

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET  
DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE SAAD DAHLEB – BLIDA 01  
FACULTE DES SCIENCES  
DEPARTEMENT D'INFORMATIQUE



Mémoire de fin d'étude  
Pour l'obtention du diplôme de master en informatique  
Spécialité : Traitement automatique de la langue

**Thème :**

**Conception et réalisation d'un assistant vocal et  
textuelle pour les maisons intelligentes**

**Présenté par :**

Mlle. MEHTOUK Selma

Mlle REKKIS Selma

**Promotrice :**

Mme TEBBI.H

**Encadreur :**

Mr LICHOURI.M

**Devant le jury composé de :**

Mr. Cherif Zahar

Mme. Arkam

Président

Examinatrice

## **Remerciement**

Ce mémoire a été réalisé dans le cadre de fin d'études pour l'obtention du diplôme d'ingénieur en informatique.

Nous tenons à remercier d'abord ALLAH de nous avoir donné la force, le courage et la patience d'achever ce travail. Nous souhaitons remercier nos parents pour leurs sacrifices, leur soutien et compréhension durant toutes nos années d'étude.

Nous souhaitons remercier toutes les personnes qui nous ont apporté leur aide au cours de ce travail, et notamment : madame TEBBI.H et Mr LICHOURI.M pour le temps qu'ils nous ont consacré, leurs conseils et leurs remarques lors de l'élaboration de cette étude.

Nous adressons nos sincères remerciements aux membres du jury qui ont pris de leur temps pour évaluer ce travail. Nous exprimons notre gratitude à l'ensemble du corps enseignant, techniques et administratifs du département d'informatique à l'université de Blida, pour leur disponibilité et leur gentillesse.

Et finalement, nos sincères remerciements à tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste travail.

# Dédicace

Je dédie ce travail

A ma chère mère

Quoi que je fasse ou je dise, je ne saurai point te remercier comme il se doit, ta présence à mes coté a été toujours ma source de force, merci d'être une femme si exceptionnelle qui me pousse à travailler dur pour devenir meilleure

A mon cher père

Tu as toujours été à mes côtés pour me soutenir et m'encourager, que ce travail traduit ma gratitude et mon remerciement

A mes très chers frères *ISLAM* et *HOUSSAM*

A mes sœurs *SOUMIA* et *SARAH*

A toute ma famille

Puisse Dieu vous donne santé, bonheur et réussite

Et à mes petites princesses *RAHMA* et *ROUA*

SELMA.M

# RESUME

L'augmentation d'utilisation du mobile a fait apparaître le besoin de plus en plus d'applications mobiles dans divers domaines en se basant sur plusieurs modalités de communication (Parole, Texte, Langage des Signes).

Parmi ces applications, on trouve l'assistant personnel intelligent qui a pour tâche d'accéder à des informations à partir d'une variété de sources en ligne : la circulation routière, les nouvelles, les cours boursiers, l'agenda de l'utilisateur, les prix de vente au détail de divers produits, etc. Ces informations seront fournies par la machine (mobile) sur demande de l'utilisateur (Parole ou Texte).

Malgré que ces derniers temps, on a vu l'apparition de multitude d'applications personnelles intelligentes tel que Google assistant, Alexa ça reste restreint aux langues occidentales. Alors que pour la langue Arabe et dans les pays Arabes on trouve que l'application selma et RafiQ Afin de combler ce vide et faciliter l'accès à l'information pour les citoyens arabes, nous allons appliquer les techniques de traitement automatique de la langue en se basant sur les systèmes de compréhension automatique de l'énoncé textuel.

**Mot clé :** Classification, Compréhension de l'énoncé, Assistant personnel, Arabe, Sémantique, Texte, Parole, Maison Intelligente

# ABSTRACT

The increase in mobile use has led to the need for more and more mobile applications in various fields based on several communication methods (Word, Text, and Sign Language).

These applications include the Intelligent Personal Assistant, which has the task of accessing information from a variety of online sources: weather, traffic, news, stock prices, the user's calendar, the retail prices of various products, etc. This information will be provided by the machine (mobile) on request of the user (Word or Text).

Although in recent times, we have seen the appearance of a multitude of intelligent personal applications such as Google assistant Alexa, it remains restricted to western languages. While for the Arabic language and in the Arab countries we find the application Selma and RafiQ.

In order to fill this gap and facilitate access to information for Arab citizens, we will apply the techniques of automatic language processing based on the systems of automatic understanding of the textual statement.

**Key words:** Classification, Understanding the statement, Personal assistant, Arabic, Semantic, Text, Word, Smart Home.

## الملخص

أدت الزيادة في استخدام الهاتف المحمول للحاجة إلى المزيد والمزيد من تطبيقات الهاتف المحمول في مختلف المجالات استناداً إلى عدة أساليب اتصال (الكلمة، النص، ولغة الإشارة).

وتشمل هذه التطبيقات المساعد الشخصي الذكي، الذي يتولى مهمة الوصول إلى المعلومات من مجموعة متنوعة من المصادر على الإنترنت: الطقس، والمرور، والأخبار، وأسعار الأسهم، والتقويم الزمني للمستخدم، وأسعار التجزئة لمختلف المنتجات، وما إلى ذلك. وسوف تقدم هذه المعلومات بواسطة الجهاز المحمول، بناء على طلب المستخدم (كلمة أو نص)

وعلى الرغم من أننا شهدنا في الآونة الأخيرة ظهور العديد من التطبيقات الشخصية الذكية مثل Google Assistant، Alexa، إلا أنها تظل مقتصرة على اللغات الغربية. وفيما يتعلق باللغة العربية وفي البلدان العربية، نجد الطلب RafiQ وSelma. ولسد هذه الفجوة وتيسير وصول المواطنين العرب إلى المعلومات، سنطبق تقنيات التجهيز التلقائي للغات على أساس نظم الفهم التلقائي للبيان النصي.

**كلمات مفتاحية:** التصنيف، فهم البيان، مساعد شخصي، العربية، الدلالة، النص، الكلمة، المنزل الذكي.

# TABLE DE MATIERE

Introduction générale.....	1
Chapitre I : concepts fondamentaux d'intelligence artificiel .....	3
2. Intelligence artificielle (IA) : .....	4
2.1. Définitions : .....	4
2.2 Comment fonctionne l'Intelligence Artificielle ?.....	6
2.3 Les domaines d'application IA : .....	6
2.4 Historique d'intelligence artificielle [11] :.....	7
2.5 IA descendante et ascendante : .....	9
3. Machine Learning : .....	9
3.1 Apprentissage supervisé :.....	10
3.2 Apprentissage non supervisé [14]: .....	11
3.3 Renforcement : .....	11
4. Deep Learning:.....	11
4.1 RESEAUX DE NEURONES (NEURAL NETWORKS 'NN') :.....	13
4.1.1 Fonctionnement du réseau de neurone : .....	13
4.2 L'architecture de DL : .....	14
4.2.1 L'apprentissage non supervisé: .....	14
4.2.2L'apprentissage supervisé: .....	14
4.3 Quelques applications des RNN: .....	17
4.3.1 Long Short-Term Memory (LSTM):.....	18
5. La différence entre IA, ML, DL : .....	19
6. Conclusion :.....	20
Chapitre II : Etat de l'Art sur les assistants virtuels intelligents .....	21
1. Introduction : .....	22
2.Traitement automatique de langage (TAL):.....	22
2.1 Historique du TAL : .....	23
2.2 L'objectif du traitement automatique des langues : .....	24
2.3 Les approches du TAL : .....	24
2.4 TAL dans IA : .....	25
3.Qu'est-ce qu'un assistant virtuel intelligente? .....	25
3.1 Méthode d'interaction : .....	25
3.2 Historique d'AVI [25]:.....	26
3.3 Les assistants Virtuels populaires :.....	27

