
bénéfique.

Nous remercions particulièrement les membres de
jury qui évalue notre travail.

Et le personnel de notre département ainsi que
nos enseignants qui ont fait prouvent
disponibilité durant nos études.

Enfin, nous remercions toute personne ayant
contribuée de près ou de loin à la réalisation de ce
projet.

Dédicaces

*Je dédie ce modeste à ma mère qui ont veillé nuit et jour pour
que je puisse arriver à ce stade de mes études*

*Je dédie ce modeste effort à l'esprit pour de mon père, et je
souhaite de mon dieu de l'introduit le paradis.*

*A ma promotrice, Mm ELGHERS SABRINA sans qu'elle ce
modeste travail n'aurait pas vu le jour.*

A ma binôme ZINEB

Aussi à mes frères et mes sœurs et leurs enfants.

Et je n'oublier pas mes chers amis

*A tous mes enseignants qui ont grandement contribué à
l'enrichissement de mes connaissances.*

Et à tous ceux qui me connaissent

*Pour terminer, un grand merci à mon époux **Bilel** a sa
patience, ces encouragements tout au long de préparation de
ce travaille*

MERIEM



Dedicace

*À travers ce modeste mémoire, je tiens à présenter mes sincères
dédicaces à mes très chers parents, j'espère que vous serez
toujours fiers de moi.*

A mes frères et mes sœurs, A toute ma famille.

A ma promotrice, Mm ELGHERS SABRINA

A mon binôme meriem.

*A l'ensemble des amis que j'ai connu pendant mes études et à
ceux qui m'ont prodigué leurs vifs conseils, encouragements et
témoigné de leur amitié.*

*A tous mes professeurs depuis la maternelle jusqu'à
l'université.*

A tous que j'aime et m'aiment.



ZINEB

Sommaire

| | |
|---|----------|
| Introduction générale..... | 1 |
| Chapitre I : Etat de l'art sur les réseaux de capteurs médicaux..... | 3 |
| 1. Introduction..... | 4 |
| 2. Les Capteurs..... | 4 |
| 2.1 Définition de capteur..... | 4 |
| 2.2. Les différents types de capteurs..... | 4 |
| 2.3. Architecture physique d'un capteur..... | 5 |
| 3. Les réseaux de capteur..... | 6 |
| 3.1. Architecture du réseau..... | 6 |
| 3.2. Contraintes de conception d'un réseau de capteur..... | 7 |
| 3.3. Les avantages d'un réseau de capteur sans fil..... | 8 |
| 3.4. Quelques applications des réseaux de capteur..... | 8 |
| 3.4.1. Applications militaires..... | 8 |
| 3.4.2. Applications à la sécurité..... | 8 |
| 3.4.3. Applications environnementales..... | 9 |
| 3.4.4. Applications commerciales..... | 10 |
| 3.4.5. Applications médicales et vétérinaires..... | 10 |
| 3.5. L'Internet des Objets (Internet of Things)..... | 10 |
| 3.5.1. Objets connectés : des applications multiples..... | 11 |
| 4. Les réseaux de capteur médical..... | 12 |
| 4.1. Exemples des capteurs dans le domaine médical..... | 12 |
| 4.2. WBAN (Wireless Body Area Network)..... | 13 |
| 4.3. Applications de suivi médical..... | 14 |
| 4.4. Besoins médicaux..... | 14 |
| 4.5. Indicateurs des applications de suivi médical..... | 15 |

| | |
|--|-----------|
| 4.6. Exemple des applications utilisant WSN dans le domaine médical..... | 16 |
| 5. Conclusion..... | 17 |
| Chapitre II : Etat de l'art sur les techniques de l'apprentissage automatiques..... | 18 |
| 1. Introduction..... | 19 |
| 2. L'apprentissage automatique..... | 19 |
| 2.1. Définition de l'apprentissage automatique..... | 19 |
| 2.2. Concepts et Sources de l'apprentissage automatique..... | 20 |
| 2.2.1. Qu'est-ce que l'apprentissage automatique..... | 20 |
| 2.2.2. Définition..... | 20 |
| 2.2.3. Modélisation..... | 20 |
| 2.2.4 Domaines d'applications de l'apprentissage automatique..... | 21 |
| 2.3. Types d'apprentissage..... | 21 |
| 2.3.1. L'apprentissage supervisé..... | 21 |
| 2.3.2. L'apprentissage non-supervisé..... | 23 |
| 2.3.3. L'apprentissage semi-supervisé..... | 24 |
| 2.3.4 L'apprentissage partiellement supervisé (probabiliste ou non)..... | 24 |
| 2.3.5 L'apprentissage par renforcement..... | 25 |
| 2.4. Facteurs de pertinence et d'efficacité..... | 25 |
| 3. Les techniques de l'apprentissage automatique..... | 26 |
| 3.1. Concept de l'analyse prédictive..... | 26 |
| 3.1.1. Algorithme des K plus proches voisins..... | 26 |
| 3.1.2. Réseaux de Neurones..... | 28 |
| 3.1.3. Classifieur Naïf Bayésien..... | 30 |
| 3.1.4. Machine à Vecteurs de Support..... | 31 |
| 3.1.5. Arbres de Décision..... | 33 |
| 3.1.5.1. Exemple..... | 33 |
| 3.1.5.2. Idée et Propriétés Générales..... | 34 |
| 3.1.5.3. Exemple Introductif..... | 34 |
| 3.1.5.4. Classification et Règles..... | 35 |
| 3.1.5.5. Définition du Formalisme : Arbres de Décision..... | 35 |

| | |
|---|-----------|
| 3.1.6. Régression logistique (SGV) | 35 |
| 3.1.7. Forêts d'arbres décisionnels (Random Forest) | 35 |
| 3.1.8. Le Boosting..... | 36 |
| 3.1.9.K_means..... | 36 |
| 3.2. Utilisations et applications des algorithmes..... | 36 |
| 4. Conclusion..... | 40 |
| Chapitre III : Le Cloud Computing..... | 42 |
| 1. Introduction..... | 42 |
| 2. Définition..... | 42 |
| 3. Les Caractéristique de Cloud Computing..... | 43 |
| 4. Les services de Cloud Computing..... | 44 |
| 4.1. IaaS : Infrastructure as a service..... | 44 |
| 4.2. PaaS: Platform as a service..... | 44 |
| 4.3. SaaS: Software as a Service..... | 45 |
| 5. Les Processuss de déploiement de Cloud Computing..... | 46 |
| 5.1. Cloud privé..... | 46 |
| 5.2. Cloud public..... | 47 |
| 5.3. Cloud communautaire..... | 47 |
| 5.4. Cloud hybride..... | 47 |
| 6. Les principaux fournisseurs de Cloud Computing..... | 47 |
| 7. Microsoft Azure..... | 48 |
| 7.1 Définition..... | 48 |
| 7.2 Description..... | 48 |
| 8. Azure ML..... | 50 |
| 8.1. Présentation de l'apprentissage automatique dans le cloud Microsoft Azure..... | 50 |
| 8.2. Azure ml studio..... | 51 |
| 9. Conclusion..... | 52 |
| Chapitre IV : Conception et Réalisation de l'application..... | 54 |
| 1. Introduction..... | 54 |
| 2. Conception..... | 54 |
| 2.1. Contexte du travail..... | 54 |
| 2.2. L'algorithme de prédiction choisis..... | 55 |

| | |
|---|----|
| 2.3. Notre processus ML..... | 56 |
| 2.4. Le web service..... | 60 |
| 2.5. Architecture proposé..... | 62 |
| 3.Realisation..... | 63 |
| 3.1. Création d'une expérience..... | 63 |
| 3.2. Importation de nos données..... | 64 |
| 3.3. Classification..... | 64 |
| 3.4. Mise en forme des données..... | 66 |
| 3.5. Création d'un processus de Machine Learning..... | 67 |
| 3.6. Publier le Web Service..... | 72 |
| 3.7. Test du Web Service..... | 73 |
| 3.8. Créer une application client C #..... | 73 |
| 4. Conclusion..... | 74 |
| Conclusion générale | 76 |

Liste des figures

- p.4 - Figure 01 : Schéma bloc d'un capteur analogique.*
- p.5 - Figure 02 : architecture physique d'un capteur.*
- p.6 - Figure 03 : Exemple d'un réseau de capteurs.*
- p.10 - Figure 04 : Application sur le contrôle de la qualité de l'eau.*
- p.12 - Figure 05 : Exemple de capteur de température.*
- p.12 - Figure 06 : Exemple de capteur de rythme cardiaque.*
- p.13 - Figure 07 : Un patient qui a cinq capteurs sur son corps.*
- p.16 - Figure 08 : Plateforme Mercury.*
- p.20 - Figure 09 : Schéma de modélisation d'une machine d'apprentissage.*
- p.21 - Figure 10 : l'apprentissage automatique.*
- p.22 - Figure 11 : l'apprentissage supervisé.*
- p.23 - Figure 12 : l'apprentissage non supervisé.*
- p.25 - Figure 13 : L'apprentissage par renforcement.*
- p.27 - Figure 14 : Schéma d'une classification par la méthode k-NN.*
- p.29 - Figure 15 : Schéma du perceptron.*
- p.30 - Figure 16 : Schéma du perceptron multi couches.*
- p.32 - Figure 17 : Principe du Séparateur à Vaste Marge (SVM).*
- p.33 - Figure 18 : Schéma d'un arbre de décision.*
- p.34 - Figure 19 : Schéma d'un arbre de décision.*
- p.35 - Figure 20 : formalisme d'un arbre de décision.*
- p.42 - Figure 21 : Les différents composants de Cloud Computing*
- p.46 - Figure 22 : Les services du Cloud Computing*
- p.51 - Figure 23 : le flux de travail dans azure machine Learning*
- p.52 - Figure 24 : la fenêtre de création d'expérimentation*
- p.55 - Figure 25 : données d'entrée et de sortie pour notre application*
- p.57 - Figure 26 : Notre processus Machin Learning.*
- p.61 - Figure 27 : Architecture client service de notre application*
- p.62 - Figure 28 : Architecture proposée*
- p.64 - Figure 29 : nouvel expérience*

p.66 - Figure 30 : Jointure des tables de données
p.66 - Figure 31 : Partition de données
p.67 - Figure 32 : Algorithme pour la classification
p.68 - Figure 33 : Sélection de la valeur à prédire
p.69 - Figure 34 : Workflow de notre processus
p.69 - Figure 35 : Visualisation de Evaluate Model
p.70 - Figure 36 : Workflow de processus avec autre algorithme
p.71 - Figure 37 : Comparaison entre les deux algorithmes
p.72 - Figure 38 : information sur le web service
p.73 - Figure 39 : Test de web service
p.73 - Figure 40 : Résultat de test de web service

Liste des tableaux

p.37 - Tableau 1 : Comparaison entre quelques algorithmes
p.58 - Table 02 : Table des données médicale d'un patient
p.58 - Table 03 : Table de classification pour RESP
p.59 - Table 04 : table de données de formation
p.65 - Table 05 : Table de classification

ملخص:

أطروحتنا هي بناء تطبيق لتوفير أفضل حل بديل للمريض القائم على المنزل أو المستشفى يستند في البداية على شبكات الاستشعار الطبية، وقادرة على قياس بعض من المعلمات الفسيولوجية ونقل البيانات. مهم للمرضين أو الأطباء حيث يجعل التدخل السريع إذا لزم الأمر.

والهدف من هذا العمل هو تحقيق منصة التنبؤ للطوارئ الطبية باستخدام البيانات التي تم جمعها من قبل بسن (شبكة الاستشعار الطبية) التي يتم تثبيتها على جسم المريض.

في هذا السياق، استخدمنا مايكروسوفت أزور آلة التعلم كوسيلة للتنمية السحابية من خلال الاستفادة من سعة تخزين البيانات ومكتبة خوارزمية كبيرة في التعلم الآلي لتحليل البيانات الطبية لدينا.

الكلمات المفتاحية: الاستشعار، بسن، تعلم الآلة، والتنبؤ، والحوسبة السحابية، والطوارئ الطبية

Résume :

Notre mémoire consiste à construire une application permettant de fournir à une personne maintenue à domicile ou à l'hôpital une meilleure solution alternative fondée dans un premier temps sur les réseaux de capteurs médicaux, capable de mesurer certains de ses paramètres physiologiques et de transmettre des données importantes aux infirmières ou médecins ce qui fait une intervention rapide en cas de nécessité.

L'objectif de ce travail est la réalisation d'une plateforme de prédiction d'une situation d'urgence médicale en utilisant les données captées par un BSN (réseau de capteur médicale) qui est installé sur le corps de patient.

Dans ce contexte nous avons utilisé Microsoft Azure machine learning comme un moyen de développement dans le cloud en exploitant sa capacité de stockage des données et sa grande bibliothèque d'algorithme en apprentissage automatique pour l'analyse de nos données médicales.

Les mots clés : capteur, BSN, apprentissage automatique, prédiction, cloud computing, situation d'urgence médicale.

Abstract :

Our thesis is to build an application to provide a home-based or hospital-based patient with a better alternative solution based initially on medical sensor networks, able to measure some of its physiological parameters and to transmit data important to nurses or doctors which makes a quick intervention if necessary.

The objective of this work is the realization of a prediction platform of a medical emergency using data collected by a BSN (Medical Sensor Network) that is installed on the patient's body.

In this context, we used Microsoft Azure machine learning as a means of cloud development by leveraging its data storage capacity and large algorithm library in machine learning for analysis of our medical data.

Keywords : sensor, BSN, machine Learning, prediction, cloud computing, medical emergency.

Introduction générale

Les applications médicales que permettent de relever et d'afficher différents paramètres physiologiques sont utilisées depuis plus de cent ans (p. ex. électrocardiogramme). Dans ces applications on utilise des fils pour relier des capteurs vers une unité qui calcule et présente des résultats. L'idée est désormais de pouvoir supprimer ces fils et surveille les résultats à distance.

Grâce au progrès dans le domaine des réseaux sans fil, des nouvelles applications sont conçues dans le domaine médical et de la santé. L'efficacité du personnel médical est augmentée en utilisant ces nouveaux outils et applications.

Dans le domaine de la santé, le suivi des patients à long terme et le suivi des personnes âgées, postopératoire et des maladies chroniques ainsi que les maisons intelligentes (des maisons équipées par des capteurs et la technologie de l'information pour permettre aux personnes qui devraient normalement être placées dans des établissements spécialisés de vivre autonome) sont devenus un objet de discussion réelle dans la communauté des chercheurs en informatique.

Les patients peuvent porter des capteurs qui surveillent les signes vitaux signalés en temps réel à leur médecin. Cela permet d'améliorer la qualité des soins de santé et d'économiser de l'argent aux patients.

Dans ce mémoire, on a proposé une architecture pour la surveillance des patient, qui sont munis de Smartphones, à partir des données captées par un BSN (Body Sensor Network). On a utilisé un algorithme sur le Cloud Microsoft Azure qui permettra la prédiction d'une situation d'urgence médicale à partir des données captées par le BSN. Le traitement se fera sur le cloud pour envoyer une alerte aux personne médicale (médecin, infirmier ou ambulance) en temps réel afin de réagir au bon moment.

Problématique

Notre problème majeur est de faire la prédiction machine des données récoltées à partir d'un BSN afin de prédire une situation d'urgence médicale.

Dans notre travail, on s'intéressera plus particulièrement aux réseaux de capteurs médicaux BSN, Un BSN (Body Sensor Networks) consiste en un ensemble de capteurs qui récoltent les données médicales d'un patient (les signes vitaux comme Pression artérielle, battement du cœur...etc.). Il est nécessaire, de veiller constamment aux signes vitaux des différents patients, pour contrôler effectivement leurs conditions physiques et de fournir des traitements urgent quand il s'agit d'une urgence, comme un changement anormal dans les battements du cœur par exemple.

Le travail qui nous a été demandé consiste à développer une plateforme de prédiction de situation d'urgence chez des patients, qui sont munis de Smartphones, à partir de données captées par un BSN.

Pour répondre à cette problématique, nous avons proposé une architecture pour la surveillance des patient, qui sont munis de Smartphones, à partir des données captées par un BSN (Body Sensor Network). On a utilisé un algorithme sur le Cloud Microsoft Azure qui permettra la prédiction d'une situation d'urgence médicale à partir des données captées par le BSN. Le traitement se fera sur le cloud pour envoyer une alerte aux personne médicale (médecin, infirmier ou ambulance) en temps réel afin de réagir au bon moment.

Organisation du mémoire

Ce mémoire est structuré en quatre chapitres comme suite :

Le premier chapitre présente des différentes notions de base liées aux capteurs, réseaux de capteur et réseau de capteur médicale.

Le deuxième chapitre de cette étude traite les notions fondamentales de l'apprentissage automatique et passer en revue les différents types d'apprentissage.

Le troisième chapitre présente quelque notion sur le cloud computing

Tandis que le dernier chapitre présente en premier la conception pour notre travail ainsi les différentes étapes de l'implémentation de notre application et les différents tests effectué et les résultats.

CHAPITRE I :

Etat de l'art sur les réseaux de capteurs médicaux.

1. Introduction

L'essor des nouvelles technologies ainsi que les progrès effectués dans les domaines des micro-électroniques, des télécommunications, des réseaux et du traitement de l'information ont entraîné l'apparition de nouveaux outils et objets communicants qui améliorent notre qualité de vie. Parmi ces objets communicants nous intéressons aux capteurs

Au cours des dernières décennies et grâce à l'avancée des systèmes embarqués et des technologies sans fil, les Réseaux de Capteurs Sans Fil ou « Wireless Sensor Network (WSN) », sont de plus en plus utilisés dans de nombreux domaines. Parmi ces domaines, nous nous intéressons aux WSN pour les applications médicales.

2. Les Capteurs

2.1 Définition de capteur

C'est un dispositif sensible à certains phénomènes physiques qu'il capte ou détecte et transformant cette grandeur physique en signal (en général électrique). [1]



Figure 01 : Schéma bloc d'un capteur analogique. [1]

2.2. Les différents types de capteurs [2]

- Les Capteurs de Lumière
- Les Capteurs de Force

-Les Capteurs de Son

-Les Capteurs de Vitesse

-Les Capteurs Météorologiques

- Les Capteurs de Position

2.3. Architecture physique d'un capteur [8]

Un capteur est composé de 3 unités :

- **L'unité d'acquisition** : l'unité d'acquisition est composée d'un capteur qui va obtenir des mesures numériques sur les paramètres environnementaux et d'un convertisseur Analogique/Numérique qui va convertir l'information relevée et la transmettre à l'unité de traitement.
- **L'unité de traitement** : l'unité de traitement est composée de deux interfaces, une interface pour l'unité d'acquisition et une interface pour l'unité de transmission. Cette unité est également composée d'un processeur et d'un système d'exploitation spécifique. Elle acquiert les informations en provenance de l'unité d'acquisition et les envoie à l'unité de transmission.
- **L'unité de transmission** : l'unité de transmission est responsable de toutes les émissions et réceptions de données via un support de communication radio.

Ces trois unités sont alimentées par une batterie comme la montre la figure ci-dessous :

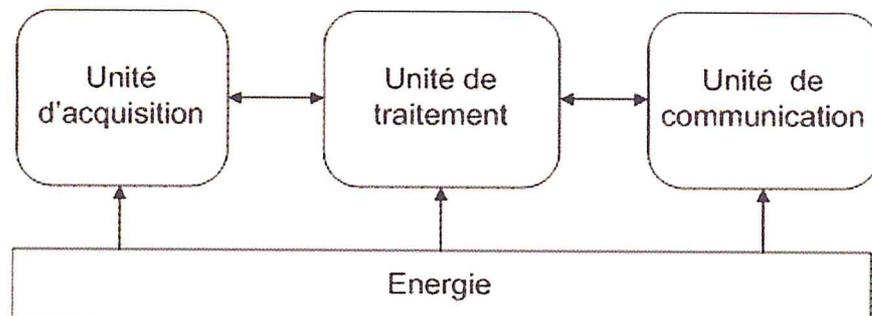


Figure 02 : architecture physique d'un capteur.

3. Les réseaux de capteur [3]

Un réseau de capteurs sans fil est un réseau ad hoc avec un grand nombre de nœuds qui sont des micro-capteurs capables de récolter et de transmettre des données environnementales d'une manière autonome. La position de ces nœuds n'est pas obligatoirement prédéterminée. Ils peuvent être aléatoirement dispersés dans une zone géographique, appelée « champ de captage » correspondant au terrain d'intérêt pour le phénomène capté.

Les données captées par les nœuds sont acheminées grâce à un routage multi-saut à un nœud considéré comme un "point de collecte", appelé nœud-puits (ou sink). Ce dernier peut être connecté à l'utilisateur du réseau (via Internet, un satellite ou un autre système). L'utilisateur peut adresser des requêtes aux autres nœuds du réseau, précisant le type de données requises et récolter les données environnementales captées par le biais du nœud puits.