3.4 Comparaison entre Google, Alexa et siri : .....	28
4. Les assistants personnels en Arabe : .....	30
5. Les assistants virtuels pour maisons intelligentes : .....	32
5.1 Exemple des assistants virtuels pour maison intelligente : .....	32
5.2 Exemple des assistants virtuels pour maison intelligente en arabe : .....	33
6. Conclusion : .....	34
<b>Chapitre III : Conception de notre plateforme.....</b>	<b>35</b>
1. Introduction : .....	36
3.1. Corpus d'anglais : .....	37
3.2. Corpus arabes : .....	38
3.3. Nettoyage de la base de données : .....	39
3.3.1 Éliminer les mots vides : .....	40
3.3.2 Normalisation : .....	40
4. Entraînement : .....	41
4.2. Les modèles proposés : .....	42
4.2.2 KNN : .....	44
4.2.3 RNN_LSTM : .....	45
6. Conclusion .....	48
<b>Chapitre IV : Implémentation et évaluation.....</b>	<b>49</b>
1. Introduction : .....	50
2. Environnement : .....	50
2.1. Le langage de programmation : .....	50
2.2. Google drive : .....	51
2.3. Google Colaboratory : .....	51
2.4. Les bibliothèques : .....	52
3. La démarche de travail : .....	54
4. Réalisation du projet : .....	59
4.1 Page d'accueil : .....	59
5. Test et résultat : .....	62
6. Conclusion : .....	64
<b>Conclusion générale.....</b>	<b>66</b>
<b>Bibliographie : .....</b>	<b>68</b>

# LISTE DES FIGURES

Figure 1.1 : Les domaines d'application IA .....	07
Figure 1.2 : les sous-ensembles d'IA.....	10
Figure1.3 : les algorithmes de machine Learning .....	10
Figure 1.4 les neurones humains et les réseaux de neurones .....	12
Figure 1.5 : architecture d'une cellule de réseaux de neurones.....	14
Figure1.6 architecture du deep Learning.....	14
Figure 1.7 L'architecture d'un réseau de neurones CNN.....	15
Figure1.8 L'architecture de réseaux de neurones récurrents.....	16
Figure 1.9 schématisation d'application RNN.....	17
Figure 1.10 le schéma d'un seul neurone LSTM.....	18
Figure2.1 la création du Tal.....	23
Figure 2.2 : les résultats d'IQ test.....	29
Figure 2.3: speakers d'alexa et google home.....	33
Figure 3.1 : l'architecture de notre model.....	36
Figure3.2 : reprente la base de donnée anglaise.....	37
Figure 3.3: schéma de collection du dataset anglaise.....	38
Figure 3.4 : processus de chargement dataset arabe.....	39
Figure3.5 : représente base de données arabe.....	39
Figure3.6 : processus d'entrainement.....	41
Figure3.7 : exemple de probabilité avec Word level.....	44
Figure3.8 : l'architecture générale du KNN.....	45
Figure 3.9 : l'architecture générale du modèle bi-lstm.....	46
Figure4.1 : logo de langage python.....	53
Figure4.2 : logo de Google drive.....	54
Figure4.3 : logo de google colab .....	54
Figure4.4 : logo flask.....	57
Figure4.5 schéma LSTM de notre modèle.....	58
Figure4.6 résultat accuracy.....	59
Figure4.7 epoch losses.....	60
Figure4.8 les démarches du travail.....	61
Figure 4.9 page d'accueil en anglais .....	62

<b>Figure4.10 page d'accueil en arabe.....</b>	<b>63</b>
<b>Figure 4.11: exécution d'une commande en anglais.....</b>	<b>64</b>
<b>Figure 4.12 résultat d'exécution en anglais.....</b>	<b>64</b>
<b>Figure 4.13 : exemple d'exécution en arabe.....</b>	<b>65</b>
<b>Figure 4.1 résultat d'exécution en arabe.....</b>	<b>65</b>

# LISTE DES TABLEAUX

Tableau 2.1 étude comparative entre les AVI.....	30
Tableau 3.1 exemple d'élimination les mots vides en anglais.....	40
Tableau 3.2 exemple d'élimination les mots vides en arabe.....	40
Tableau3.3 : configuration des hyper-paramètre du modèle bi-lstm.....	48
Tableau 4.1 les résultats d'accuracy.....	57
Tableau4.2 : exemple d'extraction l'action, l'objet et localisation.....	62
Tableau 4.3: résultat mnb pour la langue anglais.....	62
Tableau 4.4: résultat mnb arabe.....	63
Tableau 4.5: résultat KNN anglais.....	64
Tableau4.6 : résultat knn arabe.....	64
Tableau4.7 : résultat lstm anglais.....	64
Tableau 4.8: résultat lstm en arabe.....	64

# **INTROCUCTION GENERALE**

# Introduction générale

Actuellement, le monde connaît une avance technologique considérable dans tous les secteurs et cela grâce à l'informatique. La technologie est de plus en plus présente dans notre vie quotidienne. Les smartphones et les ordinateurs dominent la vie personnelle des individus,

Ce progrès technologique de ces dernière années on les voir dans plusieurs domaine à savoir : la santé, l'agriculture, l'automobile, domotique... ect, spécialement pour ce dernier, il est possible aujourd'hui de gérer la plupart des fonctions à partir d'un Smartphone ou une tablette.

La domotique regroupe les technologies de l'informatique, de l'automatique et de l'électronique [1] permettant d'améliorer le confort, la sécurité, la communication et la gestion d'énergie d'une maison,

En effet, la domotique permet par exemple d'optimiser l'utilisation de l'éclairage, du chauffage afin de réduire notre consommation en énergie. C'est pourquoi nous avons proposé de fabriquer un modèle réduit d'une maison dite « intelligente ».

Par définition, une maison intelligente est un ensemble d'objets connectés entre eux et contrôlés par un système de base. Assurant diverses fonctions (éclairage, multimédia, chauffage et climatisation, volets, sécurité...) [2]

Notre travail a pour objectif la conception et la réalisation d'un système pour une maison intelligente, ce système permet de faciliter les taches aux habitants et d'améliorer le confort, la sécurité, la communication et la gestion d'énergie de la maison. Le présent travail est structuré en quatre chapitres : Le premier chapitre présent des concepts fondamentaux d'intelligence artificiel et deep Learning. Le deuxième chapitre présente la définition des assistants virtuels personnelles et la Comparaison entre eux, et les assistants des maisons intelligentes qui existent. Le troisième chapitre concerne la modélisation et la conception de notre

plateforme. Le dernier chapitre présente la partie l'implémentation de notre plateforme, où nous présenterons l'environnement de développement, la structure de notre application et quelques interfaces de celle-ci. Enfin, nous clôturons ce mémoire par une conclusion générale et quelques perspectives.

### **Problématique :**

Les assistants virtuels sont intégrés dans la vie de tous les jours, jouant divers rôles, dans une variété d'industries. Après faire quelques recherches, nous avons trouvé qu'avec le nombre énorme des assistants pour la maison intelligent qui existent dans le marché, mais aucun assistant entre eux peut supporter l'Arabe, sauf Google assistant qui comprendre l'arabe standard. De cela nous avons décidé de créer notre assistant pour les maisons intelligentes qui peut supporter l'arabe mais dans un corpus limiter.

### **Objectif :**

L'objectif de notre travail est la mise en œuvre d'un système de compréhension des requêtes textuelles/parlées des utilisateurs afin de fournir les informations demandées ou réaliser l'action requise pour la gestion d'une maison intelligente.

Dans notre cas on a manipulé 4 fonctionnalités les plus utilisées dans une maison intelligente qui sont (heat, light, volume, music) puis on les a adaptés à la langue Arabe.