3.1. Architecture du réseau [15]

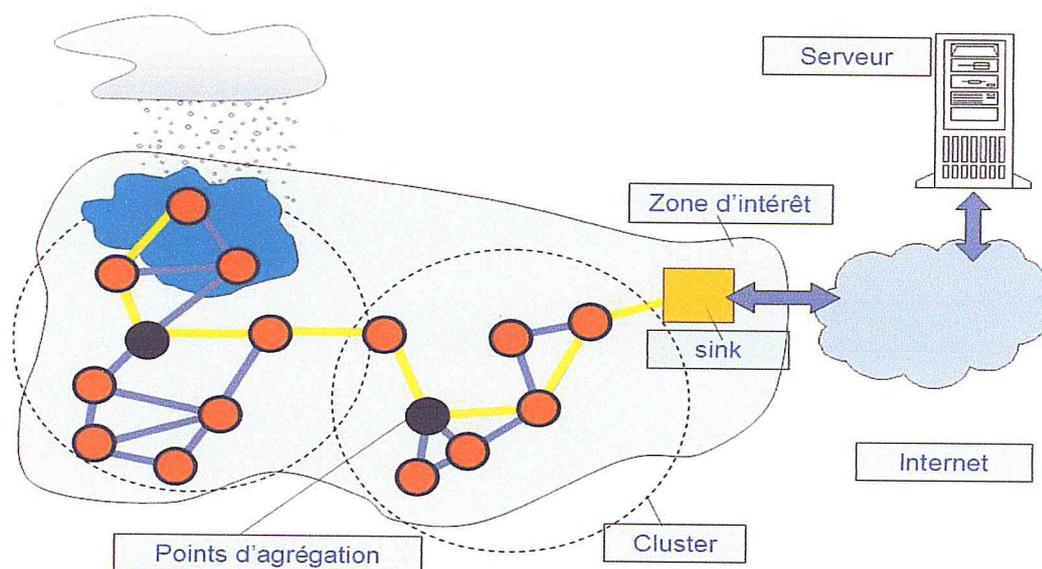


Figure 03 : Exemple d'un réseau de capteurs.

Le problème majeur dans les réseaux de capteurs est la consommation d'énergie, qui est fortement affectée par les communications entre les nœuds. Pour résoudre ce problème, des points d'agrégation sont introduits dans le réseau, comme le montre la *figure 03*. Cela réduit le nombre total de messages échangés entre les nœuds et permet d'économiser de l'énergie.

3.2. Contraintes d'un réseau de capteur [4]

Les principaux facteurs et contraintes influençant l'architecture des réseaux de capteurs peuvent être résumés comme suit :

La tolérance de fautes : c'est la capacité de maintenir les fonctionnalités du réseau sans interruptions dues à une erreur intervenue sur un ou plusieurs capteurs.

L'échelle : un nombre aussi important de nœuds engendre beaucoup de transmissions inter nodales et nécessite que le puits "sink " soit équipé de beaucoup de mémoire pour stocker les informations reçues.

Les coûts de production : Souvent, les réseaux de capteurs sont composés d'un très grand nombre de nœuds.

L'environnement : Les capteurs sont souvent déployés en masse dans des endroits tels que des champs de bataille ... Par conséquent, ils doivent pouvoir fonctionner sans surveillance dans des régions géographiques éloignées.

La topologie de réseau : Le déploiement d'un grand nombre de nœuds nécessite une maintenance de la topologie.

Les contraintes matérielles : La principale contrainte matérielle est la taille du capteur. Les autres contraintes sont que la consommation d'énergie doit être moindre.

Les médias de transmission : Dans un réseau de capteurs, les nœuds sont reliés par une architecture sans-fil

La consommation d'énergie : la durée de vie d'un capteur dépend grandement de la durée de vie de la batterie. C'est pour cette raison que les recherches actuelles se

concentrent principalement sur les moyens de réduire cette consommation.

3.3. Les avantages d'un réseau de capteur sans fil [10]

- liberté de mouvement des usagers et des terminaux.
- Moins de risque de rupture de liens que dans le filaire (coupure du câble, abimer les connecteurs et prise, etc....)
- Des économies à long terme (retour sur investissements)
- portabilité et discrétion : Les capteurs peuvent être portés sur le corps ou placés dans les environnements à surveiller.
- Facilité de déploiement et d'évolutivité.
- Temps réel et toujours fonctionnel : Les données capté permettant travailleurs d'urgence ou de soins de santé de réagir en temps réel.
- Reconfiguration et auto-organisation : Comme il n'y a aucune installation fixe, l'ajout et la suppression des capteurs reconfigurent instantanément le réseau.

3.4. Quelques applications des réseaux de capteur

D'après nos recherches on a trouvé que les applications les plus courantes de réseaux de capteur sont ceux-ci :

3.4.1. Applications militaires [3]

Comme dans le cas de plusieurs technologies, le domaine militaire a été un moteur initial pour le développement des réseaux de capteur.

Un réseau de capteurs déployé sur un secteur stratégique ou difficile d'accès, permet par exemple d'y surveiller tous les mouvements (amis ou ennemis), ou d'analyser le terrain avant d'y envoyer des troupes (détection d'agents chimiques, biologiques ou de radiations). Des tests concluants auraient déjà été réalisés par l'armée américaine dans le désert de Californie.

3.4.2. Applications à la sécurité [3]

Les structures d'avions, navires, automobiles, métros, etc. sont suivies en temps réel par des réseaux de capteurs, de même que les réseaux de circulation ou de distribution de

l'énergie. Les altérations de structure d'un bâtiment, d'une route, d'un quai, d'une voie ferrée, d'un pont ou d'un barrage hydroélectrique (à la suite d'un séisme ou du vieillissement) sont détectées par des capteurs préalablement intégrés dans les murs ou dans le béton, sans alimentation électrique ni connexions filaires. Certains capteurs ne s'activant que périodiquement peuvent fonctionner durant des années, voire des décennies. Un réseau de capteurs de mouvements peut constituer un système d'alarme distribué qui servira à détecter les intrusions sur un large secteur.

3.4.3. Applications environnementales [3]

Un réseau de capteur sans fil a été testé en Espagne (en laboratoire puis dans une rivière de l'Est de l'Espagne) pour le suivi en continu du taux de nitrates d'un cours d'eau. Il est associé à un système expert². Un triple capteur redondant modulaire permet pour chaque capteur d'améliorer la fiabilité du système, sans grands changements de coûts ou de consommation d'énergie. Ce type de réseau peut détecter, quantifier, dater, cartographier les pollutions azotées de l'eau et d'identifier leur source géographique, par exemple dans le cadre de l'application de la directive Nitrate, notamment en zone vulnérable.

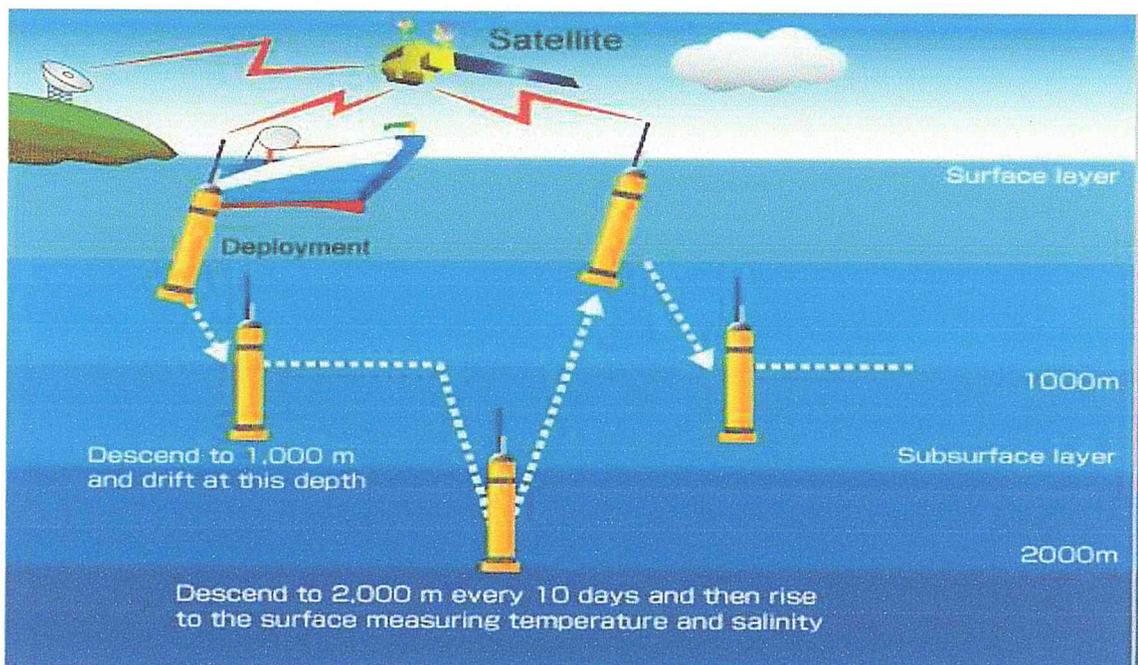


Figure 04 : Application sur le contrôle de la qualité de l'eau.

3.4.4. Applications commerciales [3]

Des nœuds capteurs pourraient améliorer le processus de stockage et de livraison (pour garantir la chaîne du froid en particulier). Le réseau ainsi formé, pourra être utilisé pour connaître la position, l'état et la direction d'un paquet ou d'une cargaison. Un client attendant un paquet peut alors avoir un avis de livraison en temps réel et connaître la position du paquet. Des entreprises manufacturières, via des réseaux de capteurs suivre le procédé de production à partir des matières premières jusqu'au produit final livré. Grâce aux réseaux de capteurs, les entreprises offrir une meilleure qualité de service tout en réduisant leurs coûts. Les produits en fin de vie vont être mieux démontés et recyclés ou réutilisés si les micros capteurs en garantissent le bon état.

3.4.5. Application médicales et vétérinaire [3]

La surveillance des fonctions vitales d'un organisme vivant est facilitée par des micro-capteurs avalés ou implantés sous la peau. Des gélules multi-capteurs ou des micro-caméras pouvant être avalé existe déjà, pouvant sans recours à la chirurgie, transmettre des images de l'intérieur d'un corps humain (avec une autonomie de 24 heures). Une récente étude présente des capteurs fonctionnant dans le corps humain, qui pourraient traiter certaines maladies. Un projet est de créer une rétine artificielle composée de 100 micro-capteurs pour corriger la vue. D'autres ambitieuses applications biomédicales sont aussi présentées, tel que : la surveillance de la glycémie, la surveillance des organes vitaux ou la détection précoce de cancers. Des réseaux de capteurs permet une surveillance permanente des patients et une possibilité de collecter des informations physiologiques de meilleure qualité, facilitant ainsi le diagnostic de quelques maladies.

3.5. L'Internet des Objets (Internet of Things) [12]

L'Internet of things repose sur les objets connectés. Un « objet connecté » possède généralement une sensibilité à son environnement par le biais de capteurs : température, position, pulsations cardiaques... Cet objet est relié à un mode de visualisation des informations, embarqué (directement sur l'objet, donc) ou distant (par l'intermédiaire d'un tableau de bord affiché sur l'écran d'un ordinateur ou d'une tablette). Il peut

également faire preuve d'une interactivité plus ou moins poussée avec d'autres objets, prendre des « décisions » par lui-même et gérer d'autres objets, qu'ils soient connectés ou non. À partir de tous ces objets connectés, l'Internet of things permet de construire des « maillages d'objets » – même à distance –, élaborant ainsi des systèmes de plus en plus intelligents.

3.5.1. Objets connectés : des applications multiples [11]

L'Internet des objets représente avant tout un formidable levier de performance, notamment pour les villes – qui deviendraient alors des « Smart Cities » – et plus généralement pour les administrations publiques. Cette performance s'articulerait autour de 3 axes :

- **Des économies budgétaires** : par exemple en Finlande, des capteurs intégrés aux poubelles permettent d'envoyer un message pour signaler que la collecte est nécessaire, ce qui a permis d'économiser 40% du budget de ramassage des déchets.
- **Des nouveaux revenus** : par exemple à New York, les écrans urbains qui fonctionnent 7 jours sur 7, 24h sur 24, permettent de surveiller la ville, de proposer des moyens de communications (ex : Wifi), de diffuser des informations, mais aussi de générer des revenus publicitaires.
- **Des avantages pour les citoyens** : par exemple à Nice, un « smart parking » avertit intelligemment les conducteurs du nombre de places restantes aux différents endroits de la ville, permettant ainsi de réduire les bouchons de 30%, d'augmenter les revenus des parkings, et de réduire les émissions de CO2.

L'Internet des objets pourrait soutenir de nombreux développements et améliorations comme par exemple l'habitat intelligent, le contrôle optimisé des consommations énergétiques et naturelles, le stationnement intelligent, une tarification routière plus adaptée, etc.

Il permettrait aussi d'accroître les performances administratives, en améliorant la productivité des agents et en réduisant les coûts d'exploitation. Cisco cite par exemple le secteur de la Défense.

4. Les réseaux de capteur médical

C'est ce que nous intéresse dans notre étude car c'est avec les informations qu'on va obtenir à partir de ces réseaux de capteur médicale (BSN) on arrive à prédire les situations d'urgence médicale d'un patient.

Les réseaux des capteurs sans fil médicaux sont utilisés aujourd'hui dans la médecine pour surveiller certains signes vitaux comme la température, la pression artérielle ou le rythme cardiaque, etc. Donc ça permettent non seulement d'améliorer la qualité de vie des patients, mais aussi le suivi des patients en temps réel et d'intervenir le plus rapidement possible dans les cas d'urgences. [7]

4.1 Exemples des capteurs dans le domaine médical

Les capteurs les plus connus dans la surveillance médicale et dont quelques-uns sont été utilisés [10] :

- Les capteurs de température et les capteurs de pression : pour le monitoring et de pression du patient.



Figure 05 : Exemple de capteur de température. [13]

- Les capteurs de rythme cardiaque : ils collectent le rythme cardiaque, la saturation d'oxygène et les données d'électrocardiogramme (ou EKG) et les relient à un réseau sans fil vers des appareils de réception.



Figure 06 : Exemple de capteur de rythme cardiaque. [14]

- Électromyogramme (EMG) : des capteurs qui mesurent directement les signaux de contrôle musculaire.
- D'autres capteurs : accéléromètre, gyroscope, gyromètre, glucomètre, pedomètre, etc.

Les capteurs communiquent entre eux et avec la station de base avec le protocole ZigBee. La station de base (ou la passerelle) relie le réseau PAN (Personale Area Network) porté par la personne sous surveillance au réseau Internet. Cette passerelle peut-être un téléphone mobile, un PDA, un micro-ordinateur ou n'importe quel dispositif relié au PAN d'une part et à Internet d'autre part.

4.2. WBAN (Wireless Body Area Network) [9]

Le WBAN (Wireless Body Area Network) consiste en plusieurs capteurs physiologiques (ECG, pulse oxymètre, température) disposés sur le corps humain pour collecter les signes vitaux comme l'ECG, la tension artérielle, le pulse oxymètre, puis transmettre les mesures au puits (sink), PDA, mobile phone ou personale computer. Ensuite, l'information est transmise en temps réel via internet au personnel médical.

Le WBAN peut être utilisé à l'intérieur de l'hôpital pour le monitoring des patients qui sont en situation critique.

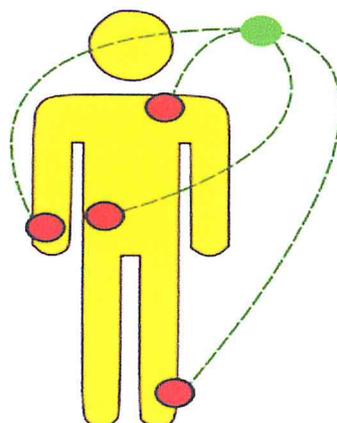


Figure 07 : Un patient qui a cinq capteurs sur son corps.

4.3. Applications de suivi médical [5]

Les applications qui permettent de relever et d'afficher différents paramètres physiologiques sont utilisées depuis plus de cent ans (p. ex. électrocardiogramme). Dans ces applications on utilise des fils pour relier des capteurs vers une unité qui calcule et présente des résultats. L'idée est désormais de pouvoir supprimer les fils et les rassembler dans le Smartphone d'un patient.

4.4. Besoins médicaux [6]

De nouveaux besoins ont ramené à chercher des actions de développement dans le domaine de la santé :

- Une population vieillissante
- Un manque de personnel hospitalier
- Une demande de soin à domicile
- Une restructuration du modèle familial

4.5. Indicateurs des applications de suivi médical [5]

Commençons par décrire les données que l'on peut obtenir lors d'une visite chez le médecin. Ces informations sont également enregistrées lorsque nous sommes sur un lit d'hôpital. Lors d'une visite chez un médecin il est possible d'obtenir les indicateurs suivants :

1-Les signaux vitaux [5]

Parmi les signaux vitaux les plus fréquemment collectés on peut souligner :

- **La pression artérielle**, avec une fréquence d'échantillonnage de plusieurs fois par jour. Normalement elle est représentée par un intervalle de valeurs avec des valeurs différentes selon qu'il s'agit de la pression systolique ou de la pression diastolique.
- **Le pouls** est représenté par une valeur entre 0 et 220.
- **La fréquence respiratoire.**
- **La température** est une valeur en degrés Celsius (ou Fahrenheit). Cependant, la valeur de la température peut changer selon l'endroit où la partie du corps utilisée pour prendre l'échantillon. Notons que la température moyenne a des petites variations selon l'endroit de mesure et l'âge du patient. Une prise de température, a besoin d'un calibrage pour déterminer les valeurs normales.

Le pouls, la fréquence respiratoire, et la température peuvent être pris avec des fréquences d'échantillonnage variables en fonction des besoins (mesures toutes les deux, quatre, ou six heures).

Il existe d'autres paramètres comme

- **la Saturation pulsée de l'hémoglobine en oxygène** ou SpO₂, mesurée grâce à des appareils comme l'oxymètre de pouls ou saturomètre. Une valeur normale de SpO₂ est autour de 95-100%, et représente le pourcentage d'hémoglobine qui est déjà liée avec l'oxygène. Ces appareils utilisent la lumière pour détecter les ions d'oxygène dans le sang.

4.6. Exemple des applications utilisant WSN dans le domaine médical

- **CodeBlue**

CodeBlue est un projet dans le domaine médical basé sur les réseaux de capteurs sans fil. Les buts spécifiques de ce projet concernent le suivi du patient pré-hôpital et dans l'hôpital pour les soins d'urgence. Ce projet possède le matériel ainsi que le logiciel suivant :

- Accéléromètre, gyroscope et électromyogramme(EMG) pour le monitoring des patients.
- Une plateforme (*figure 08*) Réseau de capteurs appelée Mercury portable pour l'analyse de mouvement haut -fidélit'

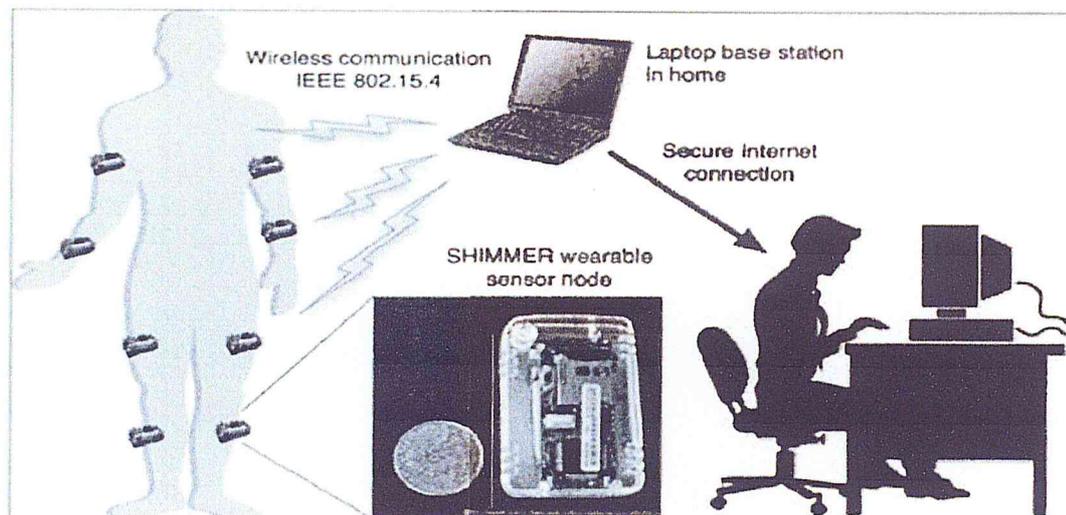


Figure 08 : Plateforme Mercury. [10]

Mercury est une plateforme de réseaux de capteurs conçus pour supporter des applications qui ont des données intensives et qui peut s'adapter aux fluctuations de la disponibilité des ressources. Les principaux défis abordés par Mercury sont : longue vie de nœud capteur pour conserver l'énergie de batterie afin que le capteur vive le plus long possible, un fonctionnement autonome et la nécessité pour le système d'ajuster automatiquement son comportement en réponse aux fluctuations de la bande passante radio et de la disponibilité de l'énergie.

La plateforme matérielle actuelle SHIMMER développée chez Intel. SHIMMER intègre un processeur TI MSP430, radio CC2420 IEEE 802.15.4, un accéléromètre triaxial et la batterie i-polymère rechargeable [10].

5. Conclusion

Nous avons défini dans ce chapitre qu'est-ce qu'un capteur ensuite on a montré la description et l'utilisation des réseaux de capteurs, plus particulièrement les BSN, les réseaux de capteur médicale qui sont le but de notre travail.

Dans le chapitre suivant on va jeter la lumière sur la notion d'apprentissage automatique ou on peut prédire des cas d'urgences à partir des données captées par le BSN.