***Chapitre I : concepts  
fondamentaux  
d'intelligence  
artificiel***

## **1. Introduction:**

Les nouvelles technologies d'intelligence artificielle (IA) sont partout, elles ont facilité la vie des individus, Les progrès de la recherche sur l'IA permettent de concentrer les efforts non seulement sur l'amélioration des capacités de l'IA, mais aussi sur la maximisation de ses bénéfices pour la société tout en respectant les valeurs éthiques.

La mise en application de l'intelligence artificielle exige des connaissances de la perception, du raisonnement, du traitement des informations et de l'apprentissage. Dans ce chapitre nous allons présenter des concepts généraux sur IA, sa définition, les domaines et les aspects d'IA et nous allons aussi parler dès les sous domaines d'intelligences artificielles et les définir.

## **2. Intelligence artificielle (IA) :**

### **2.1. Définitions :**

Selon le Larousse<sup>1</sup>, l'intelligence Artificielle se définirait comme étant «l'ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence.»[3]

En 2010, Russell<sup>2</sup>and Norvig<sup>3</sup>ont défini IA comme étant «l'étude des agents [intelligents] qui reçoivent des préventions de l'environnement et prennent des mesures. Chacun de ces agents est mis en œuvre par une fonction qui cartographie les actions, et nous couvrons différentes façons de représenter ces fonctions, comme les systèmes de production, les agents réactifs, les planificateurs logiques, les réseaux neuronaux et les systèmes décisionnels et théoriques. » [4]

---

<sup>1</sup> Dictionnaire français, [www.larousse.fr](http://www.larousse.fr)

<sup>2</sup> Informaticien connu pour sa contribution à l'intelligence artificielle

<sup>3</sup> Un scientifique américain, chercheur en informatique. Il est actuellement directeur de recherche pour Google.

« L'IA fait référence à un ensemble de méthodes et de technologies visant à faire fonctionner

intelligemment des ordinateurs ou d'autres dispositifs. L'IA consiste essentiellement en un ensemble d'algorithmes fonctionnant sur des données (généralement volumineuses). L'apprentissage machine (ML) est un sous-ensemble de l'IA qui traite des algorithmes d'extraction d'informations utiles, à partir de données complexes. Les applications de l'apprentissage machine ont eu récemment un impact inattendu dans de nombreux domaines de la science et de la technologie. Il existe une large progression de la recherche sur l'IA et sur l'augmentation probable de son impact sur le futur de la société » [5]

« L'intelligence artificielle (IA) est un ensemble de techniques permettant à des machines d'accomplir des tâches et de résoudre des problèmes normalement réservés aux humains et à certains animaux. »[6]

« IA Capable de prédire et d'adapter, l'IA utilise des algorithmes qui découvrent des modèles à partir d'énormes quantités d'informations. Prend ses propres décisions, l'IA est capable d'accroître l'intelligence humaine, de fournir des informations et améliorer la productivité. L'IA utilise des algorithmes pour construire des modèles analytiques. À partir de ces algorithmes, l'IA découvrira comment exécuter des tâches au moyen d'innombrables séries d'essais et d'erreurs. » [7]

Marvin Lee Minsky<sup>4</sup>, l'un des précurseurs de la discipline définit l'intelligence artificielle comme « la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique ». [8]

---

<sup>4</sup> Un scientifique, Il a travaillé dans le domaine des sciences cognitives et de l'intelligence artificielle.

D'après toutes ses définitions, et d'autres on verra que les définitions de l'IA vari avec le temps, et cela tout dépend du problème traité.

Maintenant, nous pouvons énoncer de manière informelle notre définition :

« IA s'agit d'une branche du domaine informatique, ayant le but la création de machines capables d'effectuer des tâches nécessitant une intelligence humaine. L'IA est la simulation de l'intelligence humaine par des machines »

## **2.2 Comment fonctionne l'Intelligence Artificielle ?**

L'IA comme l'intelligence humaine, fonctionne grâce à l'interconnexion d'un réseau de neurones. Ainsi, les experts de l'IA ont créé des équations avec des paramètres variables. Chaque équation est interconnectée avec les autres équations afin de créer la stimulation intellectuelle artificielle [9]

Selon Harry Shum, Président Exécutif de Microsoft, l'IA fonctionne seulement s'il y a présence d'une vaste quantité d'informations, d'une puissance informatique extraordinaire ; du cloud; et des algorithmes révolutionnaires, et du deep Learning ».  
[9]

## **2.3 Les domaines d'application IA :**

Après avoir vu le fonctionnement de l'IA, nous présentons dans cette partie les différents domaines d'utilisation de l'IA qui sont nombreux et touchent tous les domaines de notre vie quotidienne, On peut les résumer par (Figure 1.1), dont certains sont mentionnés ci-dessous :

- Les jeux :
- Robotique
- Transports
- La santé.... Ect

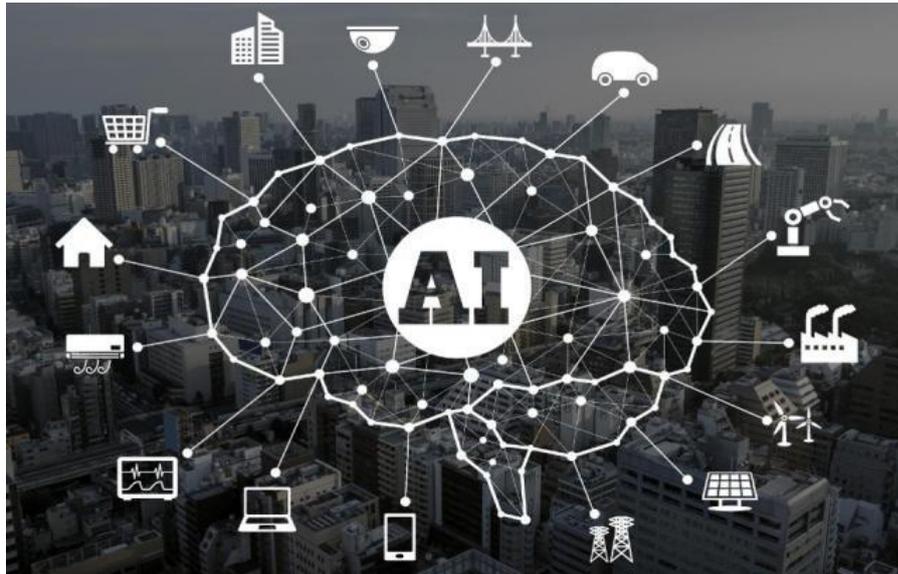


Figure 1.1 : Les domaines d'application IA [10]

## 2.4 Historique d'intelligence artificielle [11] :

Tout commence lorsque Alan Turing<sup>5</sup> demande si la machine peut « penser » comme les êtres humains, en 1950, Turing propose à la suite une expérience (test de Turing)<sup>6</sup>.

En 1952, Arthur Samuel<sup>7</sup> crée un logiciel capable d'apprendre à jouer aux échecs de manière autonome.

Le terme « intelligence artificielle » a été créé en 1955 par John McCarthy<sup>8</sup>.

L'officialisation de l'IA comme domaine scientifique date de 1956 lors d'une conférence aux États-Unis qui s'est tenue au Dartmouth Collège.

Dès le milieu des années 60, la recherche autour de l'IA sur le sol américain était principalement financée par le Département de la Défense. Dans le même temps, des laboratoires ouvrent çà et là à travers le monde.

---

<sup>5</sup> Un mathématicien et cryptologue britannique, auteur de travaux qui fondent scientifiquement l'informatique

<sup>6</sup> Le test de Turing est une proposition de test d'intelligence artificielle fondée sur la faculté d'une machine à imiter la conversation humaine

<sup>7</sup> Un pionnier américain du jeu sur ordinateur, de l'intelligence artificielle et de l'apprentissage automatique

<sup>8</sup> Pionnier de l'intelligence artificielle avec Marvin Lee Minsky

Certains experts prédisaient à l'époque que « des machines seront capables, d'ici 20 ans, de faire le travail que toute personne peut faire ».

En 1974, arrive une période appelée le « AI Winter ». Beaucoup d'experts ne réussissent pas à faire aboutir leurs projets et les gouvernements britanniques et américains réduisent leurs financements pour les académies.

Dans les années 80, le succès des systèmes experts permet de relancer les projets de recherche sur l'intelligence artificielle. Un système expert était un ordinateur capable de se comporter comme un expert (humain), mais dans un domaine bien précis.

Le développement exponentiel des performances informatiques, permet entre 1990 et 2000 d'exploiter l'IA sur des terrains jusqu'alors peu communs. On retrouve à cette époque le data mining, ou encore les diagnostics médicaux.

Dix ans plus tard, les avancées technologiques permettent un renouveau de l'intelligence artificielle. En 2008, Google réalise de formidables progrès dans le domaine de la reconnaissance de discours et lance cette fonctionnalité dans ses applications pour smartphones.

Dès le début de notre décennie, l'IA s'illustre grâce aux prouesses de Watson d'IBM. En 2011, ce super-cerveau a battu en direct les deux plus grands champions de Jeopardy. Un exercice loin d'être simple pour un ordinateur. Néanmoins, après Deep Blue, les années 2010 marquent un tournant dans la médiatisation des recherches. Nouvelle victoire de l'IA sur l'Homme en 2016, avec la victoire du système Alpha GO de Google DeepMind sur Lee Sedol, le champion de jeu de Go. L'intelligence artificielle conquiert aussi le domaine du jeu vidéo, avec notamment DeepMind Alpha Star sur Starcraft ou Open AI Five sur Dota 2.

## 2.5 IA descendante et ascendante :

Les promesses non tenues des débuts de l'IA ont amené à distinguer d'une part les machines qui non seulement mettraient en œuvre des raisonnements semblables aux raisonnements humains, mais auraient également une réelle conscience d'elles-mêmes : c'est ce qu'on appelle **l'intelligence artificielle forte** ; d'autre part les machines qui rendent de nombreux services aux humains en simulant l'intelligence humaine : **c'est l'intelligence artificielle faible**. [8]

- **l'IA faible ou descendante**: imitation fidèle d'un comportement observé et qui est reproduit à l'identique à l'aide d'un programme informatique. Il est très performant dans son domaine mais reste confiné dans celui-ci, sans possibilité d'évoluer.
- **l'IA forte (ascendante)**: cette fois le comportement humain est mimé par suite d'apprentissage et accumulation de connaissances de plus en plus complexes. La machine est dotée d'une sorte de conscience d'elle-même. Les algorithmes évoluent et échappent progressivement au cadre fixé au départ par leurs auteurs.

## 3. Machine Learning :

L'apprentissage automatique ou machine Learning est un domaine de l'intelligence artificielle (figure1.2) qui utilise des algorithmes d'apprentissage pour construire des programmes qui ont la capacité d'apprendre et d'améliorer automatiquement à partir d'expériences sans être explicitement programmé.

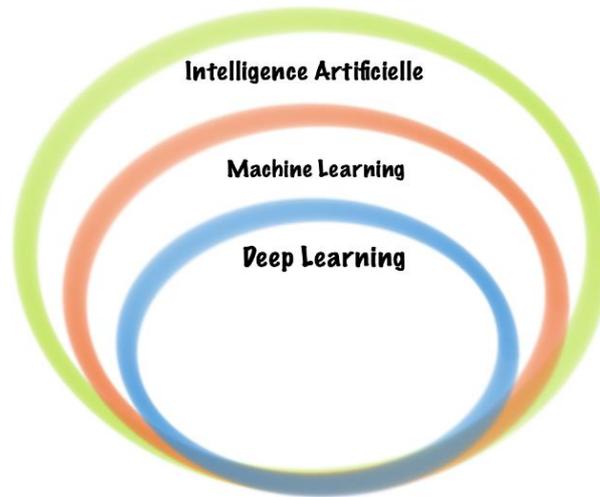


Figure 1.2 : les sous-ensembles d'IA [12]

Ces algorithmes peuvent être classés en trois catégories (figure1.3) :

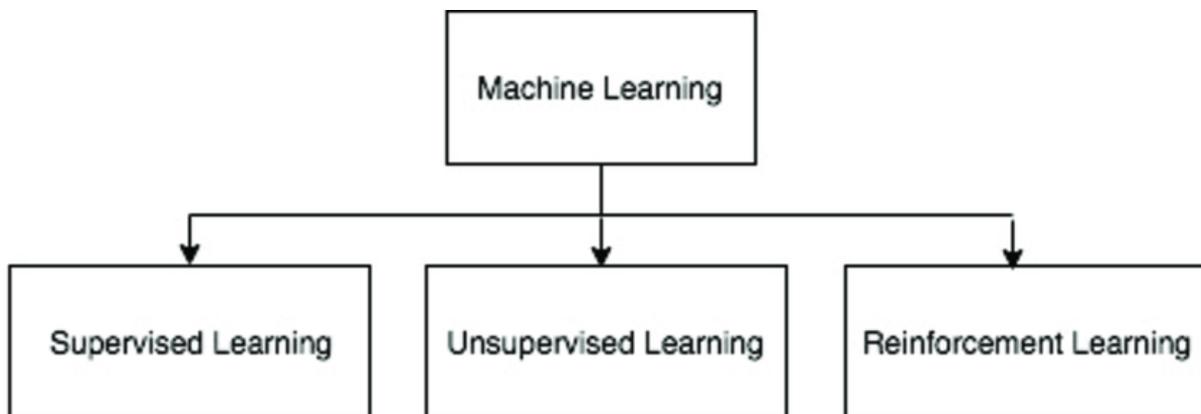


Figure1.3 les algorithmes de machine Learning [13]

### 3.1 Apprentissage supervisé :

Machine Learning avec supervision est une technologie élémentaire mais stricte, cet algorithme nécessite une variable entrée et une variable sortie, et l'ordinateur cherche des solutions pour obtenir la variable de sortie

Exemples d'apprentissage supervisé :

- KNN
- Arbre de décision.
- Naïve bayes.

- Régression.

### **3.2 Apprentissage non supervisé [14]:**

Dans la machine Learning sans supervision, on dispose d'un ensemble d'objets sans aucune valeur cible associée ; il faut apprendre un modèle capable d'extraire les régularités présentes au sein des objets pour mieux visualiser ou appréhender la structure de l'ensemble des données

#### **Type de ML sans supervision :**

Il existe deux types de machine Learning sans supervision :

**Clustering** : L'objectif consiste à trouver des regroupements dans les données.

**Association** : L'objectif consiste à identifier les règles qui permettront de définir de grands groupes de données.

Les principaux algorithmes de la machine Learning sans supervision sont :

- K-Means,
- clustering/regroupement hiérarchique
- réduction de la dimensionnalité