CHAPITRE II :

Etat de l'art sur les techniques de l'apprentissage automatiques

1. Introduction

D'après [24] L'apprentissage automatique fait référence au développement, à l'analyse et à l'implémentation de méthodes qui permettent à une machine d'évoluer grâce à un processus d'apprentissage, et ainsi de remplir des tâches qu'il est difficile ou impossible de remplir par des moyens algorithmiques plus classiques.

Dans le présent chapitre, nous commencerons par introduire les concepts fondamentaux du l'apprentissage automatique ensuite on va expliquer quelques célèbres algorithmes de prédiction.

2. L'apprentissage automatique

Le but de cette étude est de s'initier à l'apprentissage automatique et à la modélisation des systèmes de classification, mais également d'étudier et pouvoir au final concevoir des algorithmes de prédiction.

2.1. Définition de l'apprentissage automatique [18]

L'apprentissage automatique (ou artificiel) (*machin-Learning* en anglais) est un des champs d'étude de l'intelligence artificielle. L'apprentissage automatique fait référence à la capacité d'un système à acquérir et intégrer de façon autonome des connaissances.

Cette notion englobe toute méthode permettant de construire un modèle de la réalité à partir de données, soit en améliorant un modèle partiel ou moins général, soit en créant complètement le modèle.

1) L'apprentissage humain (naturel) :

De la naissance à l'âge adulte, les êtres vivants acquièrent de nombreuses capacités qui leur permettent de survivre dans leur environnement. L'apprentissage d'un langage, de l'écriture et de la lecture sont de bons exemples des capacités humaines.

-2) l'apprentissage automatique qui permet d'avoir un système qui s'optimise en fonction de l'environnement, les expériences et les résultats observés. Par exemple, la reconnaissance de caractères manuscrits, tels qu'ils apparaissent sur une enveloppe. La difficulté tient au fait que la variété des formes rencontrée est infinie. L'apprentissage par cœur n'est pas possible, et il faut donc être capable de

généraliser à partir d'un ensemble d'exemples de caractères.

2.2. Concepts et Sources de l'apprentissage automatique

2.2.1. Qu'est-ce que l'apprentissage automatique [16]

La faculté d'apprendre de ses expériences passées et de s'adapter est une caractéristique essentielle des formes de vies supérieures. Elle est essentielle à l'être humain dans les premières étapes de la vie pour apprendre des choses aussi fondamentales que reconnaître une voix, un visage familier, apprendre à comprendre ce qui est dit, à marcher et à parler.

L'apprentissage automatique est une tentative de comprendre et reproduire cette faculté d'apprentissage dans des systèmes artificiels. Il s'agit, très schématiquement, de concevoir des algorithmes capables, à partir d'un nombre important d'exemples (les données correspondant à "l'expérience passée"), d'en assimiler la nature afin de pouvoir appliquer ce qu'ils ont ainsi appris aux cas futurs.

2.2.2. Définition [18]

Un programme d'ordinateur est capable d'apprendre à partir d'une expérience E et par rapport à un ensemble T de tâches et selon une mesure de performance P, si sa performance à effectuer une tâche de T, mesurée par P, s'améliore avec l'expérience E.

2.2.3. Modélisation d'une machine d'apprentissage [22]

L'apprentissage automatique d'une machine toujours concerne un ensemble de tâches concrètes - T. Pour déterminer la performance de la machine, on utilise une mesure de la performance P. La machine peut avoir à l'avance un ensemble d'expérience E ou elle va enrichir cet ensemble plus tard.

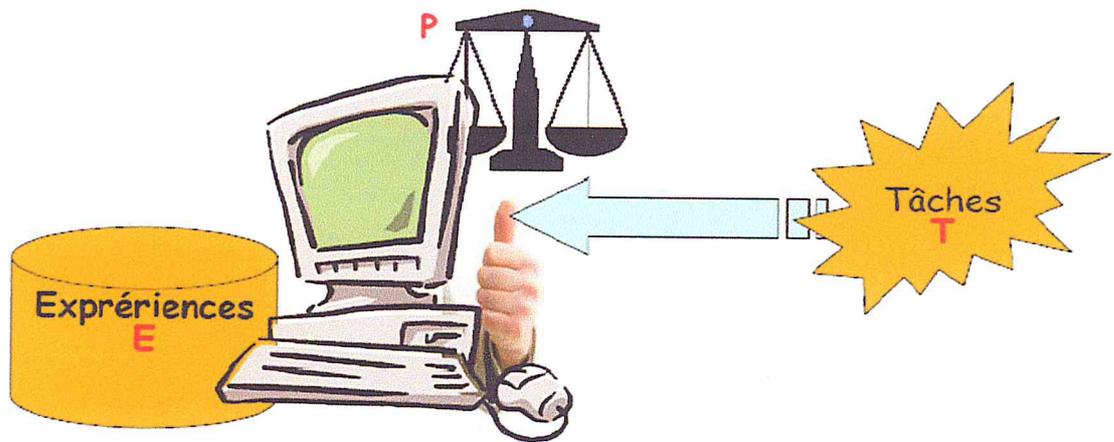


Figure09 : Schéma de modélisation d'une machine d'apprentissage.

Donc, l'apprentissage automatique pour la machine est qu'avec l'ensemble de tâches T que la machine doit réaliser, elle utilise l'ensemble d'expériences E telle que sa performance sur T est améliorée.

2.2.4 Domaines d'applications de l'apprentissage automatique [22]

L'apprentissage automatique s'applique à un grand nombre d'activités humaines et convient en particulier au problème de la prise de décision automatisée. Il s'agira, par exemple :

- D'établir un diagnostic médical à partir de la description clinique d'un patient.
- De donner une réponse à la demande de prêt bancaire de la part d'un client sur la base de sa situation personnelle.
- De déclencher un processus d'alerte en fonction de signaux reçus par des capteurs.
- De la reconnaissance des formes.
- De la reconnaissance de la parole et du texte écrit.
- De contrôler un processus et de diagnostiquer des pannes.

2.3. Types d'apprentissage

Les algorithmes d'apprentissage peuvent se catégoriser selon le mode d'apprentissage qu'ils emploient :

2.3.1. L'apprentissage supervisé

Il se base sur les points suivants :

- Si les classes sont prédéterminées et les *exemples* connus, le système apprend à classer selon un modèle de classement ; on parle alors d'apprentissage supervisé (ou d'analyse discriminante). [22]

-Un expert doit préalablement correctement étiqueter des exemples. L'apprenant peut alors trouver ou approximer la fonction qui permet d'affecter la bonne « étiquette » à ces exemples. -Parfois il est préférable d'associer une donnée non pas à une classe unique, mais une probabilité d'appartenance à chacune des classes prédéterminées (on parle alors d'apprentissage supervisé probabiliste). [22]

-L'analyse discriminante linéaire ou les SVM sont des exemples typiques.

- Autre exemple : en fonction de points communs détectés avec les symptômes d'autres patients connus (les exemples), le système peut catégoriser de nouveaux patients au vu de leurs analyses médicales en risque estimé (probabilité) de développer telle ou telle maladie. [22]

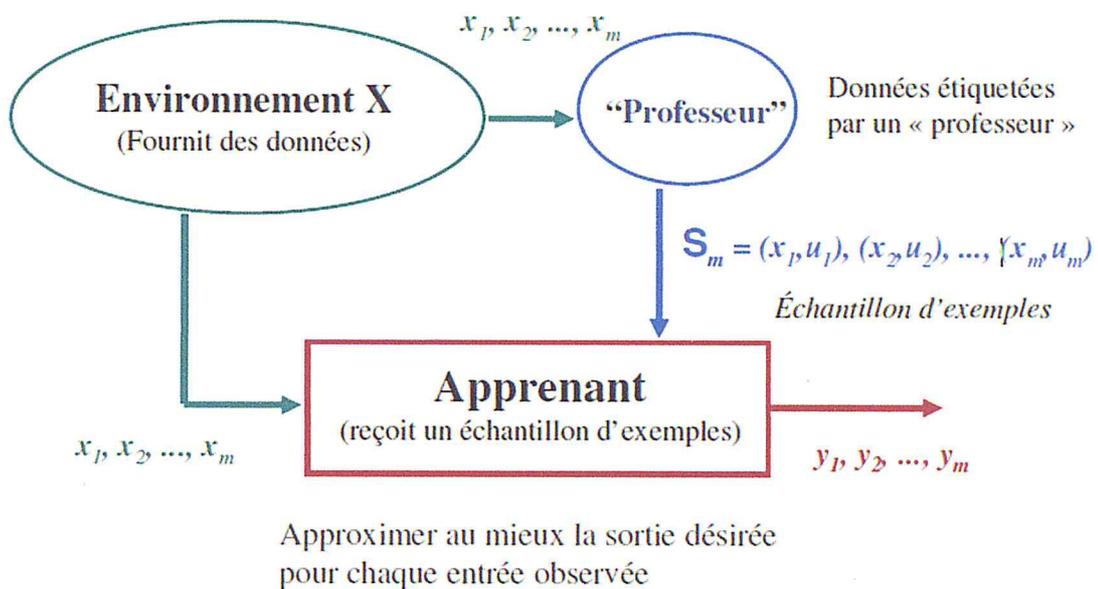


Figure10 : l'apprentissage automatique. [23]

Un expert est employé pour étiqueter correctement des exemples.

_ L'apprenant doit alors trouver ou approximer la fonction qui permet d'affecter la bonne étiquette à ces exemples [23]

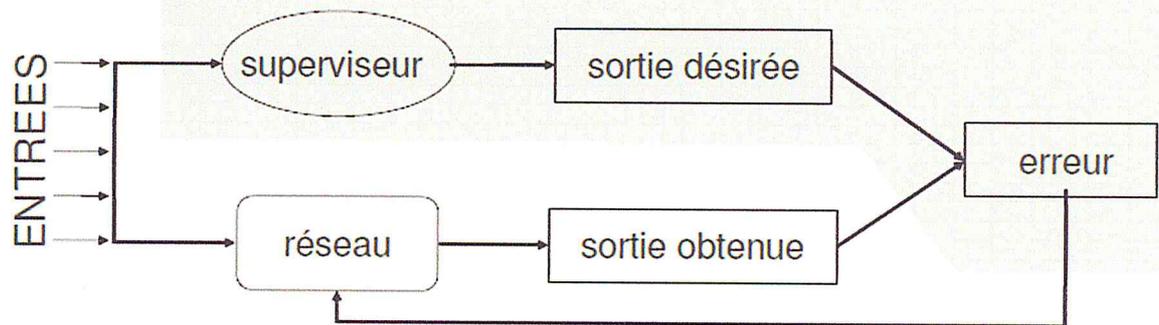


Figure11 : l'apprentissage supervisé. [23]

En fonction du résultat (comportement) supervisé que l'on veut obtenir, on peut distinguer deux types de problèmes [19] :

- **Régression** : lorsque le résultat supervisé que l'on cherche à estimer est une valeur dans un ensemble continu de réels.
- **Classification** : lorsque l'ensemble des valeurs de sortie est discret. Ceci revient à attribuer une classe (aussi appelée étiquette ou label) pour chaque vecteur d'entrée.

2.3.2. L'apprentissage non-supervisé

Quand le système ou l'opérateur ne dispose que d'exemples, mais non d'étiquettes, et que le nombre de classes et leur nature n'ont pas été prédéterminés, on parle d'apprentissage non supervisé (ou clustering).

Aucun expert n'est disponible ni requis. L'algorithme doit découvrir par lui-même la structure plus ou moins cachée des données. Le système doit ici dans l'espace de description (la somme des données) cibler les données selon leurs attributs disponibles, pour les classer en groupe homogènes d'exemples. La similarité est généralement calculée selon la fonction de distance entre paires d'exemples. C'est ensuite à l'opérateur d'associer ou déduire du sens pour chaque groupe. Divers outils mathématiques et logiciels peuvent l'aider. On parle aussi d'analyse des données en régression. Si l'approche est probabiliste (c'est à dire que chaque exemple au lieu d'être classé dans une seule classe est associé aux probabilités d'appartenir à chacune des classes), on parle alors de « *soft clustering* » (par opposition au « *hard clustering* ») [18].

Exemple : Un épidémiologiste pourrait par exemple dans un ensemble assez large de victimes de cancers du foie tenter de faire émerger des hypothèses explicatives, l'ordinateur pourrait différencier différents groupes, qu'on pourrait ensuite associer par exemple à leur provenance géographique, génétique, à l'alcoolisme ou à l'exposition à un métal lourd ou à une toxine telle que l'aflatoxine. [22]

A partir de l'échantillon d'apprentissage $S = \{(x_i)\}_{i,m}$ non étiqueté, on cherche des régularités sous-jacente [23]

_ Sous forme d'une fonction

_ Sous forme d'un modèle complexe

Afin de résumer, détecter des régularités, comprendre... $\{(x_i)\}_{i,m}$

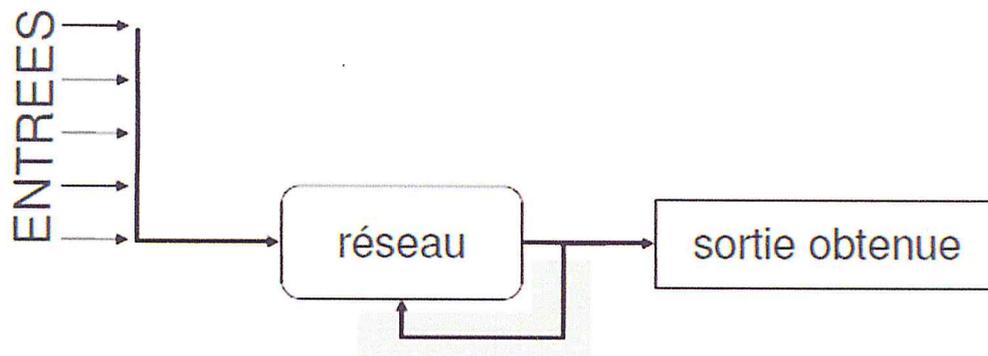


Figure12 : l'apprentissage non supervisé. [23]

2.3.3. L'apprentissage semi-supervisé [22]

Effectué de manière probabiliste ou non, il vise à faire apparaître la distribution sous-jacente des « *exemples* » dans leur espace de description. Il est mis en œuvre quand des données (ou « *étiquettes* ») manquent... Le modèle doit utiliser des exemples *non-étiquetés* pouvant néanmoins renseigner.

Exemple : En médecine, il peut constituer une aide au diagnostic ou au choix des moyens les moins onéreux de tests de diagnostics.

2.3.4 L'apprentissage partiellement supervisé (probabiliste ou non)

[22]

Quand l'étiquetage des données est partiel. C'est le cas quand un modèle énonce qu'une donnée n'appartient pas à une classe *A*, mais peut-être à une classe *B* ou *C*.

Exemple dans le domaine médical : un médecin est certain que son patient n'est pas atteint de la maladie *C* mais hésite entre la maladie *A* ou la maladie *B*.

2.3.5 L'apprentissage par renforcement [22]

L'algorithme apprend un comportement étant donné une observation. L'action de l'algorithme sur l'environnement produit une valeur de retour qui guide l'algorithme d'apprentissage.

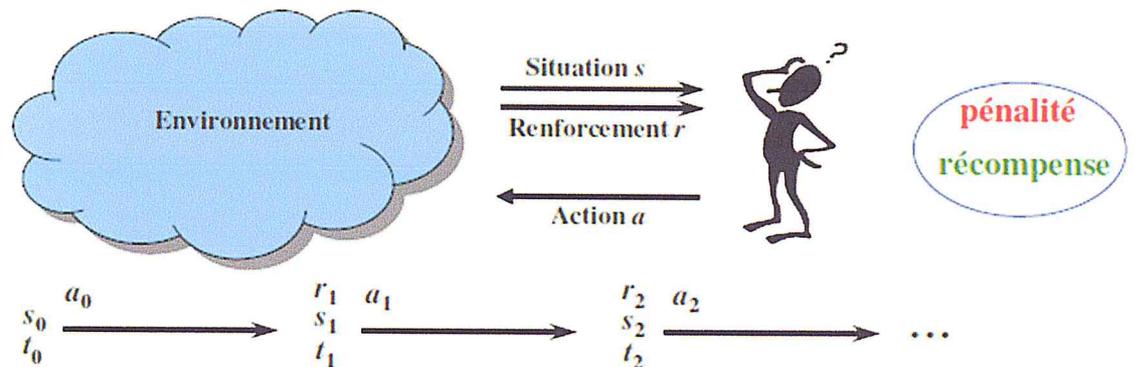


Figure 13 : L'apprentissage par renforcement. [23]

- _ L'agent apprend à se rapprocher d'une **stratégie comportementale optimale** par des interactions répétitives avec l'environnement [23]
- _ Les décisions sont prises séquentiellement à des intervalles de temps discrets [23]

2.4. Facteurs de pertinence et d'efficacité [22]

La qualité de l'apprentissage et de l'analyse dépend du besoin en amont et a priori compétence de l'opérateur pour préparer l'analyse. Elle dépend aussi de la complexité du modèle (spécifique ou généraliste) et de son adaptation au sujet à traiter. Enfin, la qualité du travail dépendra aussi du mode (de mise en évidence visuelle) des résultats pour l'utilisateur final (un résultat pertinent pourrait être caché dans un schéma trop complexe, ou mal mis en évidence par une représentation graphique inappropriée). Avant cela, la qualité du travail dépendra de facteurs initiaux contraignants, liées à la base de données :

1. **Nombre d'exemples** : moins il y en a plus l'analyse est difficile, mais plus il y en a plus le besoin de mémoire informatique est élevé et plus longue est l'analyse.
2. **Nombre et qualité des attributs** : décrivant ces exemples (La distance entre deux "exemples" numériques (prix, taille, poids, intensité lumineuse, intensité de bruit,

etc.) est facile à établir, celle entre deux attributs catégoriels (couleur, utilité, est plus délicate)

3. **Pourcentage de données renseignées** et manquantes

4. « **Bruit** » ; Le nombre et la « *localisation* » des valeurs douteuses (erreurs) ou naturellement non conformes au modèle de distribution générale des « *exemples* » sur leur espace de distribution.

3. les techniques de l'apprentissage automatique

3.1. Concept de l'analyse prédictive [20]

L'analyse prédictive utilise des formules mathématiques, appelées algorithmes, qui analysent les données historiques ou actuelles afin d'identifier des schémas ou des tendances permettant de prévoir des événements futurs. [20]

Les techniques dites de modélisation prédictive (prédictive modelling ou prédictive data mining) analysent un ensemble de données et en extraient un ensemble de règles, formant un modèle prédictif, afin d'essayer de prédire, avec la plus grande précision, le comportement de nouvelles données. Les techniques dans le domaine de la prédiction sont les plus classiques, car on peut directement les utiliser dans des applications concrètes. [19]

L'apprentissage automatique se base sur des modèles prédictifs qui apprennent à partir de données existantes afin de prévoir les résultats et comportements futurs.

Voici quelques algorithmes les plus populaires dans l'apprentissage automatique et notamment ceux qui sont utilisés dans la prédiction des données :

3.1.1. Algorithme des K plus proches voisins [19]

Le classifieur des k plus proches voisins ou *k-ppv* (k-Nearest Neighbor ou k-NN, en anglais) est l'un des algorithmes de classification les plus simples.

a) Principe de l'algorithme KNN :

Un exemple est *classifié par vote majoritaire* de ses k "voisins" (par mesure de distance), c'est-à-dire qu'il est prédit de classe **C** si la classe la plus représentée parmi ses k voisins est la classe **C**.

Un cas particulier est le cas où $k = 1$, l'exemple est alors affecté à la classe de son plus proche voisin.

L'opérateur de distance le plus souvent utilisé est la **distance Euclidienne**, cependant, en fonction du problème, on peut encore utiliser les distances de *Hamming*, de *Mahalanobis*, etc.

Remarque :

- Le choix du k est très important pour la classification.
- Il est utile de choisir des *valeurs paires* de k , pour éviter les cas d'égalité.

Exemple :

- Sur la *figure 14*, on peut voir l'effet du choix de k sur le résultat de la classification. En effet, si $k=1$; 2 ; 3 l'exemple à prédire (noté "?") serait classifié comme étant de la classe "X", mais si $k=5$, il serait classifié comme étant de la classe "O".

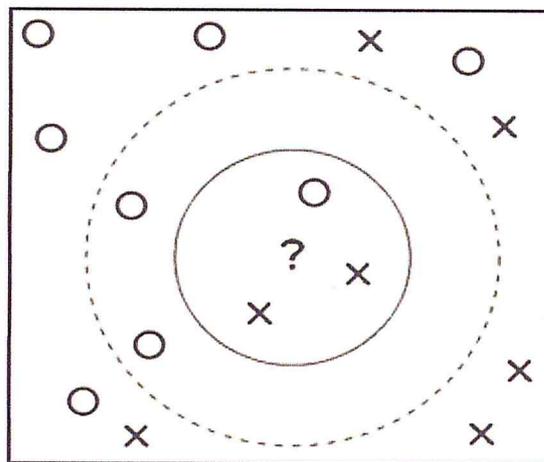


Figure 14 : Schéma d'une classification par la méthode k-NN.

b) Description formelle de l'algorithme KNN :

L'algorithme des K plus proches voisins est donné par :

Déclarations

- M : nombre de classes d'apprentissage $C = \{c_1, \dots, c_M\}$;
- N : nombre d'exemples d'entraînement $E = \{e_1, \dots, e_N\}$;
- $Ent = \{(e_i, c_k)\}$, : ensemble d'apprentissage formé par les couples (e_i, c_k) ; /* e_i est l'exemple d'apprentissage et c_k sa classe d'appartenance */
- ex : exemple test /* dont on cherche la classe d'appartenance */

Début

< On cherche à classer e_x ? > ;

Pour Chaque exemple (e_i, w) *Ent Faire*

< Calculer la distance $D(e_i, e_x)$ entre e_i et e_x > ;

FPour

< Trier les échantillons e_i par ordre croissant des distances > ;

Pour les k plus proches e_i de e_x (les k premières – ayant les plus petites- $D(e_i, e_x)$)

Faire

< Compter le nombre d'occurrences de chaque classe > ;

FPour

< Attribuer à e_x la classe c_j la plus fréquente > ; /*Celle qui apparait le plus souvent*/

Fin.