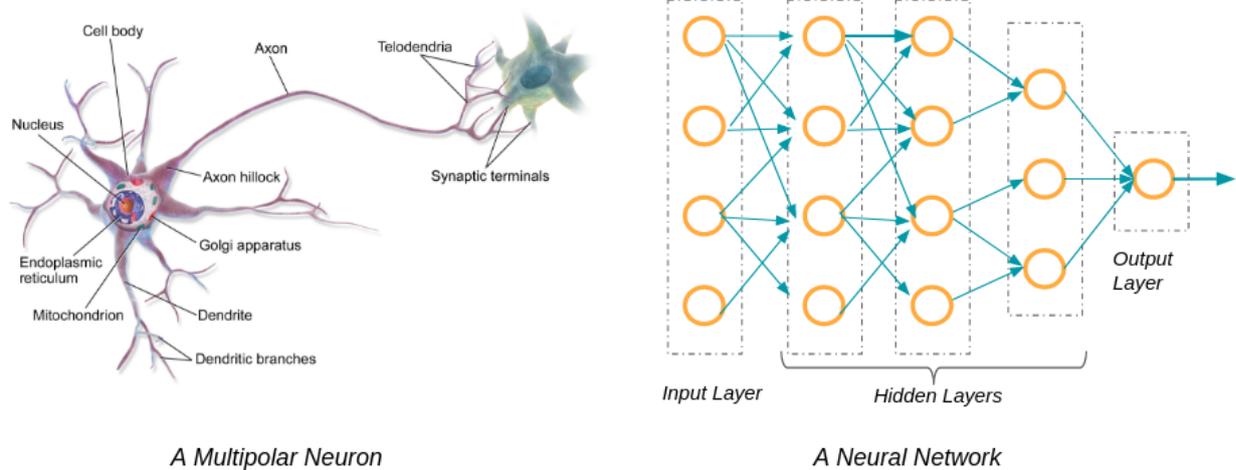
### **3.3 Renforcement :**

Dans ML de renforcement l'algorithme apprend en essayant encore et encore d'atteindre un objectif précis. Il pourra essayer toutes sortes de techniques pour y parvenir. Le modèle est récompensé s'il s'approche du but, ou pénalisé s'il échoue. [15]

## **4. Deep Learning:**

L'apprentissage automatique est un sous domaine de l'intelligence artificielle, Le Deep Learning est lui-même un sous ensemble de l'apprentissage automatique [figure1.2].

Selon Yann LECUN<sup>9</sup> spécialiste de machine Learning, « Les cerveaux humains et animaux sont “ profonds ” (figure 1.4), dans le sens où chaque action est le résultat d'une longue chaîne de communications synoptiques (de nombreuses couches de traitement). Nous recherchons des algorithmes d'apprentissage correspondants à ces “ architectures profondes ”. Nous pensons que comprendre l'apprentissage profond ne nous servira pas uniquement à construire des machines plus intelligentes, mais nous aidera également à mieux comprendre l'intelligence humaine et ses mécanismes d'apprentissages » [15]



**Figure 1.4 les neurones humains et les réseaux de neurones [16]**

Le Deep Learning est déjà utilisé dans un certain nombre de secteurs et investit de nombreux aspects de notre vie courante à savoir : la reconnaissance d'images, les voitures autonomes, la modération automatique des réseaux sociaux, assistant personnel et les chabots...

L'apprentissage profond, repose sur le traitement de quantités massives de données et ce grâce à des réseaux de neurones artificiels dont la structure imite celle du cerveau humain. Ces réseaux sont composés de centaines de « couches » de neurones, chacune recevant et interprétant les informations de la couche

<sup>9</sup> Un chercheur en intelligence artificielle et vision artificielle

précédente. Plus le nombre de neurones est élevé, plus le réseau est dit « profond ».

Dans DL, la machine est capable d'apprendre par elle-même, on peut le définir comme une architecture de réseau multi-neuronal contenant un grand nombre de paramètres et de couches.

## 4.1 RESEAUX DE NEURONES (NEURAL NETWORKS 'NN') :

Les réseaux de neurones, communément appelés des réseaux de neurones artificiels sont des imitations simples des fonctions d'un neurone dans le cerveau humain pour résoudre des problématiques d'apprentissage de la machine (Machine Learning)

### 4.1.1 Fonctionnement du réseau de neurone :

Un réseau de neurones fonctionne de manière similaire au réseau de neurones du cerveau humain. Il est composé d'une suite de couches dont la première couche correspond à la couche d'entrée (input layer), et la dernière est celle de la sortie (output layer).

Chaque couche est composée de neurones, leurs sorties représentent l'entrée de la couche suivante. Ceci est illustré dans la Figure ci-dessus.(figure1.5)

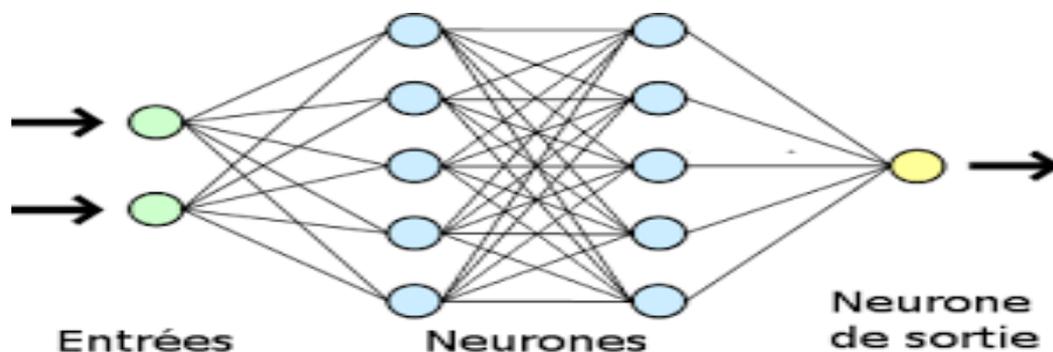


Figure 1.5 : architecture d'une cellule de réseaux de neurones [17]

## 4.2 L'architecture de DL :

Le nombre d'algorithmes utilisés dans Deep Learning est large et varié (figure 1.6)

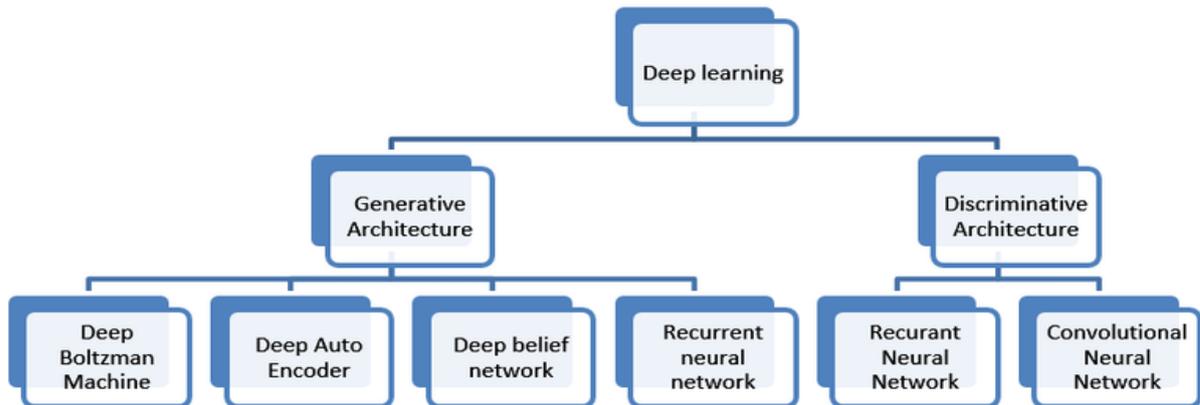


Figure1.6 architecture du deep Learning [18]

### 4.2.1 L'apprentissage non supervisé:

L'apprentissage non supervisé fait référence à l'espace du problème dans lequel il n'y a pas d'étiquette cible dans les données utilisées.

L'architecture de DL non supervisée peut être divisée en quatre parties :

- les cartes auto-organisées
- les auto-encodeurs
- les machines Boltzmann restreintes.
- RNN

### 4.2.2 L'apprentissage supervisé:

L'apprentissage supervisé est le domaine de problème dans lequel la cible à prévoir est explicitement étiquetée dans les données utilisées pour la formation.