3.1.2. Réseaux de Neurones [19]

Les réseaux de neurones artificiels sont construits à partir du modèle biologique des neurones. Leurs fondements proviennent des premiers travaux sur les réseaux de neurones qui conçurent un modèle simplifié du neurone biologique, appelé le neurone formel. Ce dernier peut être représenté comme un automate à n entrées et une sortie.

Le premier modèle de réseau de neurones complet a été nommé Perceptron (*figure 15*). Dans ce modèle, une fonction de combinaison associe les entrées du neurone (par exemple en les sommant), puis compare ce résultat à une valeur seuil, via une fonction d'activation.

Typiquement, la fonction d'activation renvoie une valeur binaire, relativement au fait qu'elle ait été activée ou pas.

Les paramètres importants de ce modèle sont les coefficients synaptiques (c'est-à-dire les poids), le seuil, et la façon de les ajuster lors de l'apprentissage.

En effet, il faut choisir un mécanisme permettant de calculer ces paramètres et de les faire converger vers une valeur assurant une classification aussi proche que possible de l'optimale, lors de la phase d'apprentissage.

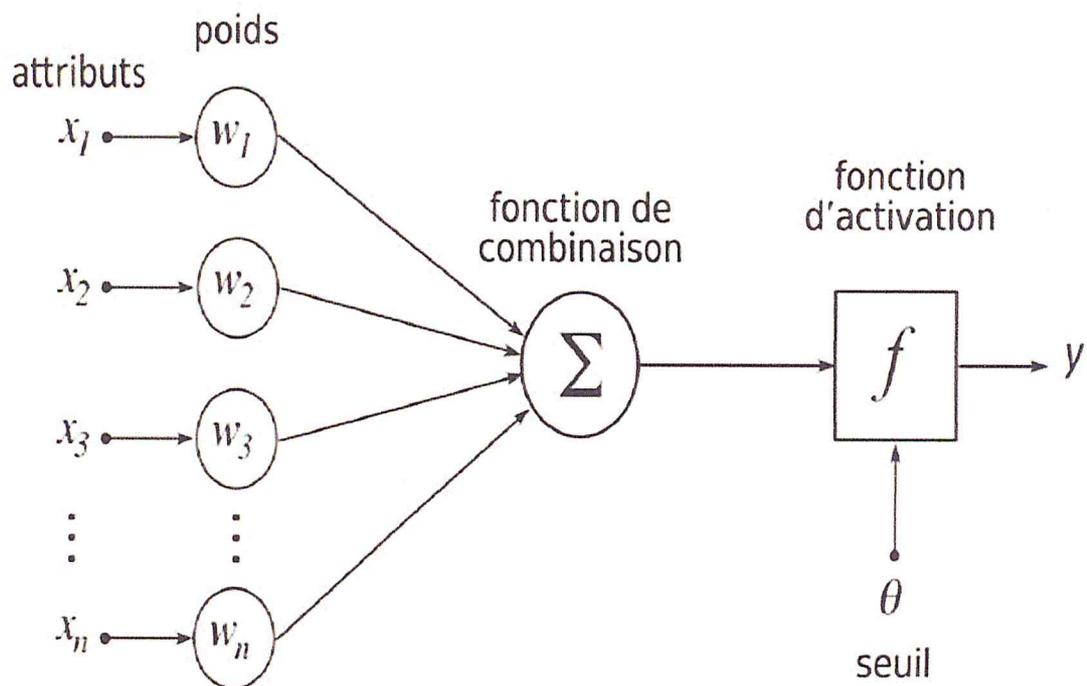


Figure 15 : Schéma du perceptron.

On peut représenter la sortie du réseau par l'équation suivante (avec f la fonction d'activation) :

$$y = f\left(\sum_{i=1}^n (w_i x_i) - \theta\right)$$

Remarque : Dans sa première version, le neurone formel était implémenté avec une fonction à seuil (pour la fonction d'activation), mais de nombreuses versions existent (fonction linéaire par morceaux, sigmoïde, Gaussienne, etc.), pouvant ainsi fournir plusieurs valeurs de sortie.

Amélioration : Le modèle du perceptron multicouche (MLP, Multi-Layer Perceptron, en anglais), reprend le principe du perceptron, en ajoutant plusieurs vecteurs de poids (appelées couches cachées), afin d'augmenter les combinaisons possibles.

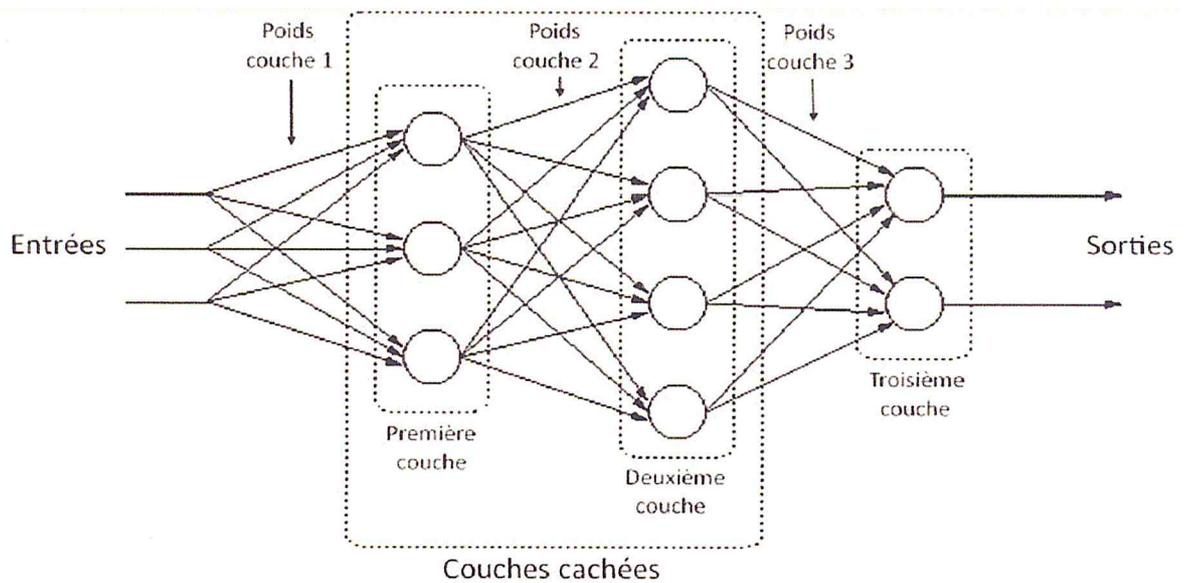


Figure 16 : Schéma du perceptron multi couches.

Les neurones d'une couche sont reliés aux neurones des couches adjacentes par des liaisons pondérées. Ainsi, le poids de chacune de ces liaisons est l'élément clef du fonctionnement du réseau.

Remarque : On peut considérer que la fonction d'activation est la même à chaque couche, mais qu'elle varie d'une couche à l'autre.

3.1.3. Classifieur Naïf Bayésien

La méthode de classification naïve bayésienne est un algorithme d'apprentissage supervisé (supervised machine Learning) qui permet de classer un ensemble d'observations selon des règles déterminées par l'algorithme lui-même. Cet outil de classification doit dans un premier temps être entraîné sur un jeu de données d'apprentissage qui montre la classe attendue en fonction des entrées. Pendant la phase d'apprentissage, l'algorithme élabore ses règles de classification sur ce jeu de données, pour les appliquer dans un second temps à la classification d'un jeu de données de prédiction. Le classificateur bayésien naïf implique que les classes du jeu de données d'apprentissage soient connues et fournit, d'où le caractère supervisé de l'outil. [21]

La classification naïve bayésienne repose sur l'hypothèse que les attributs sont fortement (ou naïvement) indépendants. Elle est basée sur le théorème de Bayes qui ne s'applique que sous cette hypothèse. [19]

Théorème de Bayes est donné par : $P(x/y) = \frac{P(y/x)P(x)}{P(y)}$, avec $P(x/y)$ est la probabilité conditionnelle d'un événement x sachant qu'un autre événement y de probabilité non nulle s'est réalisé. [19]

Dans le cas d'une classification, on pose H l'hypothèse selon laquelle un "vecteur d'attributs X (représentant un objet) appartient à une classe C ", et l'on suppose que l'on cherche à estimer la probabilité $P(H/X)$, c'est-à-dire la probabilité que l'hypothèse H soit vraie, considérant X . [19]

Ainsi, si l'on a un nouvel exemple $x = (x_1, x_2, \dots, x_n)$ dont on veut trouver la classe, on va chercher la probabilité maximale d'appartenance à cette classe :

$$P(x)^* = \operatorname{argmax} P(x_1, x_2, \dots, x_n / H) * P(H)$$

Cette équation est directement déduite du théorème de Bayes. En effet, comme l'objectif est de faire une maximisation et que le dénominateur ne dépend pas de x , on peut le supprimer. [19]

Cependant, cette probabilité pourrait être beaucoup trop compliquée à estimer, si l'on considère le nombre de descriptions possibles. C'est alors que l'on utilise l'hypothèse d'indépendance, qui nous permet de décomposer la probabilité conditionnelle en un produit de probabilités conditionnelles. Le classifieur devient alors : [19]

$$\text{Class}(x) = \operatorname{arg max} \prod_{i=1}^n P(x_i/H) * P(H)$$

3.1.4. Machine à Vecteurs de Support

Cette technique tente de séparer linéairement les exemples positifs des exemples négatifs dans l'ensemble des exemples. Chaque exemple doit être représenté par un vecteur de dimension n . [22]

Idée : En effet, l'idée principale des SVM est de reconsidérer le problème dans un espace de dimension supérieure, éventuellement de dimension infinie. Dans ce nouvel espace, il est alors probable qu'il existe un hyperplan séparateur linéaire. Si c'est le cas, les SVM cherchent parmi l'infinité des hyperplans séparateurs celui qui maximise la marge entre les classes (*figure 17*). [19]

On applique au vecteur d'entrée x une transformation non-linéaire Φ , afin de décrire les données dans un autre espace. L'espace d'arrivée $\Phi(x)$ est appelé **espace de redescription**. On cherche alors l'**hyperplan séparateur optimal**, d'équation $\Phi(x) = \alpha \Phi(x) + \beta$, dans l'espace de redescription. [19]

Procédé : Pour cela, on cherche *les équations des hyperplans parallèles* qui passent par **les vecteurs supports**, c'est-à-dire les attributs les plus proches de la frontière interclasse. On en déduit l'équation de l'hyperplan optimal, équidistant de ces hyperplans. [19]

En pratique, on ne connaît pas la transformation Φ , on construit donc plutôt directement la fonction h , appelée *fonction noyau*. Le *théorème de Mercer* explicite les conditions que h doit satisfaire pour être une fonction noyau : elle doit être *symétrique et semi-définie positive*.

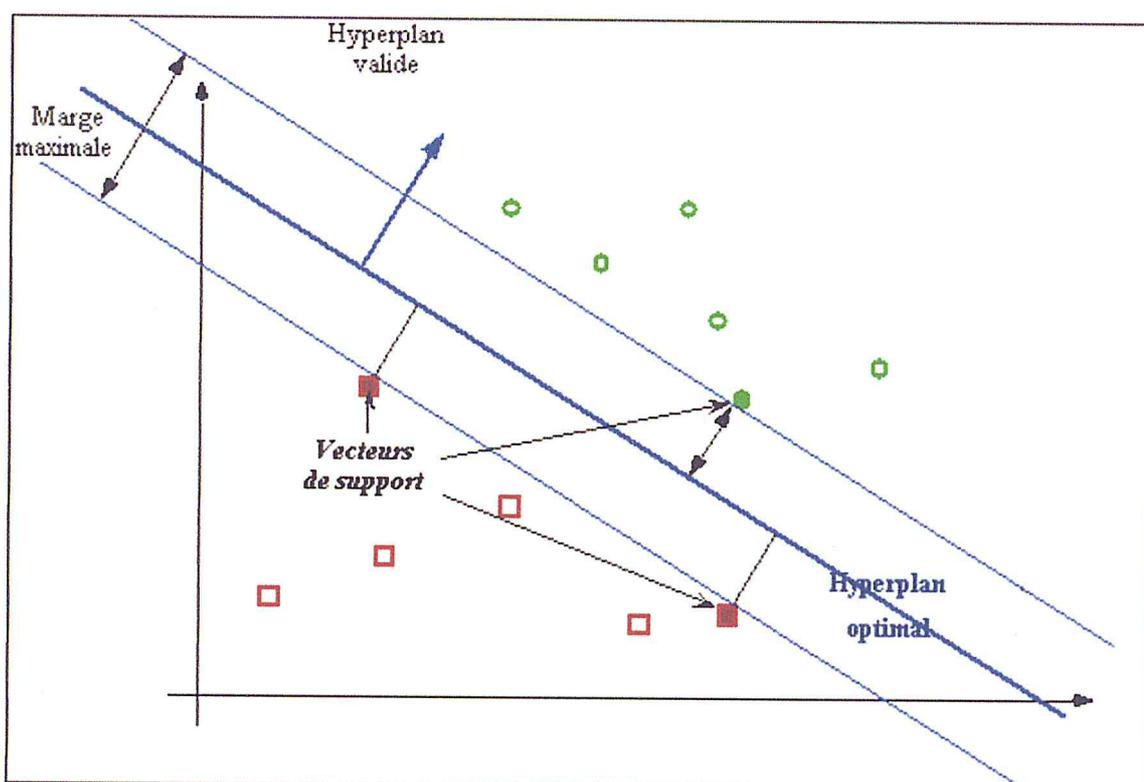


Figure 17 : Principe du Séparateur à Vaste Marge (SVM). [19]

L'intérêt des SVM est la sélection de Vecteurs Supports qui représentent les vecteurs discriminant grâce auxquels est déterminé l'hyperplan. Les exemples utilisés lors de la recherche de l'hyperplan ne sont alors plus utiles et seuls ces vecteurs supports sont utilisés pour classer un nouveau cas. [19]

Remarque. Lorsque l'on utilise les SVM, le *premier paramétrage* consiste donc à *choisir une fonction noyau*, parmi de nombreuses candidates : *Linéaire, polynomial, gaussien*, etc. [19]

3.1.5. Arbres de Décision

Les arbres de décision sont un outil très populaire de classification. Leur principe repose sur la construction d'un arbre de taille limitée. La racine constitue le point de départ de l'arbre et représente l'ensemble des données d'apprentissage. Puis ces données sont segmentées en *plusieurs sous-groupes*, en fonction d'une variable discriminante (un des attributs). [19]

3.1.5.1. Exemple

Sur la *Figure 18*, la première *variable discriminante* est la température corporelle. Elle divise la population en deux sous-groupes : les personnes dont la température est supérieure à 37°C et les autres. Le processus est ensuite réitéré au deuxième niveau de l'arbre, où les sous-populations sont segmentées à leur tour en fonction d'une autre *valeur discriminante* (dans l'exemple, c'est la toux).

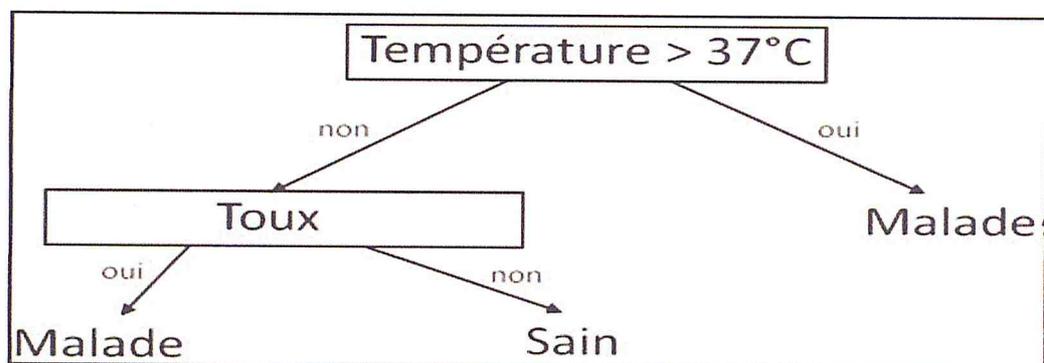


Figure 18 : Schéma d'un arbre de décision. [19]

Une fois l'arbre construit à partir des données d'apprentissage, on peut prédire un nouveau cas en le faisant descendre le long de l'arbre, jusqu'à une feuille. Comme la feuille correspond à une classe, l'exemple sera prédit comme faisant partie de cette classe.

Dans l'exemple, on peut en déduire qu'une personne qui a une température $< 37^{\circ}\text{C}$ et qui a de la toux est prédite comme *malade*, tandis qu'une personne qui a une température $< 37^{\circ}\text{C}$ mais pas de toux est considérée comme *saine*.

Lors de la création de l'arbre, la première question qui vient à l'esprit est le choix de la variable de segmentation sur un sommet. Pourquoi par exemple avons-nous choisi la variable "température" à la racine de l'arbre ? Il nous faut donc une mesure afin

d'évaluer la qualité d'une segmentation et sélectionner la meilleure variable sur chaque sommet.

L'intérêt des arbres de décision est en premier lieu leur lisibilité. En effet, il est très simple de comprendre les décisions de l'arbre une fois celui-ci créé, ce qui n'est pas toujours le cas pour les autres classifieurs que nous verrons. D'autre part, l'algorithme de création des arbres de décision fait automatiquement la sélection d'attributs jugés pertinents, et ce, même sur des volumes de données importants. [19]

3.1.5.2. Idée et Propriétés Générales [19]

Diviser récursivement et le plus efficacement possible les individus de l'ensemble d'apprentissage par des tests définis à l'aide des variables jusqu'à ce que l'on obtienne des sous-ensembles d'individus ne contenant (presque) que des exemples appartenant à une même classe ! A la base, trois opérations seront nécessaires :

1. **Décider si un nœud est terminal** (tous les individus sont dans la même classe).
2. **Sélectionner un test à associer à un nœud** (utiliser critères statistiques).
3. **Affecter une classe à une feuille** (nœud terminal) - la classe majoritaire !

3.1.5.3. Exemple Introductif [19]

Décider si un patient est **malade** ou **bien-portant (sain)** selon sa **température** et s'il a la **gorge irritée**. A partir des exemples (des patients), des attributs ou variables (Température et Gorge-irritée), des classes (malade ou sain), construire l'arbre de décision qui aura :

- 2 classes (malade et bien-portant).
- 2 variables (température et gorge-irritée).

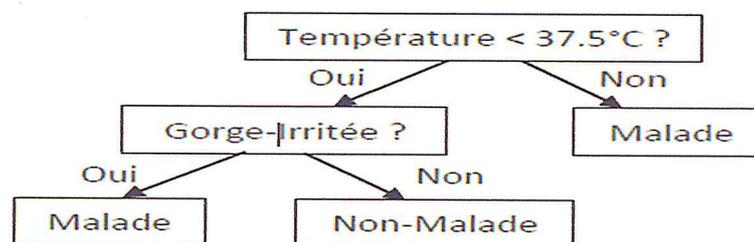


Figure19 : Schéma d'un arbre de décision.

3.1.5.4 Classification et Règles [19]

L'arbre de Décision (AD) permet de classer un nouvel exemple : (37.2, oui), c'est-à-dire, Température=37.2 et Gorge-Irritée=oui, comme appartenant à la classe **malade**.

L'AD peut être traduit en un système de règles ; lesquelles pouvant être considérées comme le pseudocode ou l'algorithme de l'AD :

- Si (Temp. < 37.5) et (Gorge-irritée) Alors malade.
- Si (Temp. < 37.5) et Non (Gorge-irritée) Alors Sain.
- Si (Temp. ≥ 37.5) Alors malade.

3.1.5.5. Définition du Formalisme : Arbres de Décision

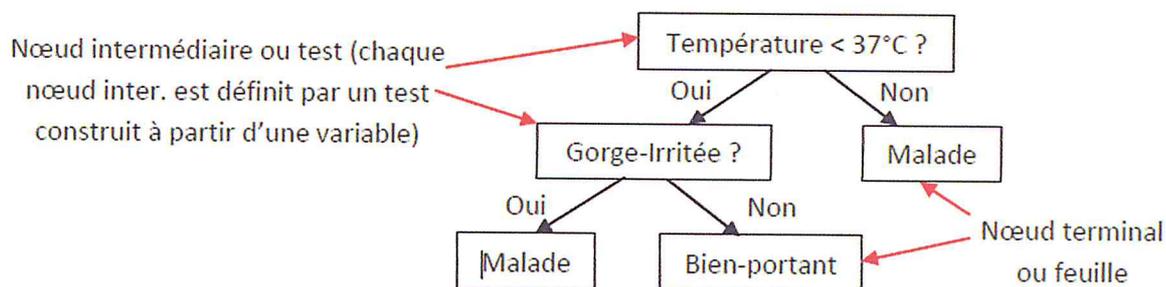


Figure 20 : formalisme d'un arbre de décision. [19]

3.1.6. Régression logistique (SGV) [17]

Son domaine de prédilection est celui de la prédiction de probabilité d'apparition d'évènements comme par exemple dans le domaine de la fraude bancaire ou encore dans celui de la publicité pour le ciblage de clients fidélisables.

3.1.7. Forêts d'arbres décisionnels (Random Forest) [17]

C'est une application de graphe en arbres de décision permettant ainsi la modélisation de chaque résultat sur une branche en fonction de choix précédents. On prend ensuite la meilleure décision en fonction des résultats qui suivront. On peut considérer cela comme une forme d'anticipation.

3.1.8. Le Boosting [17]

Il s'agit d'une méthode de classification émettant des hypothèses qui sont au départ de moindre importance. Plus une hypothèse est vérifiée, et plus son indice de confiance augmente et donc elle prend plus d'importance dans la classification.

3.1.9. K_means : [25]

L'algorithme k-means mis au point par McQueen en 1967, un des plus simples algorithmes d'apprentissage non supervisé, appelée algorithme des centres mobiles, il attribue chaque point dans un cluster dont le centre (centroïde) est le plus proche. Le centre est la moyenne de tous les points dans le cluster, ses coordonnées sont la moyenne arithmétique pour chaque dimension séparément de tous les points dans le cluster c'est-à-dire chaque cluster est représentée par son centre de gravité.

*Ces méthodes sont souvent combinées pour obtenir diverses variantes d'apprentissage.

L'utilisation de tel ou tel algorithme dépend fortement de la tâche à résoudre (classification, estimation de valeurs, etc.). [17]

3.2. Utilisations et applications des algorithmes

Dans le but de prendre une idée générale sur l'algorithme qui peut être appliqué dans notre application nous avons créé le tableau comparatif entre différents algorithmes suivant :

| Algorithme | Utilisations | Applications |
|-----------------------|---|--|
| Régression logistique | -outil pour la classification à deux classes et à classes multiples. | <p>-En médecine, elle permet par exemple de trouver les facteurs qui caractérisent un groupe de sujets malades par rapport à des sujets sains.</p> <p>-Dans le domaine des assurances, elle permet de cibler une fraction de la clientèle qui sera sensible à une police d'assurance sur tel ou tel risque particulier.</p> <p>-Dans le domaine bancaire, pour détecter les groupes à risque lors de la souscription d'un crédit.</p> <p>-En économétrie, pour expliquer une variable discrète. Par exemple, les intentions de vote aux élections.</p> |
| Réseaux de Neurones | <p>-couvre les problèmes de classes multiples, à deux classes et de régression.</p> <p>-généralement utilisés dans des problèmes de nature statistique.</p> | <p>- pour la classification d'espèces animales par espèce étant donnée une analyse ADN.</p> <p>- reconnaissance de motif ; par exemple pour la reconnaissance optique de</p> |

| | | |
|-------------------------------|--|--|
| | | <p>caractères (OCR), et notamment par les banques pour vérifier le montant des chèques, par La Poste pour trier le courrier en fonction du code postal, etc. ; ou bien encore pour le déplacement automatisé de robots mobiles autonomes.</p> <ul style="list-style-type: none"> -estimations boursières. -en météorologie. -en auscultation des ouvrages hydrauliques. |
| Machine à Vecteurs de Support | -Les SVM peuvent être utilisés pour résoudre des problèmes de discrimination, c'est-à-dire décider à quelle classe appartient un échantillon, ou de régression, c'est-à-dire prédire la valeur numérique d'une variable. | -bio-informatique, recherche d'information, vision par ordinateur, finance...etc. |
| K-Means Clustering | -K-Means est utile lorsque on a une idée du nombre de partitions existantes dans notre espace. | -Il est utilisé notamment dans la bibliothèque OpenCV précédemment mentionnée vis-à-vis du domaine de la « computer vision ». Dans Mahout, on l'utilise typiquement afin de déterminer quelle suggestion |

| | | |
|---------------------------|---|---|
| | | peut être faite à un utilisateur en fonction des préférences que d'autres utilisateurs similaires ont eues (utilisateurs du même groupe). |
| Classifieur Naïf Bayésien | -C'est un type de classification Bayésienne probabiliste simple basée sur le théorème de Bayes avec une forte indépendance (dite naïve) des hypothèses. | -la classification de documents et l'élaboration de filtres anti-spam. |
| Les arbres de décision | -C'est un algorithme de classification et de régression utilisé pour la modélisation prédictive d'attributs discrets et continus. | -la sécurité, la fouille de données, la médecine, etc. |
| Boosting | -Les algorithmes de boosting sont l'un des algorithmes les plus utilisés dans les compétitions de science de données -Ce sont des méthodes de modélisation d'ensemble pour | -Classification d'un courrier électronique comme SPAM ou non. |

| | | |
|--|---------------------------------------|--|
| | améliorer la puissance de prédiction. | |
|--|---------------------------------------|--|

Tableau 1 : Comparaison entre quelques algorithmes.

A partir de ce tableau on peut voir que les arbres de décision sont l'un des méthodes les plus utilisée dans la médecine et qui sont utilisé pour la modélisation prédictive de tout type de données discrets et continus ce qui fait que les arbres de décision peuvent être utilisés comme algorithme de prédiction dans notre solution.

4. Conclusion

Dans ce présent chapitre, nous avons présenté les différentes méthodes de prédiction en apprentissage automatique et les différents domaines où ils peuvent intervenir. Le choix d'un tel algorithme à appliquer dans un problème dépend à la nature des donnés et selon la problématique posée. Le choix final d'un tel algorithme à utiliser va être décidé dans la conception.

Dans le chapitre suivant, nous allons définir le concept du cloud qui sera une plateforme de test des différents algorithmes qui sont déjà implémentés sur ML Azure à partir de nos données médicales.

CHAPITRE III :

Le Cloud Computing

1. Introduction

Dans ce chapitre, nous allons présenter les notions et les concepts fondamentaux du Cloud Computing.

La technologie de l'internet se développe plus rapidement de manière progressive. Dans ces dernières années il y a un nouveau concept, Son but est d'améliorer les services dans le domaine TIC (Technologies de l'information et de la communication), il s'agit du « Cloud

Computing ». Ce dernier est un concept qui regroupe plusieurs technologies servant à délivrer différents services. Il peut être schématisé par un ensemble indéterminé de ressources informatiques interconnectées et une coopération dans un réseau bien défini.

L'accès aux services se fait à la demande par une application standard facilement disponible à savoir, un navigateur Web.

2. Définition [26]

NIST (National Institute of Standards and Technology) l'Institut national des normes et de La technologie, défini le Cloud Computing comme suit : « Le Cloud Computing est l'ensemble des disciplines, pratiques, technologies et modèles commerciaux utilisés pour délivrer comme un service à la demande et par le réseau des capacités informatiques (logiciels, plateformes, matériels) ».

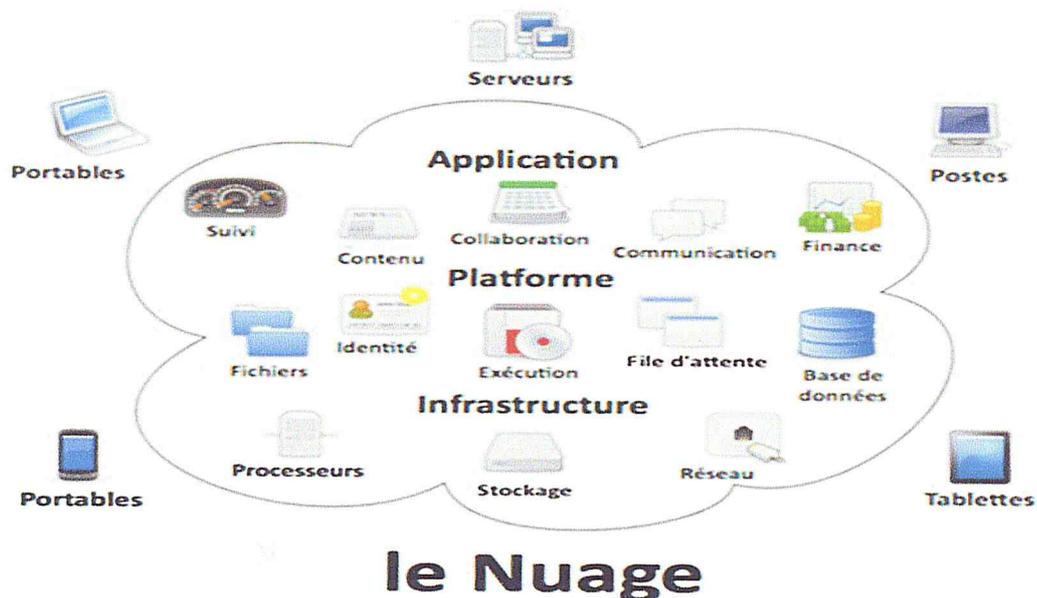


Figure 21 : Les différents composants de Cloud Computing. [26]

-Un service web s'agit d'une technologie permettant à des applications de dialoguer à distance via Internet, et ceci indépendamment des plates-formes et des langages sur lesquelles elles reposent. Pour ce faire, les services Web s'appuient sur un ensemble de protocoles Internet très répandus (XML, HTTP), afin de communiquer. Cette communication est basée sur le principe de demandes et réponses, effectuées avec des messages XML. [27]

Les services web sont décrits par des documents WSDL (Web Service Description Language), qui précisent les méthodes pouvant être invoquées, leurs signatures et les points d'accès du service (URL, port.). Les services Web sont accessibles via SOAP, la requête et les réponses sont des messages XML transportés sur HTTP. [27]

3. Les Caractéristique de Cloud Computing [26]

- ✓ **Elasticité** : pour les utilisateurs de Cloud Computing a des capacités illimitées, qui peuvent être augmentées ou réduites selon l'usage.
- ✓ **pay-as-you-use** : les clients du Cloud vont donc recevoir des tarifications, les plus précises possibles, pour les ressources qu'ils utilisent. Par exemple la tarification d'Amazon, pour le micro machine virtuelle à 0.02\$ par heure et 3.10\$ par heure pour les grosses machines virtuelles.
- ✓ **Self-service (à la demande)** : capacité à fournir une ressource informatique automatiquement, sans requérir d'interaction humaine côté fournisseur.
- ✓ **Mesure de la qualité de services** : évaluer et garantir un niveau de performance et de disponibilité adapté aux besoins spécifiques des clients.
- ✓ **Accès réseau universel** : L'accès aux ressources est très rapide via l'internet.
- ✓ **Mise en commun de ressources** : Datacenter fournissant les ressources (machines, stockage, etc.) pour les différents clients en mode partagé.
- ✓ **Multi-tenancy** : La capacité de fournir un service simultanément à plusieurs clients, cela permet également d'augmenter l'utilisation des ressources IT déployée.
- ✓ **Disponibilité** : La haute disponibilité de la plate-forme est obtenue par redondance et par la capacité de se remettre rapidement en cas des problèmes.

- ✓ **Green IT** : Green IT est un concept qui désigne l'ensemble des nouvelles technologies à faible impact environnemental. Le Cloud Computing, basé sur la virtualisation, qu'il réduise les risques des machines physiques sur l'environnement et la consommation énergétique.

4. Les services de Cloud Computing

Le Cloud Computing propose trois types de services principaux :

4.1. IaaS : Infrastructure as a service [26]

IaaS (Infrastructure as a service) en Français "L'infrastructure en tant que service", est un service principal dans le Cloud, Le service fourni ici sont les différentes composants informatiques comme des espaces de stockages, équipements réseaux, des unités centrales, etc.

Les utilisateurs peuvent accéder à ces services à la demande via l'internet sans restriction, comme un travail local.

Les services IaaS sont situés dans :

- ✓ Centres des données ; parmi ses avantages sont : Éviter d'achat et la gestion de matériel,
- ✓ Offre une grande flexibilité,
- ✓ La possibilité d'installer tout type de logiciel,
- ✓ Contrôle total des systèmes (administration à distante par SSH ou RDP "Remote

Desktop").

Parmi les prestataires d'IaaS, on peut citer : Amazon avec EC2 qui est le roi de ce genre d'infrastructure.

4.2. PaaS: Platform as a service [26]

PaaS (Platform as a service) en Français " plate-forme en tant que service", est un modèle composé de tous les éléments et les services nécessaires pour faciliter les développements des applications où PaaS prépare des environnements spécialisés pour aide les utilisateurs dans la construction, la livraison, l'extension de leurs projets. L'utilisateur loue une plateforme sur laquelle il peut développer, tester et exécuter ses applications, parce que PaaS évite d'acheter et d'installer des logiciels,

aussi il ne gère ni ne contrôle l'infrastructure sous-jacente, mais contrôle les applications déployées.

Parmi ses avantages :

- _ Facilités à gérer des développements des applications,
- _ Le déploiement est automatisé,
- _ Ne nécessite pas une infrastructure.

Il y a plusieurs fournisseurs des services PaaS, les principaux sont : Microsoft avec AZURE,

Google avec Google App Engine et Orange Business Services.

4.3. SaaS: Software as a Service [26]

SaaS (Software as a Service) en Français " L'application en tant que service" est le modèle le plus utilisé dans le monde après le service d'email, est un modèle de distribution des logiciels et les applications qu'ils sont hébergées dans des centres de données, qu'il donne la possibilité des clients pour consommation ces application à la demande via l'internet avec une facturation à l'usage réel.

Ces applications, prêts à l'emploi, et ne nécessitent pas de maintenance, d'installation de logiciel, le maintenir et le mettre à jour, toutes ces opérations sont effectuées par le fournisseur d'application, dans SaaS l'utilisation d'application reste transparente pour les utilisateurs, qui ne se soucient ni de la plateforme, ni du matériel.

Parmi ses avantages :

- _ Pas d'installation,
- _ Pas de mise à jour,
- _ Plus de License,
- _ Paiement à l'usage,
- _ Facilité de faire le test de nouveaux logiciels.

Les prestataires de solutions SaaS les plus connus sont : Google avec Google Drive et Microsoft Online Services qui intègre des produits de communication et de collaboration (Exchange, SharePoint, Live Meeting, Communications Server), HipChat, Trello, Intercom, HootSuite, Buffer ,Microsoft Azureetc.

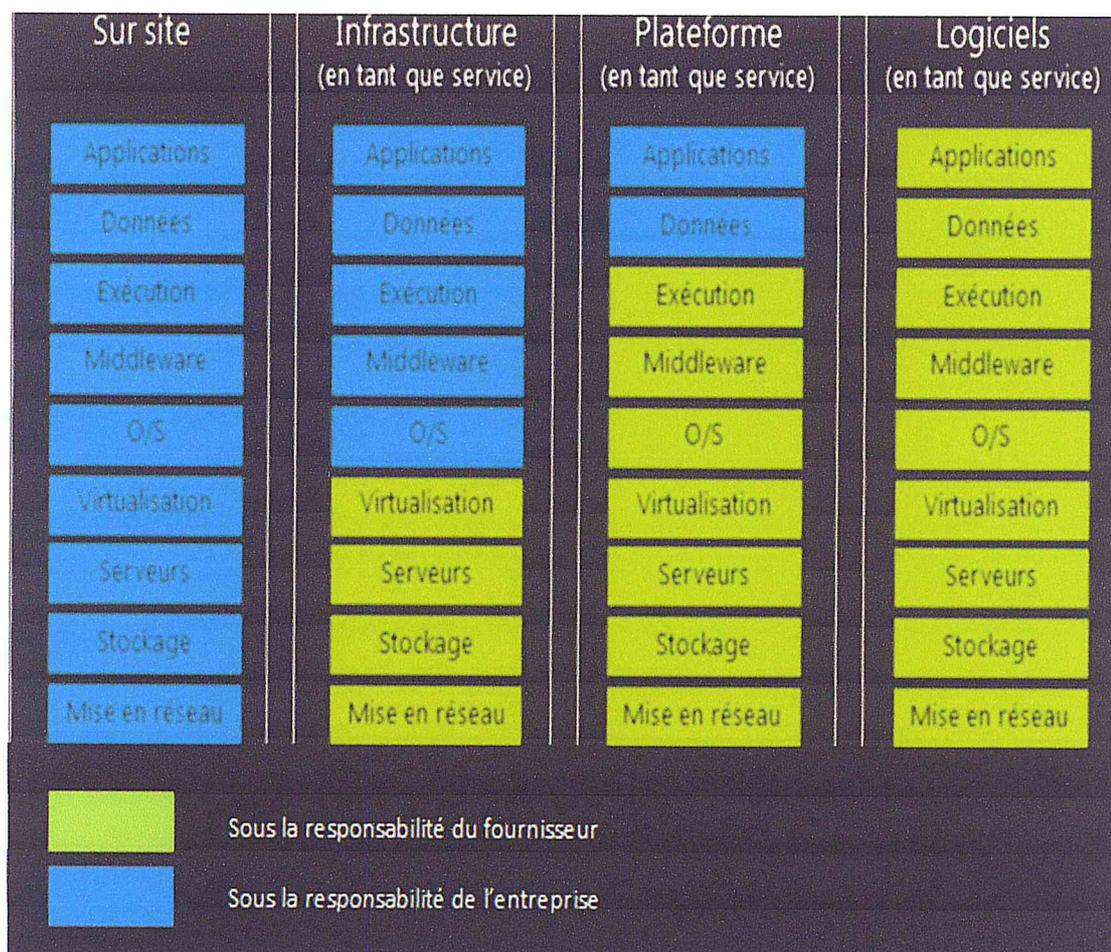


Figure 22 : Les services du Cloud Computing. [26]

5. Les Modèles de déploiement de Cloud Computing

On peut distinguer quatre type principaux dans le Cloud sont : le Cloud privé, le Cloud public, le Cloud hybride (Mixte) et le Cloud communautaire, nous allons expliquer chaque type de manier plus détaillé.

5.1. Cloud privé

En anglais " Private Cloud " est un ensemble des services et des ressources disponible à un seul client par exemple une entreprise, le Cloud privé peut être géré par l'entreprise elle-même, ou bien avec sa branches, dans ce cas il s'appelle "Le Cloud privé Interne", en d'autre façons il peut être géré par un prestataire externe au bénéfice de l'entreprise, dans ce cas s'appelle "Le Cloud privé Externe ", qu'il accessible via des réseaux sécurisés de type VPN (Virtual Private Network), par exemple Amazon Virtual Private Cloud (Amazon VPC). [26]

5.2. Cloud public

En anglais "Public Cloud" est un ensemble des services et des ressources accessibles par Internet et géré par un prestataire externe, ces ressources et services sont partagés entre plusieurs clients, qu'ils les utilisent à la demande et à tout moment sans savoir où elles existent, aussi ces services peuvent être gratuits ou payants.

Puisque les services sont payants, il existe des contrats SLA (Service Level Agreement) entre les clients et les fournisseurs, SLA est un document qui définit la qualité de service requise entre les deux. Quelques exemples de Cloud public : Amazon Elastic Compute Cloud (EC2) et Amazon S3, Sun Cloud, IBM's Blue Cloud, Google AppEngine et Windows Azure Services Platform. [26]

5.3. Cloud communautaire :

En anglais " Cloud Community" est un ensemble des infrastructures, qu'ils permettent à plusieurs clients ou organisations de partager les différentes ressources, ces ressources sont généralement spécifiques à des organisations, ce type de Cloud peut être géré par ces organisation elles-mêmes ou par des fournisseurs externes.

Aussi il permet à un groupe des utilisateurs à créer leur propre Cloud avec des caractéristiques de Cloud privé tel que la sécurité, ressources dédiées et un coût réduit. [26]

5.4. Cloud hybride :

En anglais " Hybrid Cloud" est une composition de deux ou trois types de Cloud (Privé, communautaire et public).

Par exemple l'utilisation des applications dans un Cloud public mais ces applications nécessitent des données stockées sur un Cloud privé. [26]

6. Les principaux fournisseurs de Cloud Computing

- Amazon Web Services (IaaS).
- Amazon Web Services (PaaS).
- Google AppEngine (PaaS).
- force.com (PaaS).
- salesforce.com (SaaS).
- Google Apps (SaaS).
- Et le cloud Microsoft Azure (PaaS et IaaS). [26]

Dans notre étude on se base sur le cloud Microsoft Azure comme plateforme de déploiement de notre solution.

7. Microsoft Azure

7.1 Définition [30]

Microsoft Azure (Windows Azure jusqu'en 2014) est la plate-forme applicative en nuage de Microsoft. Son nom évoque le « cloud computing », ou informatique en nuage.

7.2 Description [30]

Il s'agit d'une offre d'hébergement (applications et données) et de services (workflow, stockage et synchronisation des données, bus de messages, contacts...). Un ensemble d'API permet d'utiliser et d'accéder à cette plate-forme et aux services associés. Ces API sont exposées au travers d'un portail web (<https://portal.azure.com>) qui permet de gérer l'ensemble des services Azure.