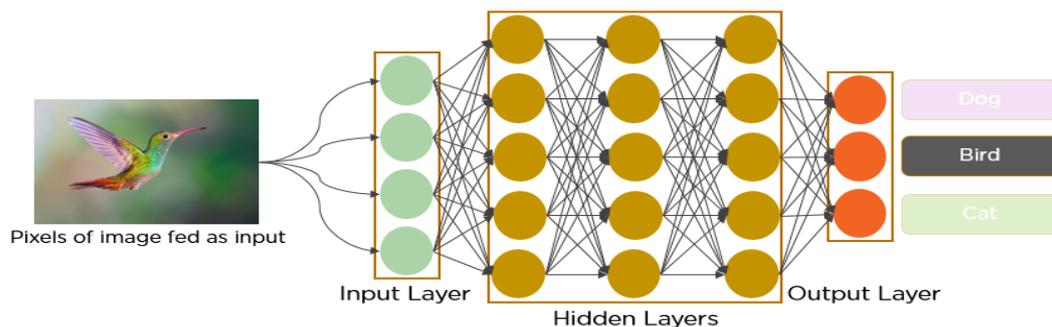
Dans cette partie, nous présentons deux des architectures d'apprentissage profond supervisé les plus courantes, **les réseaux neuronaux convolutifs** et **les réseaux neuronaux récurrents**.

- **Réseaux de neurones convolutifs :**

Ces réseaux sont utilisés pour toute application liée à l'image ou à la vidéo, comme la reconnaissance faciale ou la catégorisation des images.

« Un filtre est une matrice de nombres randomisés. Dans un CNN, les filtres sont multipliés par rapport aux représentations matricielles des bouts d'une image, ce qui permet de la balayer efficacement pixel par pixel et d'obtenir la valeur moyenne de tous les pixels adjacents, pour ainsi détecter les caractéristiques les plus importantes », explique Ajay Divakaran<sup>10</sup>, [13]

Un CNN est composé de multiples couches qui composent le réseau pour comprendre l'image de mieux en mieux à chaque passage. La première couche détecte les contours et les cadres, la deuxième couche reconnaît les formes et la troisième couche reconnaît les objets (figure1.7). La force de ce modèle est sa capacité à détecter les éléments, peu importe où ils apparaissent dans l'image.



**Figure 1.7 L'architecture d'un réseau de neurones CNN [13]**

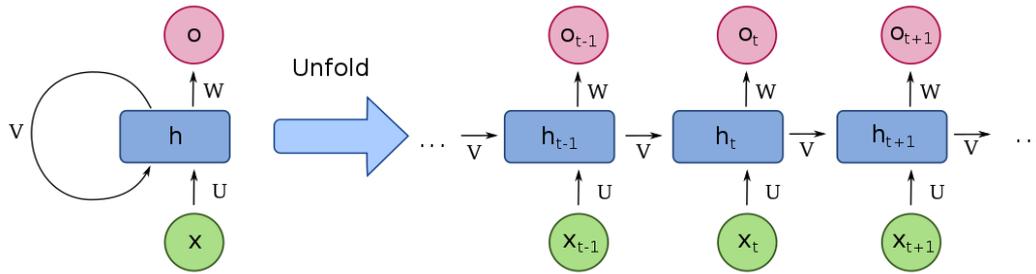
- **Les réseaux de neurones récurrents :**

Les réseaux de neurones récurrents (RNN) sont conçus pour fonctionner avec des données séquentielles. Les données séquentielles (peut être des séries chronologiques) peuvent être sous forme de texte, audio, vidéo, etc.

---

<sup>10</sup> Directeur technique du laboratoire de vision et d'apprentissage du Center for Vision Technologies de SRI International

L'architecture standard d'un RNN est illustrée dans la figure 1.8 ci-dessus



**Figure1.8 L'architecture de réseaux de neurones récurrents [19]**

Comme indiqué dans le diagramme ci-dessus. Les entrées sont reliées entre elles, La fonction  $h(t)$  a une valeur qui est déterminée par la valeur de  $h(t-1)$  et l'entrée  $x(t)$ .

D'abord, il prend le  $X(0)$  de la séquence d'entrée et ensuite il sort  $h(0)$  qui avec  $X(1)$  est l'entrée pour l'étape suivante. Ainsi,  $h(0)$  et  $X(1)$  sont les entrées pour l'étape suivante. De même,  $h(1)$  à partir de la prochaine est l'entrée avec  $X(2)$  pour l'étape suivante et ainsi de suite. De cette façon, il garde en mémoire le contexte tout en s'entraînant.

La formule pour l'état actuel est :

$$h_t = f(h_{t-1}, x_t) \quad (1)$$

Application de la fonction d'activation :

$$h_t = \tanh(w_{hh} h_{t-1} + w_{xh} x_t) \quad (2)$$

$W$  est poids,  $h$  est le vecteur caché unique,  $W_{hh}$  est le poids à l'état caché précédent,  $W_{hx}$  est le poids à l'état d'entrée actuel, **tanh** est la fonction d'activation, qui implémente une non-linéarité qui écrase les activations de  $[-1.1]$

La sortie :

$$Y_t = w_{hy} h_t \quad (3)$$

$Y_t$  est l'état de sortie.  $W_{hy}$  le poids de l'état de sortie.

### 4.3 Quelques applications des RNN:

- *One to many* : Le RNN reçoit **une unique entrée** et retourne **plusieurs sorties**, l'exemple classique de ce procédé est la légende d'image
- *Many to one* : On a plusieurs entrées et il y a une unique sortie. Une illustration de ce mode est l'**analyse de sentiment sur des textes**. Cela permet d'identifier un sentiment à partir d'un groupe de mots, déterminer le mot qui manque pour finir la phrase reçue en entrée.
- *Many to many*, on peut prendre plusieurs entrées et obtenir plusieurs sorties. Nous n'avons pas forcément le même nombre de neurones d'entrée et de sortie. Nous pouvons citer ici, la traduction de texte

Dans la figure 1.9 nous allons schématiser les applications de RNN

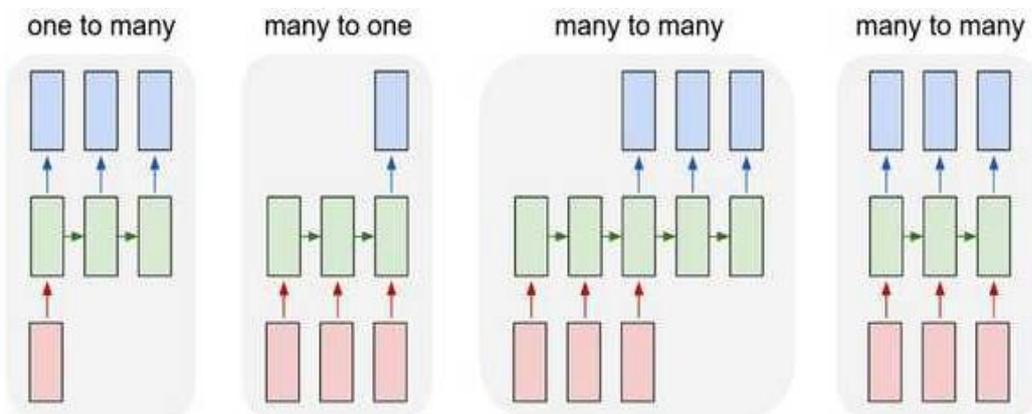


Figure 1.9 schématisation d'application RNN [20]

Malgré les performances présentées par les réseaux de neurones récurrents, néanmoins ils possèdent certaines limites :

- Modèle à mémoire courte : Il ne prend en considération que les séquences proches.
- Temps de traitement très important.
- Le modèle ne prend pas en considération les séquences ultérieures à une cellule donnée, ce qui est un handicap dans le TAL.