La plate-forme Windows Azure correspond aux offres d'informatique en nuage publics de type PaaS et IaaS de Microsoft. Elle est composée des éléments suivants :

- ❖ Windows Azure :
 - Les sites Web Windows Azure (PaaS). On peut y déployer (via FTP, WebDeploy, Git ou Team Foundation Services) des applications Web en PHP, Node.js, ASP.NET. On peut associer à un site Web une base de données relationnelle qui peut être : Windows Azure SQL Database /MySQL.
 - ✓ Les rôles applicatifs (services cloud, de type PaaS) :
 - Web Role (pour exécuter des applications Web dans Internet Information Services IIS)
 - Worker Role (pour faire tourner l'équivalent de services Windows)
 - Quand les deux rôles précédents ne conviennent pas, il est également possible de charger soi-même une image de machine virtuelle. Il s'agit du VM Role. Il est à noter que ces instances de VM role sont sans état

- (stateless) à savoir que si la machine virtuelle redémarre, elle repart de l'image qu'on lui a fournie initialement.
- Les machines virtuelles dont l'état est persistant (IaaS) qui peuvent être sous différents types de systèmes d'exploitation : Windows Server /Linux.
 - ✓ Le réseau virtuel (IaaS) qui permet de configurer les plages d'adresses dans lesquelles les machines virtuelles, voire les instances de rôles seront hébergées. Le réseau virtuel peut également être connecté par réseau privé virtuel (VPN) au réseau de l'entreprise.
 - ✓ le stockage (Windows Azure storage):
 - Les blobs (binary large object) : équivalent des fichiers
 - Les tables (non relationnelles) : ensembles de clefs/valeurs
 - Les queues qui sont une forme de MOM (Middleware Oriented Messages)
 - Les lecteurs (drives) qui permettent de voir depuis un rôle applicatif un disque virtuel (VHD) stocké dans un blob (voir plus haut)
 - ❖ Windows Azure SQL database (anciennement appelée SQL Azure) est un serveur de bases de données relationnelles (l'équivalent du service SQL Server Database Services à demeure). On utilise des bases de données en tant que service, et on n'a pas la vue sur des machines virtuelles, ce qui est très conforme à la notion de plate-forme en tant que service (PaaS).
 - ❖ D'autres services de type middleware et gestion d'identité tels que :
 - ✓ Le bus de services (Service Bus) : connectivité vers des Web Services qui ont une connexion sortante vers Internet (et non entrante). Cela autorise par exemple à une application hébergée dans des Web Roles ou Worker Roles de Windows Azure d'appeler des services Web à demeure dans les locaux de l'entreprise, de façon à relier l'application Azure au reste du système d'information de l'entreprise.
 - ✓ Windows Azure Active Directory : annuaire pour la gestion des identités qui peut être optionnellement lié à Windows Active Directory par réplication des comptes utilisateurs (sans les mots de passe car on dispose aussi de la fédération d'identité). Il s'agit ici d'un service commun avec Office 365.

- ✓ Access Control Services (ACS) : gestion du contrôle d'accès à Service Bus suivant des mécanismes standards tels qu'OAuth et les Simple Web Tokens (SWT) pour les services REST.
- ✓ Cache : cache distribué en mémoire vive, utilisable pour stocker des objets sérialisables. L'avantage de cette seconde option est de ne pas consommer des ressources supplémentaires et d'utiliser plutôt la mémoire vive non utilisée des machines déjà en place.
- ❖ Windows Azure Marketplace :
 - ✓ DataMarket : place de marché pour les données.
 - ✓ Applications : place de marché pour les applications, services, composants, formations, services autour de la plate-forme Azure.
- ❖ Windows Azure Mobile Services permet de créer les services Web et la base de données en nuage pour une application mobile telle que Windows 8 (tablettes), Windows Phone, iOS (iPhone, iPad), Android. Il est très simple d'utilisation, et le code côté serveur s'écrit en Javascript.
- ❖ Windows Azure Media Services est un service de gestion de médias tels que les vidéos et sons. Il permet typiquement de transcoder des vidéos, de les exposer sur Internet, d'exposer des chaînes en live.
- ❖ Hadoop On Azure est une distribution du célèbre Framework open source Apache Hadoop (dans le domaine du big data) sur Windows Azure (et sur Windows Server).

8. Azure ML

8.1. Présentation de l'apprentissage automatique dans le cloud

Microsoft Azure : [28]

Azure Machine Learning est un service d'analyse prédictive sur le cloud qui permet de créer et de déployer rapidement des modèles prédictifs sous forme de solutions d'analyse.

En exploitant une bibliothèque d'algorithmes prêts à l'emploi, les utiliser pour créer des modèles sur un ordinateur connecté à Internet et déployer une solution prédictive

rapidement. Commençant avec des solutions et des exemples prêts à l'emploi dans la galerie Cortana Intelligence.

Azure Machine Learning fournit non seulement des outils pour modéliser des analyses prédictives, mais également un service entièrement pris en charge, qu'on peut utiliser pour déployer nos modèles prédictifs sous la forme de services web prêts à l'emploi.

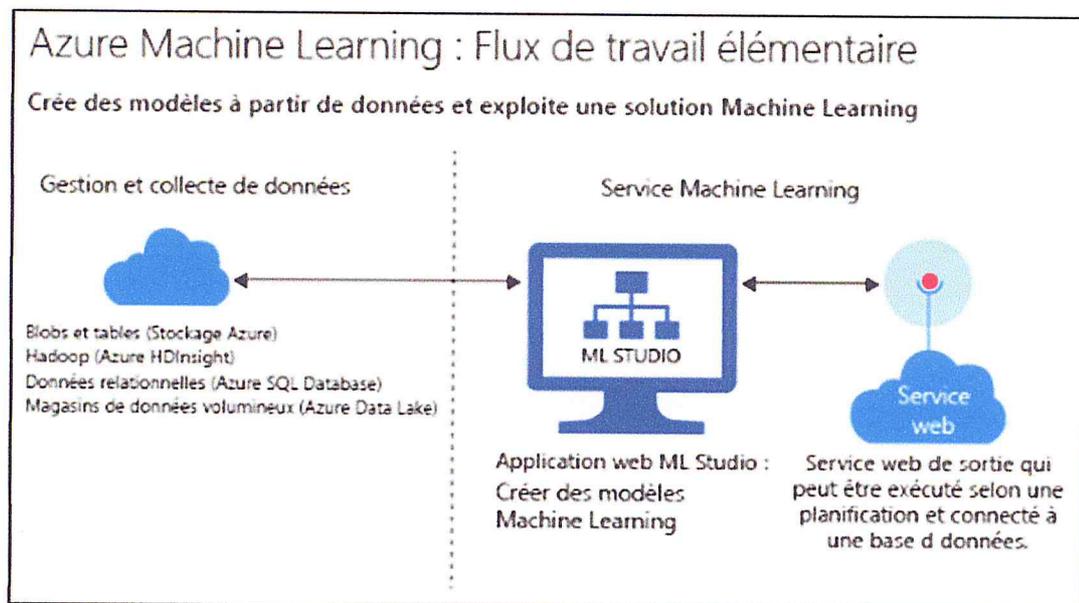


Figure 23 : le flux de travail dans azure machine Learning.

8.2. Azure ml studio : [29]

Azure Machine Learning Studio fournit un canevas interactif qui permet de développer, d'exécuter, de tester et d'itérer une expérience représentant un modèle d'analyse prédictive. Il existe un large éventail de modules capables d'effectuer les opérations suivantes :

- Entrer des données dans notre expérience
- Manipuler les données
- Former un modèle à l'aide d'algorithmes d'apprentissage automatique
- Noter le modèle
- Évaluer les résultats
- Sortir les valeurs finales

Une fois que nous sommes satisfaits de notre expérience, nous pouvons la déployer en tant que nouveau service web Azure Machine Learning afin de permettre aux utilisateurs d'envoyer de nouvelles données et de recevoir les résultats en retour.

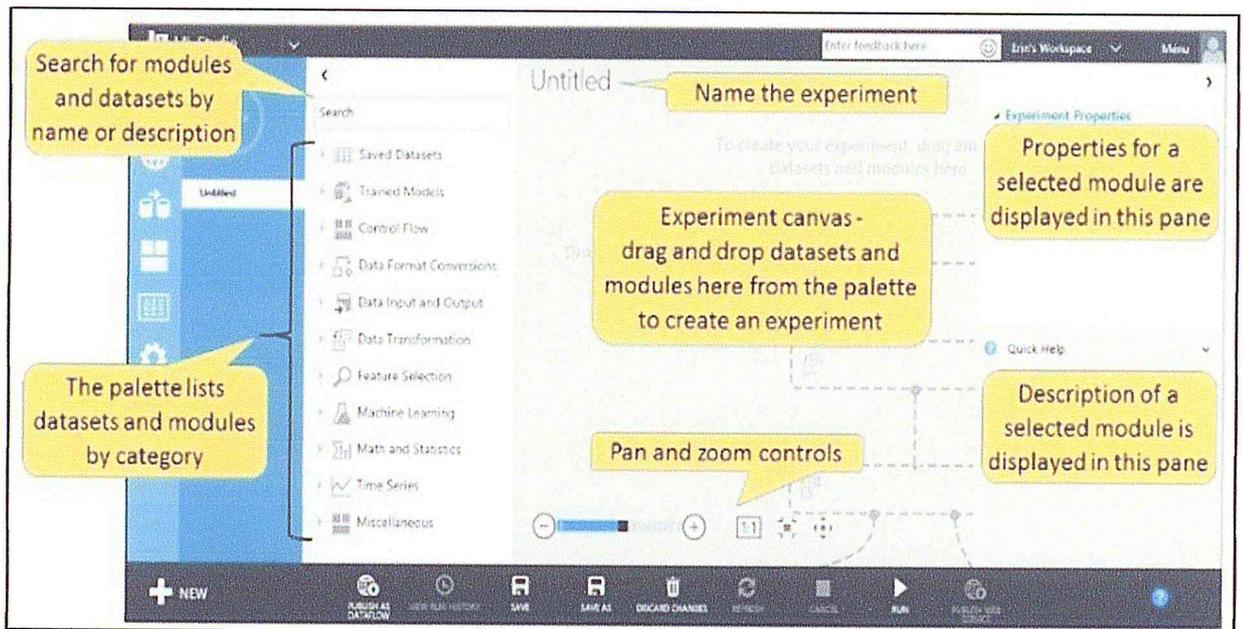


Figure 24 : la fenêtre de création d'expérimentation.

9. Conclusion

Nous avons présenté dans ce chapitre le Cloud Computing, ses caractéristiques, ses différents services et modèles de déploiement et quelques fournisseurs, y compris la plate-forme Microsoft Azure.

Nous présentons notre conception et réalisation basée sur la plateforme Microsoft Azure ML dans le chapitre suivant.

CHAPITRE IV :

Conception et Réalisation de l'application

1. Introduction

Dans ce chapitre, nous présentons notre solution qui est une plateforme de prédiction d'une situation d'urgence médicale. En effet, nous présenterons d'abord, la conception de cette application ou on trouve le modèle ainsi que l'architecture de notre solution, ensuite comme deuxième partie la réalisation de l'application.

2. Conception

2.1. Contexte du travail

Dans notre travail on s'intéresse aux personnes âgées qui vivent seules ou les personnes atteintes de maladies chroniques qui ont besoin d'une surveillance permanente. Notre solution permet le contrôle continu de ces personnes.

Dans ce travail, on propose une architecture qui répond à notre besoin qui est la prédiction d'une situation d'urgence médicale pour réagir au bon moment utilisant les paramètres du patient. Ces paramètres comprennent les signes vitaux (pression artérielle, respiration, rythme cardiaque et spo2). Ils sont primordiaux pour la surveillance médicale.

Parmi les signes vitaux surveillés dans notre travail :

- **ABP_{mean}** (mean Arterial Blood Pressure) : c'est une pression artérielle moyenne chez un individu pendant un seul cycle cardiaque
- **ABP_{sys}** (Systolic Arterial Blood Pressure) : La pression artérielle systolique est définie comme la pression maximale dans les artères, qui se produit près du début du cycle cardiaque (c'est-à-dire dans la contraction).

Valeur normale : 110-140 mmHg

- **ABP_{dias}** (Diastolic Arterial Blood Pressure) : La pression diastolique est la pression la plus basse dans le cycle cardiaque (c'est-à-dire dans la relaxation)

Valeur normale : 60-80 mmHg

- **HR** (Heart rate) : c'est le nombre de battements cardiaques (ou pulsations) par unité de temps (généralement la minute). C'est une notion quantitative qui peut aussi se définir en nombre de cycles par seconde, par l'inverse de la période.

Valeur normale : 60 à 100 /min (moyenne de 72/min)

- **RESP** (respiration) : désigne à la fois les échanges gazeux résultant de l'inspiration et de l'expiration de l'air (rejet de dioxyde de carbone CO₂ et absorption

de dioxygène O₂) et la respiration cellulaire qui permet, en dégradant du glucose grâce au dioxygène, d'obtenir de l'énergie.

Valeur normale : 12 à 20 cycles par minute

- SpO₂ (saturation « pulsée » en oxygène) : c'est une estimation de la quantité d'oxygène dans le sang.

Valeur normale : situées entre 95 et 100%

Ces signes vitaux sont utilisés dans l'algorithme de prédiction choisi comme paramètre d'entrée et la sortie sera parmi les trois valeurs « faible, normale ou élevé » pour la fréquence respiratoire.

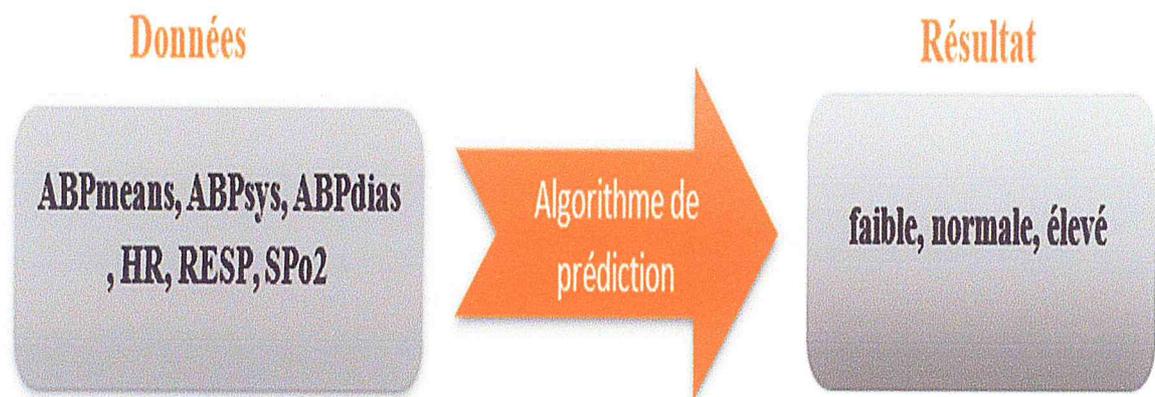


Figure 25 : données d'entrée et de sortie pour notre application.

2.2. L'algorithme de prédiction choisis

La classification et la régression sont deux types d'algorithmes d'apprentissage par machine supervisés. La régression est utilisée pour prédire une valeur continue. La classification prédit une réponse d'un ensemble défini de catégories des valeurs discrètes, comme une couleur (rouge, bleu ou vert) dans notre cas on a 3 classe à prédire (faible, normal, élevé).

Pour notre solution, parce que l'objectif est de prédire une classe nous avons choisi un modèle de classification, et le choix final de l'algorithme sera selon le taux de prédiction on va choisir l'algorithme qui donne le meilleur taux de prédiction.

C'est quoi le taux de prédiction :

Après que le modèle apprendre à faire la prédiction selon la première partie (la partie majoritaire) des données qui est utilisée dans la formation de notre modèle. La deuxième partie sera utilisée pour évaluer la performance de notre modèle formé et on va avoir une matrice de confusion qui contient la valeur souhaité dans les lignes et la valeur à prédire dans les colonnes, selon cette matrice on peut connaitre le taux de prédiction qui est la somme de transversal de cette matrice qui représente les bons résultats.

2.3. Notre processus ML

On va présenter un processus machin Learning suivant 7 étapes dans la **figure 26** suivante :

Pour prédire une situation d'urgence on a besoins des données médicale d'un patient captées par un BSN dans notre cas les données sont les signes vitaux suivant :

ABPmeans, ABPsys, ABPdias, HR, RESP, SPo2

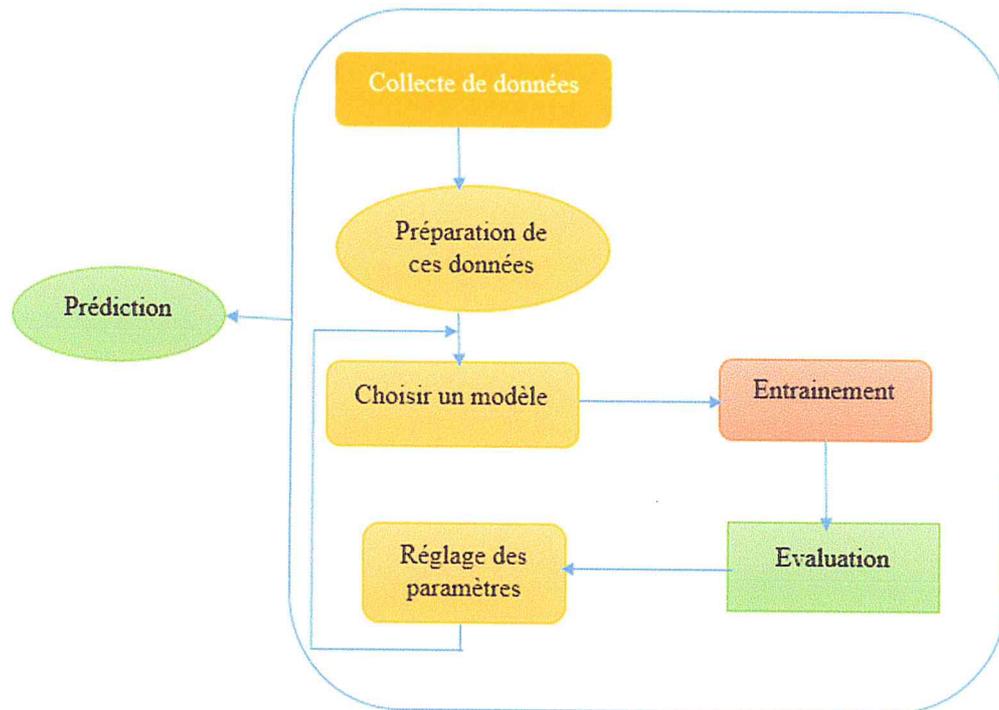


Figure 26 : Notre processus Machin Learning.

1-Collecte de données

Première étape réelle d'apprentissage automatique. Cette étape est très importante car la qualité et la quantité de données qu'on va rassembler détermineront directement la qualité de modèle prédictif

Dans notre expérience on s'intéresse à prédire le cas de la fréquence respiratoire « RESP » de ce patient prédisant la situation selon 3cas possible

Valeur RESP faible si dans l'intervalle [0-11].

Valeur RESP normale si dans l'intervalle [12-20].

Valeur RESP élevé si dans l'intervalle [21-28].

On a deux table et on va faire la jointure entre les deux pour qu'on puisse avoir une table qui contient les signe vitaux (ABPmeans, ABPsys, ABPdias, HR, RESP, SPo2) et s'il s'agit d'une valeur RESP faible normale ou élevé.

| Time (hh:mm:ss dd/mm/yyyy) | Date | ABPmean (mmHg) | ABPsys (mmHg) | ABPdias (mmHg) | HR (bpm) | RESP (bpm) | SpO2 (%) |
|-------------------------------|------|-------------------|------------------|-------------------|-------------|---------------|-------------|
| [14:34:07 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |
| [14:34:08 26/01/1995] | | 103 | 177 | 63 | 89 | 6 | 97 |
| [14:34:09 26/01/1995] | | 103 | 177 | 63 | 89 | 6 | 97 |
| [14:34:10 26/01/1995] | | 103 | 177 | 63 | 89 | 6 | 97 |
| [14:34:11 26/01/1995] | | 103 | 177 | 63 | 89 | 6 | 97 |
| [14:34:12 26/01/1995] | | 103 | 177 | 63 | 89 | 6 | 97 |
| [14:34:13 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |
| [14:34:14 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |
| [14:34:15 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |
| [14:34:16 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |
| [14:34:17 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |
| [14:34:18 26/01/1995] | | 103 | 177 | 63 | 89 | 7 | 97 |

Table 02 : Table des données médicales d'un patient.

| | |
|----|---------------------|
| 0 | valeur RESP faible |
| 1 | valeur RESP faible |
| 2 | valeur RESP faible |
| 3 | valeur RESP faible |
| 4 | valeur RESP faible |
| 5 | valeur RESP faible |
| 6 | valeur RESP faible |
| 7 | valeur RESP faible |
| 8 | valeur RESP faible |
| 9 | valeur RESP faible |
| 10 | valeur RESP faible |
| 11 | valeur RESP faible |
| 12 | valeur RESP normale |
| 13 | valeur RESP normale |
| 14 | valeur RESP normale |
| 15 | valeur RESP normale |
| 16 | valeur RESP normale |
| 17 | valeur RESP normale |
| 18 | valeur RESP normale |
| 19 | valeur RESP normale |
| 20 | valeur RESP normale |
| 21 | valeur RESP élevé |
| 22 | valeur RESP élevé |
| 23 | valeur RESP élevé |
| 24 | valeur RESP élevé |
| 25 | valeur RESP élevé |
| 26 | valeur RESP élevé |
| 27 | valeur RESP élevé |
| 28 | valeur RESP élevé |

Table 03 : Table de classification pour RESP.

| Time (hh:mm:ss) | Date (dd/mm/yyyy) | ABPmean (mmHg) | ABPsys (mmHg) | ABPdias (mmHg) | HR (bpm) | RESP (bpm) | SpO2 (%) | detection |
|--------------------|----------------------|-------------------|------------------|-------------------|-------------|---------------|-------------|--------------------|
| [14:34:07 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |
| [14:34:08 | 26/01/1995] | 103 | 177 | 63 | 89 | 6 | 97 | valeur RESP faible |
| [14:34:09 | 26/01/1995] | 103 | 177 | 63 | 89 | 6 | 97 | valeur RESP faible |
| [14:34:10 | 26/01/1995] | 103 | 177 | 63 | 89 | 6 | 97 | valeur RESP faible |
| [14:34:11 | 26/01/1995] | 103 | 177 | 63 | 89 | 6 | 97 | valeur RESP faible |
| [14:34:12 | 26/01/1995] | 103 | 177 | 63 | 89 | 6 | 97 | valeur RESP faible |
| [14:34:13 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |
| [14:34:14 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |
| [14:34:15 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |
| [14:34:16 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |
| [14:34:17 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |
| [14:34:18 | 26/01/1995] | 103 | 177 | 63 | 89 | 7 | 97 | valeur RESP faible |

Table 04 : table de données de formation.

2- Préparation de ces données

Nous devons diviser les données en deux parties. La première partie (70%), utilisée dans la formation de notre modèle, sera la majorité de l'ensemble de données. La deuxième partie (30%) sera utilisée pour évaluer la performance de notre modèle formé.

3- Choisir un modèle

La prochaine étape de notre flux de travail consiste à choisir un modèle. Il existe de nombreux modèles que les chercheurs ont créés au fil des années. Certains sont très bien adaptés aux données d'image, d'autres pour les séquences (comme le texte ou la musique), certains pour les données numériques, d'autres pour les données textuelles. Dans notre cas on va utiliser un modèle de classification. Et le choix final de l'algorithme est le **Multiclass Decision Forest** car il a donné les meilleurs résultats, avec les tests mentionnés dans la partie réalisation

4- Entraînement

Maintenant, nous passons à ce qui est souvent considéré comme l'essentiel de l'apprentissage par machine – l'entraînement. Dans cette étape, nous utiliserons nos données pour améliorer progressivement la capacité de notre modèle à prédire si une valeur respiratoire est faible, normale ou élevée.

5- Évaluation

Une fois l'entraînement est terminé, il est temps de voir si le modèle est bon, en utilisant l'évaluation. C'est là que l'ensemble de données que nous mettons de côté

plus tôt entre en jeu. L'évaluation nous permet de tester notre modèle contre des données qui n'ont jamais été utilisées pour l'entraînement. Cette métrique nous permet de voir comment le modèle peut fonctionner contre les données qu'il n'a pas encore vues. Ceci est censé être représentatif de la façon dont le modèle pourrait fonctionner dans le monde réel.

C'est là où on va utiliser la deuxième partie qui contient 30%des données.

6-Réglage des paramètres

Une fois le modèle a effectué une évaluation, il est possible de voir si on peut améliorer l'entraînement quelque façon que ce soit.

Un exemple est le nombre de fois que nous traversons l'ensemble de données de l'entraînement pendant l'entraînement. Dans notre application on a pas besoins d'utilisé ces paramètres

7- Prédiction

L'apprentissage par machine utilise les données pour répondre aux questions. Donc prédiction, ou inférence, est l'étape où nous pouvons répondre à certaines questions. C'est le résultat de l'apprentissage, le but du travail c'est tout le processus.

Nous pouvons enfin utiliser notre modèle pour prédire si la fréquence respiratoire de ce patient est faible, normal ou élevé, compte tenu de son historique respiration.

2.4. Le web service :

Définition : Lorsqu'on parle de Web Service, on parle généralement de service pouvant être interrogés via le protocole du Web.

On désigne souvent par le terme de Web service les API offertes par certains fournisseurs de services.

Les API offertes par ces fournisseurs de services permettent souvent de réaliser les même opérations que celles qu'il est possible de réaliser via leur site web mais en utilisant un langage de programmation.

Il arrive même que dans ce cas le site web soit un client d'API Web qui permet de réaliser les opérations demandées.

D'un point de vue technique, on distingue principalement deux types de Web services (tous les deux utilisent le protocole HTTP) :

- Les web services SOAP. SOAP est un protocole qui s'insère « ~au-dessus de HTTP~ » utilisant massivement XML et qui permet à un programme client de faire des appels à distance.
- Les web services REST. Les web services rest consistent simplement à utiliser les différents verbes de HTTP (GET, POST, PUT, DELETE) pour accéder, modifier, créer et supprimer des ressources.

Déployez le processus en tant que service Web prédictif – Lorsqu'on est satisfait de notre processus, on peut le déployer en tant que service Web pour prédire la fréquence respiratoire en utilisant de nouvelles données.

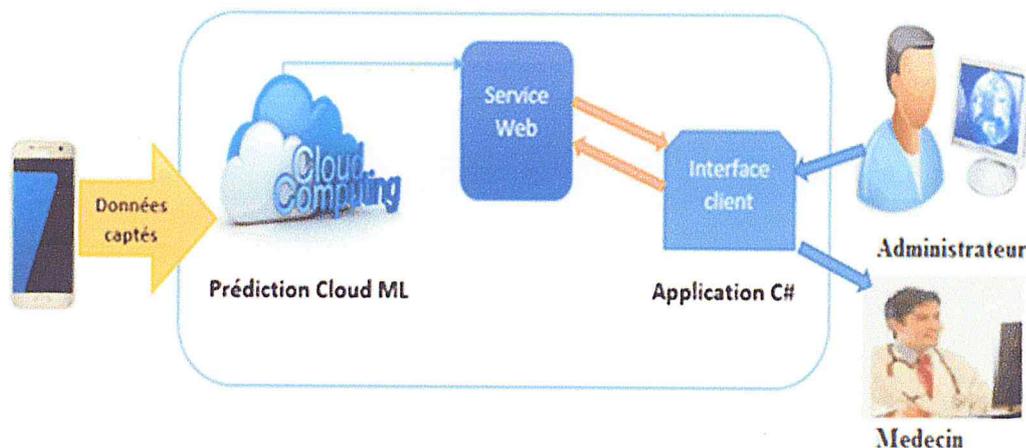


Figure 27 : Architecture client service de notre application.

Azure ML utilise les données envoyées à partir de smart phone de patient qui sont captés par le BSN pour faire la prédiction sur la situation de « RESP » ensuite à l'aide d'une application cliente développé sous Visual Studio se fait l'appelle à notre service web en utilisant son API key et qui envoie une alerte au mail de médecin si la valeur prédite représente une situation d'urgence et c'est l'administrateur qui donne le mail de médecin à l'application.

2.5. Architecture proposé

Pour notre application, on propose l'architecture décrite dans la *figure 28*.

-Les capteurs sont installés sur le corps du patient pour enregistrer les signes vitaux (pression artérielle, respiration, rythme cardiaque et spo2) du patient.

-Ils collectent et envoient les données périodiquement à une station de base (utilisée sous forme de passerelle dans notre cas c'est le smartphone du patient) qui va les transmettre à son tour à un service de Cloud à travers l'internet.

-Ces données seront alors disponibles sur un serveur de Cloud computing qui est Azure ML qui va faire la prédiction de la valeur respiratoire.

Et enfin on va voir si cette valeur représente une situation d'urgence (si elle est faible ou élevée) ou pas (si elle est normale) et envoyer une alerte au médecin selon le cas.



Figure 28 : Architecture proposée.

Les étapes de cette architecture sont comme suite

1. Les capteurs acquièrent les données puis les envoient à la passerelle : les capteurs sont placés au corps de patient et la passerelle représente le smartphone du patient
2. La passerelle envoie les données au cloud computing : l'envoi des données se fait via l'internet par un réseau internet WAN ou 3G.

Dans notre cas, on s'intéresse par les étapes 3, 4 et 5

3. Les données sont récupérées dans le cloud et stockées dans notre serveur
4. Analyse des données récupérées par un moteur de règles ici il s'agit d'Azure ML, on va utiliser un algorithme de prédiction qui va utiliser l'historique de patient pour prédire la valeur future d'un des paramètres surveillés (la fréquence respiratoire) et détecte si la valeur prédite est normale ou présente une situation d'urgence (faible ou élevé)
5. envoyer une alerte selon le cas, après l'analyse de données si la valeur prédite représente une situation d'urgence (faible ou élevé) une alerte est envoyée au médecin de ce patient ou à l'ambulance, ça se fait avec une application cliente qui utilise le service web de notre application
6. visualiser l'historique du patient : (cette étape n'a pas été implémentée dans notre application) le médecin ou l'infirmier qui fait le suivi de ce malade peut visualiser les données de son malade à tout moment

3. Réalisation

Dans ce contexte nous voulons prédire le cas de la fréquence respiratoire et pour cela nous présentons les étapes de notre expérience dans Azure ML.

3.1. Création d'une expérience :

- Nous avons créé une nouvelle expérience à partir d'Azure ML Studio.

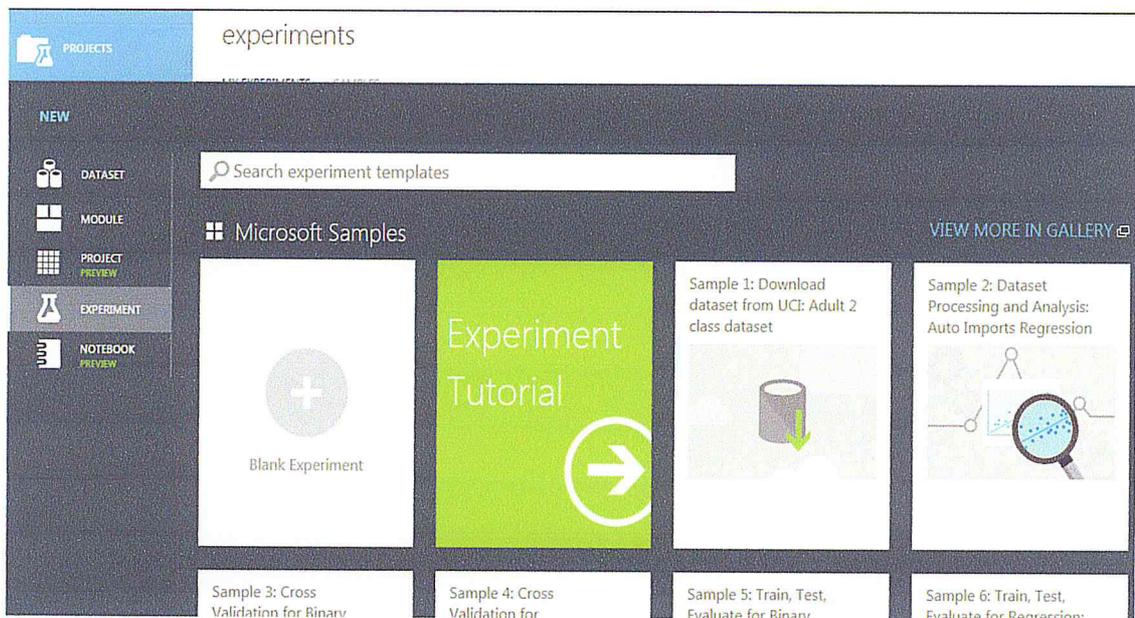


Figure 29 : nouvel expérience.

3.2. Importation de nos données :

-Il existe différentes méthodes pour charger des données dans l'expérience Azure ML selon l'endroit où les données d'entrée sont stockées. Les données peuvent être chargées à partir du système de fichiers local ou du module Reader, qui peut accéder à des données provenant de plusieurs autres emplacements de stockage persistants tels que la base de données Azure SQL, l'URL Web, etc., dans notre cas on charge notre fichier « basemedicale.CSV » à partir du système de fichiers local.

3.3. Classification

-Nous allons classifier le cas de la fréquence respiratoire avec les attributs « valeur RESP faible », « valeur RESP normale », et « valeur RESP élevé ».

-Il existe plusieurs types de classification selon le résultat souhaité, afin d'avoir un résultat multi classe sur le cas de la fréquence respiratoire (RESP) nous avons initié le fichier CSV suivant :

| RESP | detection |
|------|---------------------|
| 0 | valeur RESP faible |
| 1 | valeur RESP faible |
| 2 | valeur RESP faible |
| 3 | valeur RESP faible |
| 4 | valeur RESP faible |
| 5 | valeur RESP faible |
| 6 | valeur RESP faible |
| 7 | valeur RESP faible |
| 8 | valeur RESP faible |
| 9 | valeur RESP faible |
| 10 | valeur RESP faible |
| 11 | valeur RESP faible |
| 12 | valeur RESP normale |
| 13 | valeur RESP normale |
| 14 | valeur RESP normale |
| 15 | valeur RESP normale |
| 16 | valeur RESP normale |
| 17 | valeur RESP normale |
| 18 | valeur RESP normale |
| 19 | valeur RESP normale |
| 20 | valeur RESP normale |
| 21 | valeur RESP eleve |
| 22 | valeur RESP eleve |
| 23 | valeur RESP eleve |
| 24 | valeur RESP eleve |
| 25 | valeur RESP eleve |
| 26 | valeur RESP eleve |
| 27 | valeur RESP eleve |
| 28 | valeur RESP eleve |
| 29 | valeur RESP eleve |
| 30 | valeur RESP eleve |

Table 05 : Table de classification.

-Nous allons le corrélérer avec le fichier « basemedicale.CSV » afin de prédire le cas de la fréquence respiratoire si elle va être **faible**, **normale** ou bien **élevé** en fonction des signes vitaux : ABPmean, ABPsys,ABPdias, HR, SpO2 .

Donc les opérations suivantes sont nécessaire :

- **Join** : Jointure entre les deux tables sur la colonne RESP.
- **Select Columns in dataset** : sélection des colonnes ABPmean, ABPsys, ABPdias, HR, SpO2 hormis les colonnes RESP (nous n'en avons plus besoin de la colonne RESP car nous souhaitons prédire la colonne détection).

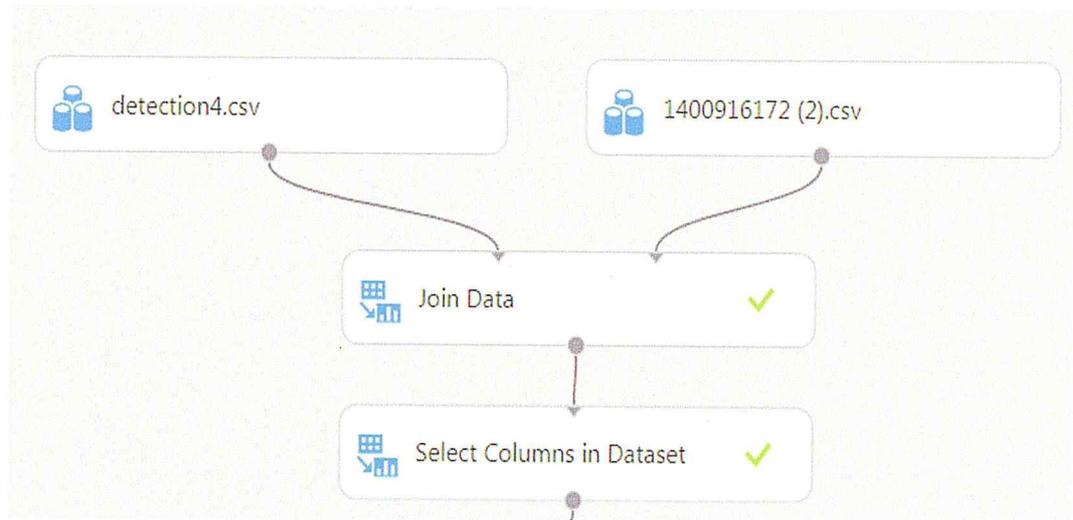


Figure 30 : Jointure des tables de données.

3.4. Mise en forme des données

L'onglet **Data Transformation** permet de mettre en forme les données, nous allons simplement partitionner le jeu de données en deux ensembles l'un pour la formation et l'autre pour le test Pour cela, nous allons appliquer le module Split.

Nous allons définir une fraction de lignes dans le premier ensemble de données de sortie à 0,7 pour diviser les données à 70/30, soit 70% pour l'entraînement et 30% pour les tests.

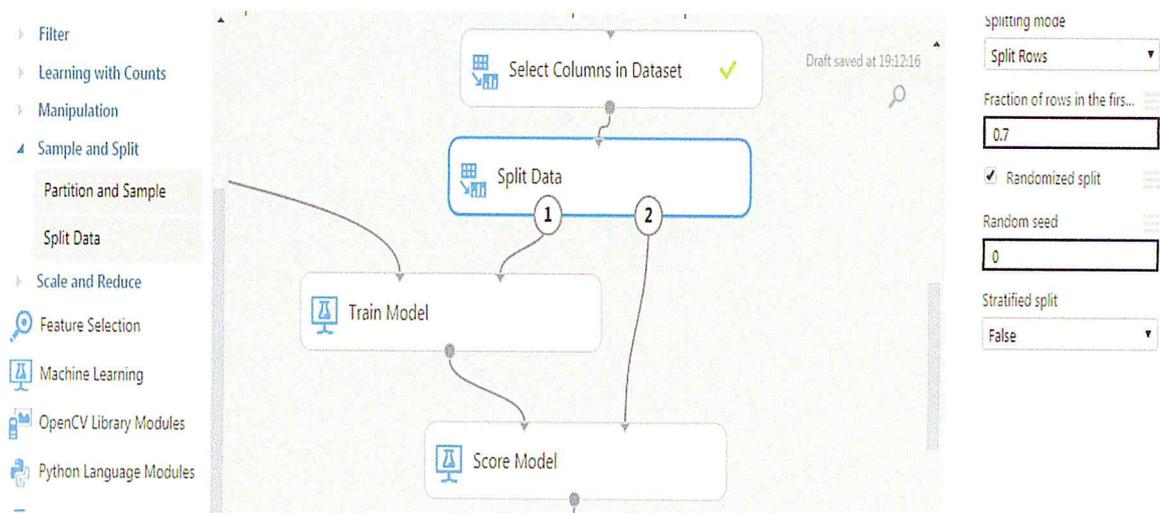


Figure 31 : Partition de données.

3.5. Création d'un modèle de Machine Learning

L'onglet Machine Learning comprend un grand nombre d'algorithmes pré-implémentés.



Figure 32 : Algorithme pour la classification.

1. Choisir un modèle de classification appelé **décision forest** grâce au module **Multiclass décision forest** sous **Initialize Model | Classification**,
 2. L'entraîner sur le premier sous-ensemble de données avec le module **Train Model** sous **Train**,
 3. Evaluer ses performances sur le second sous-ensemble avec les modules **Score Model** sous **Score** et **Evaluate Model** sous **Evaluate**.
- **Initialisation du modèle**

Commençant par sélectionner le module **Multiclass décision forest**. On utilise ce module à partir de la fenêtre principale.

- **Entraînement du modèle**

Nous allons maintenant entraîner notre modèle choisit sur les données d'entraînement à l'aide du module **Train Model**.

On accède au propriétés de module **Modèle de train**, puis on sélectionne la colonne de détection. C'est la valeur que notre modèle va prédire.

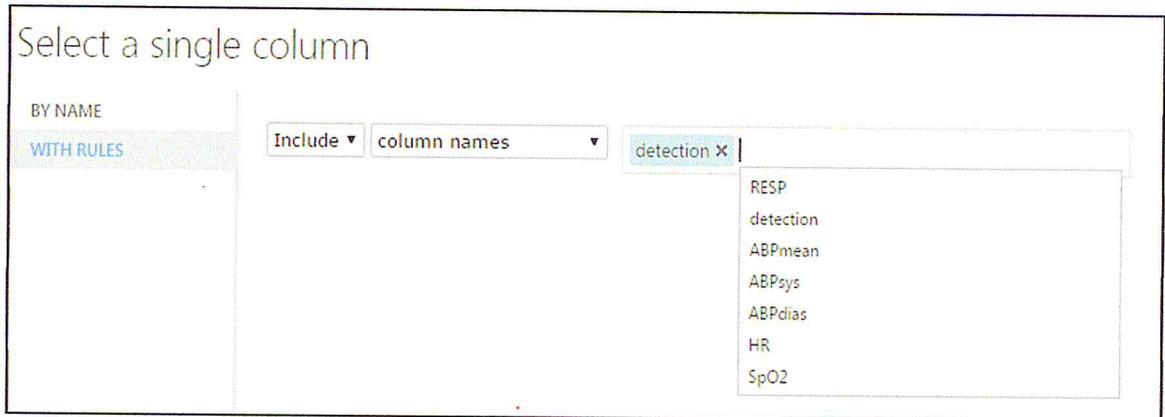


Figure 33 : Sélection de la valeur à prédire.

- **Evaluation du modèle**

Il nous reste à appliquer ce modèle entraîné sur les données de test.

Enfin, afin de tester les performances de notre algorithme, on utilise le module **Evaluate Model**.