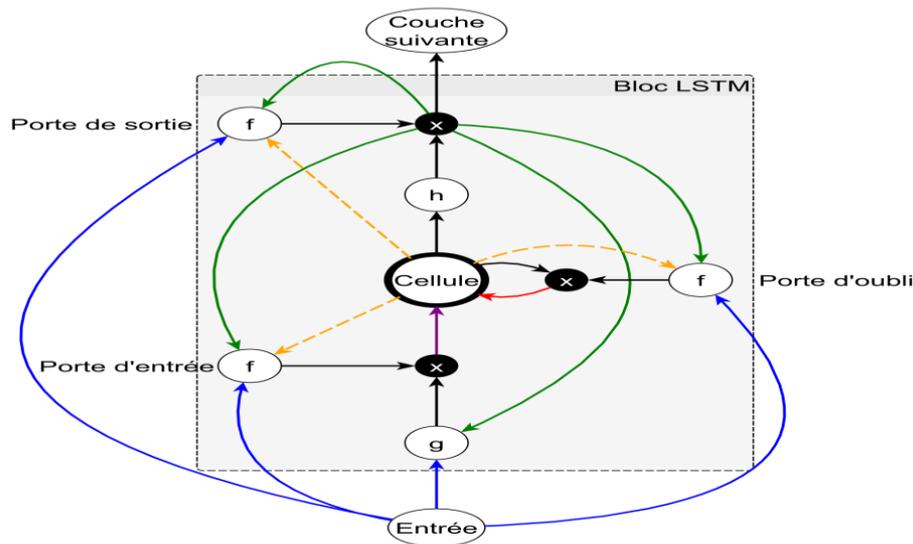
### 4.3.1 Long Short-Term Memory (LSTM):

Les unités de mémoire à court terme (LSTM), développées par les informaticiens Sepp Hochreiter<sup>1</sup> et Jurgen Schmidhuber<sup>2</sup> en 1997, constituent une solution à ce problème.

Les RNN avec des unités LSTM peuvent diviser les données en cellules mémoire à long terme et à court terme. En conséquence, ils peuvent faire la distinction entre ce qui doit être rappelé et ce qui doit être sauvé et réinjecté dans le réseau.

Les LSTM reposent sur un mécanisme de mémoire interne piloté par des **portes de contrôle**. Tous les éléments étant **différentiables**, ces réseaux s'apprennent par une **rétro-propagation** à travers le temps classique.[13]

Ce type de RNN est très utilisé en traitement du langage naturel.



**Figure 1.10 le schéma d'un seul neurone LSTM [21]**

LSTM (Long Short Term Memory) possède une mémoire interne appelée cellule (ou cell).

La cellule permet de maintenir un état aussi longtemps que nécessaire. Cette cellule consiste en une valeur numérique que le réseau peut piloter en fonction des situations (figure 1.10). [21]

Comme on peut le constater sur le schéma, la cellule mémoire peut être pilotée par **trois portes** de contrôle qu'on peut voir comme des vannes :

- **la porte d'entrée** décide si l'entrée doit **modifier le contenu** de la cellule.
- **la porte d'oubli** décide s'il faut **remettre à 0** le contenu de la cellule.
- **la porte de sortie** décide si le contenu de la cellule doit **influer sur la sortie** du neurone. [21]

Les équations régissant les trois portes de contrôle sont donc les suivantes ; elles sont l'application de la somme pondérée suivie de l'application d'une activation, typiquement la sigmoïde : [21]

**Porte d'entrée :**

$$a_l^t = \sum_{i=1}^I w_{il} x_i^t + \sum_{h=1}^H w_{hl} z_h^{t-1} + \sum_{c=1}^C w_{cl} s_c^{t-1} \quad (1)$$

$$b_l^t = f(a_l^t) \quad (2)$$

**Porte d'oubli :**

$$a_{\emptyset}^t = \sum_{i=1}^I w_{i\emptyset} x_i^t + \sum_{h=1}^H w_{h\emptyset} z_h^{t-1} + \sum_{c=1}^C w_{c\emptyset} s_c^{t-1} \quad (3)$$

$$b_{\emptyset}^t = f(a_{\emptyset}^t) \quad (4)$$

**Porte de sortie :**

$$a_w^t = \sum_{i=1}^I w_{iw} x_i^t + \sum_{h=1}^H w_{hw} z_h^{t-1} + \sum_{c=1}^C w_{cw} s_c^{t-1} \quad (5)$$

$$b_w^t = f(a_w^t) \quad (6)$$

## 5. La différence entre IA, ML, DL :

Pour mettre les choses au clair : l'intelligence artificielle (IA), la machine Learning (ML) et le deep Learning (DL) sont trois choses différentes. L'intelligence artificielle est une science comme les mathématiques ou la

biologie. Elle étudie les moyens de construire des programmes et des machines intelligentes qui peuvent résoudre de manière créative les problèmes, l'IA a toujours été considéré comme une prérogative humaine.

L'apprentissage automatique est un sous-ensemble de l'intelligence artificielle qui fournit aux systèmes la capacité d'apprendre et d'améliorer automatiquement à partir de l'expérience sans être explicitement programmé.

Le deep Learning, est un sous-ensemble de la machine Learning, qui utilise les réseaux neuronaux pour analyser différents facteurs avec une structure similaire au système neuronal humain.

## **6. Conclusion :**

Dans ce chapitre nous avons expliqué les notions de base liées à l'Intelligence Artificielle, l'apprentissage automatique, les réseaux neurones, les réseaux de neurones approfondis. Ce dernier aspect est considéré comme étant la plus récente avancée technologique de l'IA.

Notre travail s'articule sur le Traitement Automatique de la langue naturel, nous cherchons à proposer une solution plus performante à cet aspect en incluant les approches du Deep Learning.

Nous présentons dans le chapitre suivant notre travail basé justement sur l'approche du Deep Learning et les réseaux de neurones approfondis.

# **Chapitre II : Etat de l'Art sur les assistants virtuels intelligents**

## **1. Introduction :**

Les ordinateurs savent très bien interpréter les langages formels, mais un des plus grands challenges de l'informatique est de créer des solutions capables de comprendre le langage naturel. Une discipline de l'intelligence artificielle, le **Traitement Automatique du Langage Naturel**, qui se consacre aux interactions entre les machines et l'humain s'attache précisément à cet objectif.

Les assistants personnels intelligents prennent de l'importance dans la vie quotidienne. À l'avenir, avec les améliorations des techniques d'intelligence artificielle, il est prévu que la gestion individuelle de la vie sera assistée par des assistants personnels intelligents.

Ce chapitre représente notre état d'art et se concentrera sur l'introduction de différents domaines de recherche concernant notre travail, séparé en trois sections. La première section est une introduction du traitement automatique du langage, y compris la définition ainsi qu'un bref aperçu historique, suivi par les approches et les domaines du TAL

La deuxième section, comprend une définition d'un assistant personnel, et une étude comparative entre les travaux existants, et enfin nous parlera sur des assistants des maisons intelligentes.