Le workflow de notre processus est représenté par l'expérience de la **figure 34** :

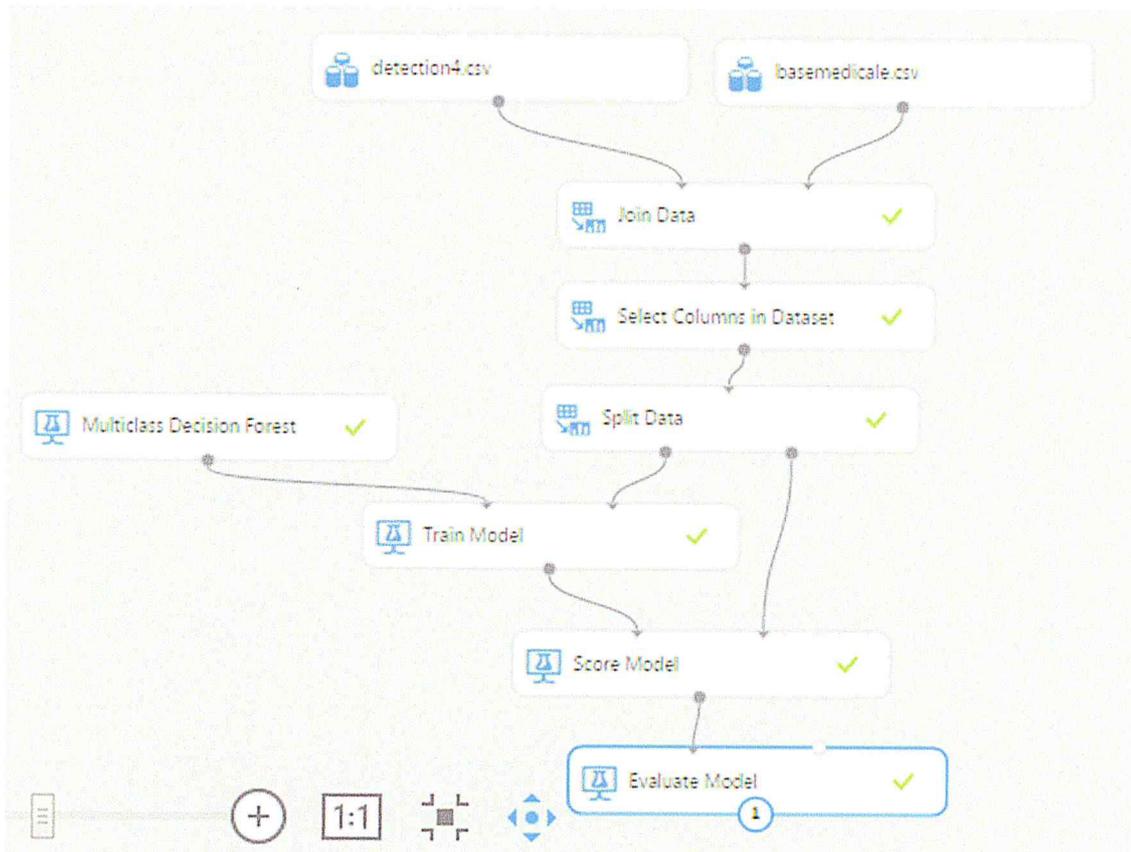


Figure 34 : Workflow de notre processus.

En sortie du module **Evaluate Model**, il est possible de visualiser les performances du classifieur sous forme de matrice de confusion. Il est présenté comme suite :

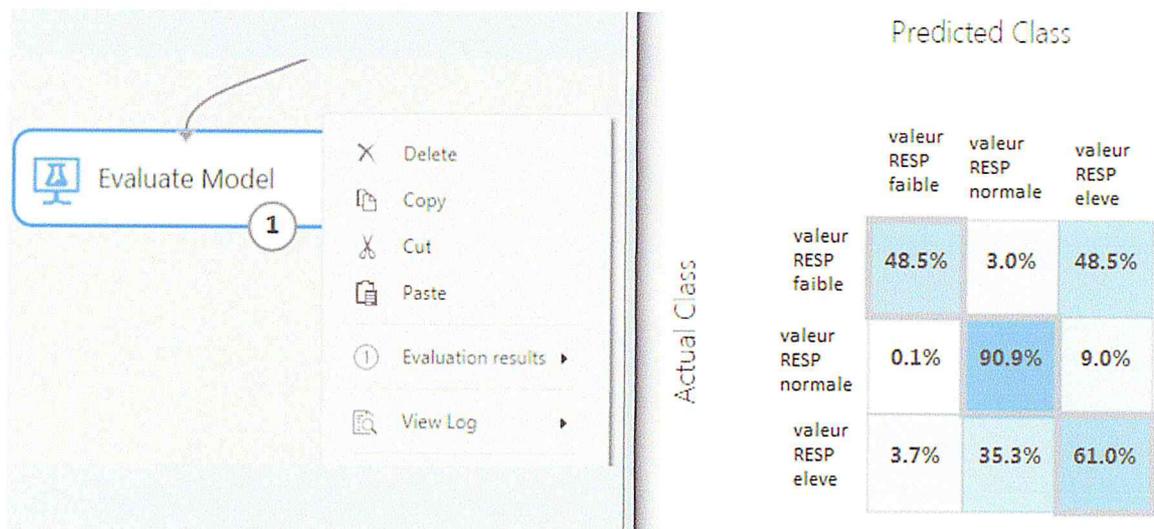


Figure 35 : Visualisation de Evaluate Model.

Notons que la deuxième entrée de **Evaluate Model** peut être reliée à un autre jeu de données sur lequel a été appliqué un autre algorithme d'apprentissage automatique afin de comparer les performances des 2 algorithmes.

Comparons les résultats avec d'autres algorithmes afin d'avoir le meilleur taux de prédiction :

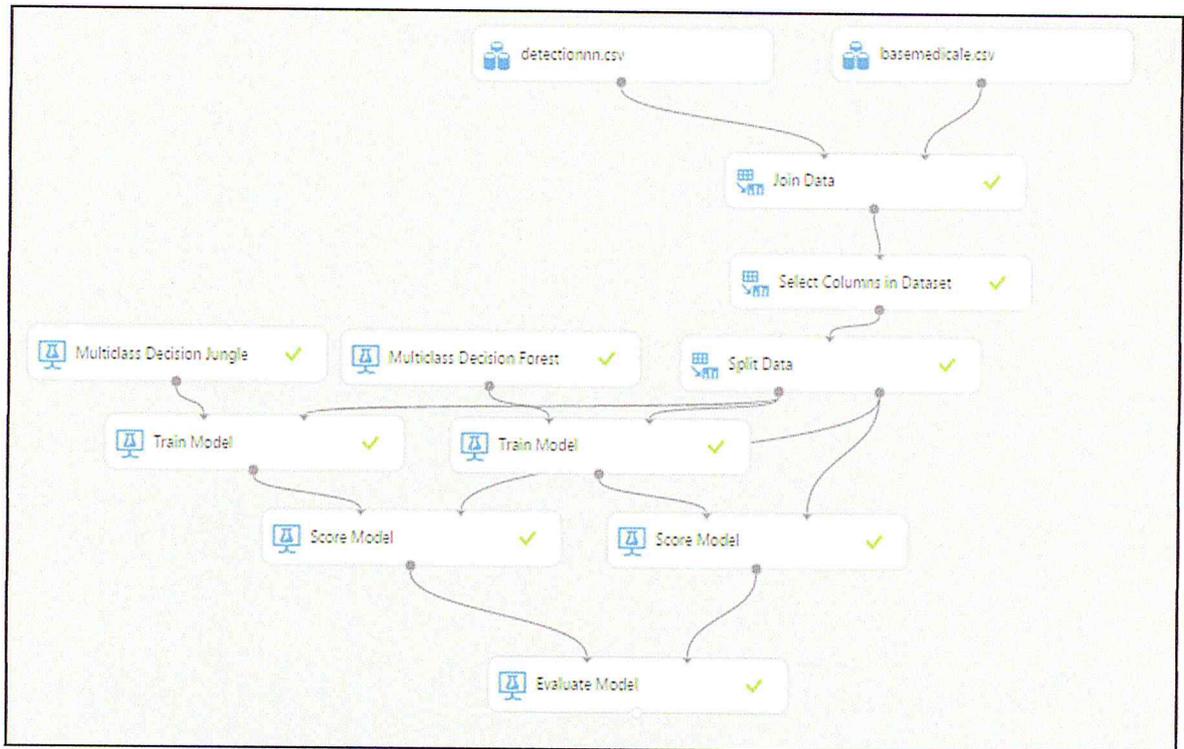


Figure 36 : Workflow de processus avec autre algorithme.

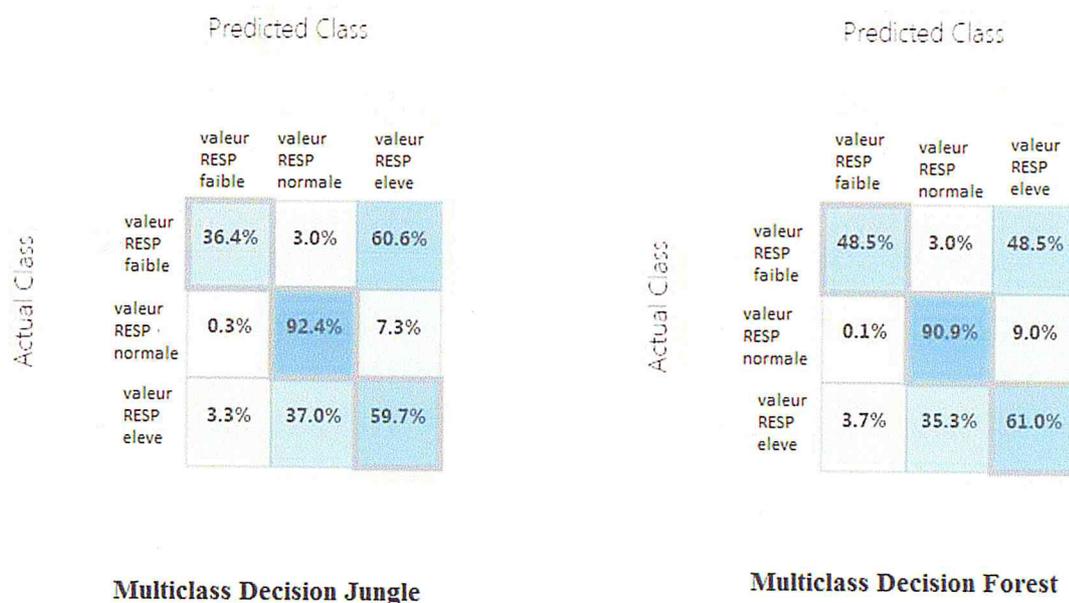


Figure 37 : Comparaison entre les deux algorithmes.

-Rappelons que le modèle entraîné (avec une partie des données sources) est évalué à l'aide des données source, une comparaison entre les prédictions et les données réelles est donc effectuée pour mesurer l'efficacité du modèle à prédire de nouveaux jeux de données.

-Le classifieur « Multiclass Decision Forest » fournit au global de meilleures prédictions, en effet plus les données situées sur la transversale sont hautes moins le modèle se trompe.

Le classifieur « Multiclass Decision Forest » a trouvé pour des valeurs RESP faible :

- 48.5 % des valeurs RESP faible : bonne prédiction
- 3 % valeurs RESP normale : mauvaise prédiction
- 48.5 % valeurs RESP élevé : mauvaise prédiction

Le classifieur « Multiclass Decision Jungle » a trouvé pour des valeurs RESP faible :

- 36.4 % des valeurs RESP faible : bonne prédiction.
- 3 % valeurs RESP normale : mauvaise prédiction.

- 60.6 % valeurs RESP élevé : mauvaise prédiction.

3.6. Publier le Web Service

Pour que d'autres personnes puissent utiliser le processus de prédiction que nous avons développé, nous pouvons le déployer en tant que service web sur Azure.

Jusqu'à présent, nous avons réalisé l'expérience avec la formation de notre modèle. Mais le service déployé n'effectue plus l'apprentissage ; il va produire de nouvelles prédictions en évaluant l'entrée de l'utilisateur en fonction de notre modèle.

Déployer comme un service web classique

On va déployer un service web classique dérivé de notre expérience, utilisant **Déployer le service web**. Machine Learning Studio déploie l'expérience en tant que service web il amène au tableau de bord associé à ce service web.

Depuis cette page, on peut exécuter un test simple du service web, en outre, il contient des informations sur la création d'applications pouvant accéder au service web).

prédiction sur le cas de la fréquence respiratoire [predictive exp.]

DASHBOARD CONFIGURATION

General **New Web Services Experience** Preview

Published experiment

[View snapshot](#) [View latest](#)

Description

No description provided for this web service.

API key

Default Endpoint

| API HELP PAGE | TEST | APPS | LAST UPDATED |
|----------------------------------|---|--|---------------------|
| REQUEST/RESPONSE | Test Test Preview | Excel 2013 or later Excel 2010 or earlier workbook | 8/6/2017 4:10:20 AM |
| BATCH EXECUTION | Test Preview | Excel 2013 or later workbook | 8/6/2017 4:10:20 AM |

Figure 38 : information sur le web service.

3.7. Test du Web Service :

Il est possible de tester dès à présent l'API en rentrant les différentes valeurs des signes vitaux afin de prédire le cas de la fréquence respiratoire s'il est faible, élevé ou normale.

Test prédiction sur le cas de la fréquence respiratoire. [Predictive Exp.] Service

Enter data to predict

ABPMEAN
100

ABPSYS
170

ABPDIAS
60

HR
80

SPO2
90

Figure 39 : Test de web service.

Résultat :

```
✓ prédiction sur le cas de la fréquence respiratoire. [Predictive Exp.] test returned [\"100\", \"170\", \"60\", \"80\", \"90\", \"0\", \"0.527113970588235\", \"0.472886029411765\", \"valeur RESP faible\"]...
```

Figure 40 : Résultat de test de web service.

La RESP serai « faible » avec une probabilité de 0.52.

3.8. Créer une application client C

On a créé une application cliente dans Visual Studio qui appelle notre service web et prédire une situation d'urgence concernant la fréquence respiratoire(RESP).

Cette application permettre de :

- Appeler notre service web (prédire la valeur de la fréquence respiratoire et dire si elle est faible, normale ou élevée)
- Détecter si cette valeur présente une situation d'urgence ou pas (l'urgence est on cas la valeur est élevée ou faible).
- Si la valeur présente une situation d'urgence un message d'alerte est envoyé au mail de médecin qui surveille ce patient.

4. Conclusion :

Dans ce chapitre nous avons pu créer un processus d'analyse prédictive permettant de prédire le cas de la fréquence respiratoire, nous avons présenté aussi la réalisation de notre application.

CONCLUSION

GÉNÉRALE

Conclusion générale :

Les réseaux de capteurs, Machine Learning et Cloud Computing, ont été identifiés parmi les technologies clefs de l'avenir et ce en raison de l'incroyable potentiel applicatif qu'elles renferment.

Notre problème majeur était de faire une analyse des données récoltées à partir d'un BSN, placé sur un patient muni d'un smartphone afin de prédire une situation d'urgence médicale.

Ce mémoire a débuté par une étude générale sur les réseaux de capteurs, en mettant en relief les différents concepts et techniques mis en œuvre et surtout les réseaux de capteurs médicaux.

Dans la deuxième partie du mémoire, nous avons expliqué les différents concepts de l'apprentissage machine pour qu'on puisse prendre une idée globale sur les algorithmes et surtout les algorithmes de prédiction, ensuite on a étudié le cloud computing afin maîtriser ces différents concepts, notamment ceux de Microsoft Azure ML. Enfin dans le dernier chapitre de notre mémoire on a présenté la conception et la réalisation de notre modèle de prédiction.

Ce projet nous a permis d'étudier et détailler l'apprentissage machine, le cloud computing et les réseaux de capteur médicale.

Comme perspective pour notre travail, on peut envisager les points suivants :

- Améliorer notre modèle en étudiant d'autres attributs (pression artérielle, SpO2, pouls, battement de cœur).
- Tester d'autres algorithmes de prédiction.
- Améliorer d'autres aspects de l'application.

Bibliographie

- [1] Chapitre 1 Caractéristiques d'un capteur. Consulté le 10 février 2017 de <https://www.fichier-pdf.fr/2015/02/07/chapitre-01-668-h14/>.
- [2] Les Capteurs. Consulté le 10 février 2017 de <http://physiquemangin.pagesperso-orange.fr/BTSSE/cours/capteursetudiant.pdf>.
- [3] Réseau de capteur sans fil. Consulté le 20 décembre 2016. de https://fr.wikipedia.org/wiki/R%C3%A9seau_de_capteurs_sans_fil.
- [4] Réseaux de Capteurs Sans Fils. Consulté le 28 février 2004 de https://moodle.utc.fr/pluginfile.php/16775/mod_resource/content/0/support-SIT60.pdf.
- [5] Juan, G.B. (2013). Réseaux de capteurs pour applications de suivi médical. Thèse de doctorat. Université Toulouse, Toulouse.
- [6] Capteurs Biomédicaux. Consulté le 28 février 2004 de http://www-master.ufr-info-p6.jussieu.fr/2005/IMG/pdf/GRETR_biocapteurs-haduong-carlier.pdf.
- [7] MAKKE, A. (2014). Détection d'attaques dans un système WBAN de surveillance médicale à distance. Thèse de doctorat. Université Paris Descartes, Paris.
- [8] Les réseaux de capteurs sans fil. Consulté le 10 mars 2004 de <http://www-igm.univ-mlv.fr/~dr/XPOSE2006/Bunel/Presentation.html>.
- [9] Belmeliani, I. (2016). Le développement d'une application mobile dédiée au streaming vidéo entre les professionnels de santé. Mémoire de master. Université Abou Bakr Belkaïd de Tlemcen, Tlemcen.
- [10] Boudra, H. (2014). Un prototype de système de télésurveillance médicale basé sur les capteurs et les réseaux de capteur sans fil. Mémoire de master. Université du Québec à Montréal. Montréal.
- [11] c'est quoi l'internet of everything ? Consulté le 15 avril 2017 de <http://www.panoptinet.com/cybersecurite-decryptee/cest-quoi-linternet-of-everything-ioe/>.
- [12] qu'est-ce que l'internet of everythings ? Consulté le 15 avril 2017 de

<http://www.cegid.com/fr/blog/qu-est-ce-que-l-internet-of-things/>.

[13] Consulté le 20 avril 2017 de <https://www.eboow.com/blog/temptraq-la-temperature-de-votre-enfant-tout-moment/>.

[14] Consulté le 20 avril 2017 de <http://www.robotshop.com/eu/fr/capteur-frequence-cardiaque-grove.html>.

[15] Makhoul, A. (2008). Réseaux de capteurs : localisation, couverture et fusion de données. Thèse de doctorat. Université de Franche-Comté, Besançon.

[16] Vincent, P. (2013). Modèles à noyaux à structure locale. Thèse de doctorat, Université de Montréal, Montréal.

[17] Apprentissage automatique (Machine Learning). Consulté le 28 avril 2017 de https://blogs.msdn.microsoft.com/big_data_france/2013/04/24/apprentissage-automatique-machine-learning/.

[18] A. Cornuéjols, L. Miclet, Y.Kodratoff, (2002). Apprentissage Artificiel : Concepts et algorithmes.

[19] Mokhtar, T. INITIATION A L'APPRENTISSAGE AUTOMATIQUE ; Support de Cours pour étudiants en Master en Intelligence Artificielle.

[20] Présentation de Machine Learning dans le cloud Azure. Consulté le 28 avril 2017 de <https://docs.microsoft.com/fr-fr/azure/machine-learning/machine-learning-what-is-machine-learning>.

[21] Classifieur bayésien naïf. Consulté le 30 avril 2017 de <https://www.xlstat.com/fr/solutions/fonctionnalites/classifieur-bayesien-naif>.

[22] Marref, N. (2013). Apprentissage Incrémental & Machines à Vecteurs Supports. Thèse de Magister, Université Hadj Lakhdar-Batna, Batna.

[23] L'apprentissage automatique. Consulté le 30 avril 2017 de <https://www.univ-tlemcen.dz/~benmammar/IA2.pdf>.

[24] Apprentissage automatique. Consulté le 29 avril 2017 de https://fr.wikipedia.org/wiki/Apprentissage_automatique.

[25] Imane, C. (2016). La Classification Des Feuilles de Vigne à Base de Descripteur Histogramme de Gradient Orienté. Thèse de Master, Université Kasdi Merbah Ouargla, Ouargla.

[26] Tahrine, H. (2013). Comparaison et mise en place des plateformes de Cloud Computing: OpenStack et Eucalyptus. Mémoire de master. Université Kasdi Merbah de Ouargla, Ouargla.

[27] Les services web. Consulté le 20 mai 2017 de <http://www-igm.univ-mlv.fr/~dr/XPOSE2004/woollams/definition.html>.

[28] Présentation de Machine Learning dans le cloud Azure. Consulté le 20 mai 2017 de <https://docs.microsoft.com/fr-fr/azure/machine-learning/machine-learning-what-is-machine-learning>.

[29] Comment faire évoluer un modèle d'apprentissage automatique d'une simple expérience en service web opérationnel. Consulté le 09 juin 2017 de <https://docs.microsoft.com/fr-fr/azure/machine-learning/machine-learning-model-progression-experiment-to-web-service>.

[30] Microsoft Azure. Consulté le 20 juin 2017 de https://fr.wikipedia.org/wiki/Microsoft_Azure.