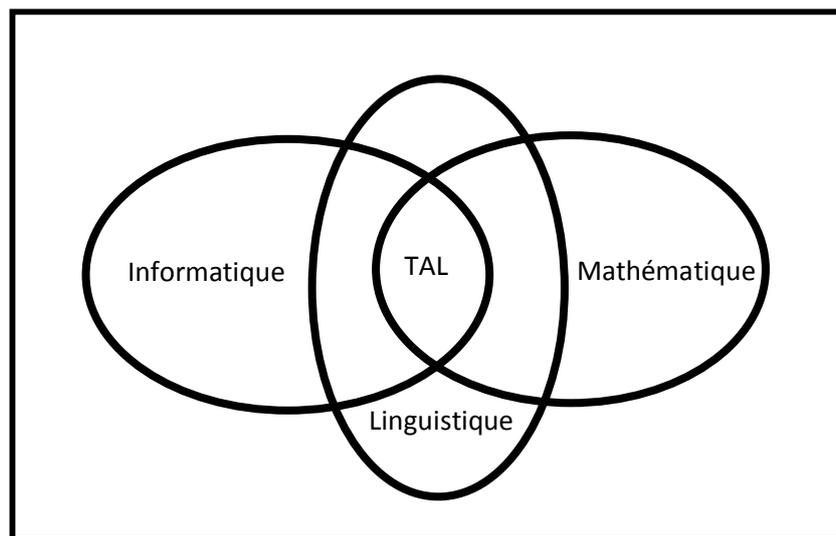
## **2. Traitement automatique de langage (TAL):**

Traitement automatique de langage (TAL) est une branche de l'informatique traitant de l'information portée par une langue, il est défini comme la manipulation automatique du langage naturel, comme la parole et le texte, par logiciel,

« Le Traitement Automatique du Langage (ou Natural Language Processing en Anglais) NLP, correspond à un cycle automatisé par l'informatique de lecture/correction/analyse de données textuelles pour en retirer différents types d'information. Une de ses déclinaisons fréquemment utilisées pour la recherche de données s'appelle le « Text Mining ». De plus, le Traitement Automatique du Langage est de nos jours souvent supporté par des algorithmes d'Intelligence Artificielle ou Machine Learning » [22]

Le traitement du langage naturel est une combinaison entre 3 domaines(Figure2.1) :

- L'informatique
- Mathématique
- linguistique



**Figure2.1 la création du Tal**

## **2.1 Historique du TAL :**

Les racines du Traitement Automatique du Langage (TAL) remontent à la conclusion de la Seconde Guerre mondiale avec la traduction automatique, ou en 1954, 60 phrases du russe à l'anglais sont traduites par ordinateur.

En 1957, Noam Chomsky<sup>11</sup> a prouvé qu'il y a des qualités similaires dans toutes les langues et a créé une sorte de grammaire qui transforme le langage réel en une forme que l'ordinateur peut comprendre. Et de cela, la révolution du TAL a été commencée.

Depuis les années 1980, une augmentation de la capacité de traitement de l'informatique, suivie du développement de l'Internet et de la communication texte numérisée (sms, emails, réseaux sociaux...), ainsi que l'émergence plus récente d'infrastructures BigData et d'algorithmes d'intelligence artificielle, a permis une explosion des capacités et des applications de traitement du langage naturel.

## **2.2 L'objectif du traitement automatique des langues :**

L'objectif du TAL est la conception de logiciels capables de traiter de façon automatique des données exprimées dans une langue. Ces données linguistiques peuvent, selon les cas, être de différents types (textes écrits, dialogues écrits ou oraux, etc.) et de taille variable. Selon la nature de l'application, le traitement peut viser à transformer des données linguistiques existantes (à des fins de correction, d'extraction d'information, de résumé, de traduction...) ou bien à en construire (génération de textes à partir d'informations). [23]

## **2.3 Les approches du TAL :**

Il existe de nombreuses façons de traiter le langage humain, les approches du TAL sont définies par les cinq points suivants :

- La reconnaissance de la parole
- La reconnaissance des lectures
- La traduction automatique
- La synthèse vocale

---

<sup>11</sup> Professeur émérite de linguistique au Massachusetts Institute of Technology.

- Recherche d'information

## 2.4 TAL dans IA :

De nos jours, le NLP a amélioré sa structure. Il est utilisé dans de nombreux appareils intelligents. Dans les dernières technologies d'IA. Le langage naturel joue un rôle essentiel. Elle détermine ce que l'utilisateur veut transmettre à l'appareil. Ensuite, elle convertit la commande en langage informatique, traite le résultat en langage humain et produit un résultat.

TAL assure une interface Homme Machine (IHM) vocale simple à utiliser et interactive.

## 3.Qu'est-ce qu'un assistant virtuel intelligente?

Les assistants virtuels intelligents (AVI) sont des **programmes artificiels intelligents**. Il s'agit d'un programme informatique intelligent qui comprend les langages naturels humains à travers des commandes vocales ou du texte et effectue des tâches pour l'utilisateur.

Un assistant personnel intelligent, aussi appelé assistant personnel virtuel est un agent logiciel qui peut effectuer des tâches ou des services pour un individu [24]

### 3.1 Méthode d'interaction :

Les assistants personnels intelligents sont à la base des logiciels à commande :

- Vocale : exemple Amazon Alexa et Siri
- Textuelle : les messageries instantanées, chat bot
- Vocal et texte : comme Google assistant

L'intérêt de ces applications est qu'à la fin l'utilisateur est censé pouvoir parler aux VPA en utilisant son langage naturel, comme il le ferait avec un autre être humain plutôt que d'avoir à utiliser des ensembles spécifiques de commandes ou un langage informatique.

### **3.2 Historique d'AVI [25]:**

La communication entre les hommes et les machines a commencé au début des années 1960. Le premier programme informatique de traitement du langage naturel, ELIZA, a été développé par Joseph Weizenbaum, professeur au MIT. Peu de temps après, une autre avancée dans la reconnaissance vocale numérique a été réalisée par IBM avec sa calculatrice à commande vocale Shoebox, présentée au grand public lors de l'exposition universelle de Seattle en 1962.

Les années 1970 ont été la décennie de la reconnaissance vocale, au cours de laquelle des entreprises et des universités, dont IBM, l'université Carnegie Mellon et le Stanford Research Institute, ont collaboré. Le résultat fut "Harpy", une machine qui maîtrisait environ 1 000 mots, avec le vocabulaire d'un enfant de trois ans capable de comprendre des phrases.

Les années 1980 ont vu naître des produits comme la machine à écrire à reconnaissance vocale d'IBM, baptisée Tangora, du nom de la dactylo la plus rapide du monde. Elle avait un vocabulaire de 20 000 mots. La naissance du premier assistant virtuel, cependant, a commencé avec IBM Simon au début des années 1990. Il s'agissait d'une technologie de reconnaissance vocale numérique qui est devenue une caractéristique de l'ordinateur personnel avec IBM et Philips.

Au début des années 2000, le premier chatbot, connu de la plupart des gens aujourd'hui, a été techniquement inventé par Colloquis, qui a lancé SmarterChild sur des plateformes comme AIM et MSN Messenger. Entièrement basé sur le texte, il était capable de jouer à des jeux, de consulter la météo, de chercher des informations et de converser avec les utilisateurs. Il est également considéré comme le précurseur du Siri d'Apple.

Le début de la décennie suivante, Apple a présenté son interface de reconnaissance vocale connue sous le nom de Siri, suivi de Google présentant Google Now, et Amazon présente Alexa et Amazon Echo. Ces assistants virtuels

n'étaient pas tels que nous les connaissons lorsqu'ils ont été introduits, mais ils ont avancé à leur forme réelle en interminable évolution. Cette évolution a commencé ce qui semble être la plus grande concurrence des plus grandes entreprises d'aujourd'hui, comme Amazon, Apple et Google.

### **3.3 Les assistants Virtuels populaires :**

Avec l'évolution actuelle du monde de la technologie. Le monde des assistants vocaux se développe à toute vitesse. Il existe déjà un certain nombre d'assistants virtuels. Un exemple d'assistant virtuel actuel disponible sur le marché est discuté dans cette section avec les tâches qu'ils peuvent fournir et leurs inconvénients.

- **Siri** : est un logiciel d'assistant personnel qui interagit avec l'utilisateur grâce à une interface vocale, reconnaît les commandes et agit sur celles-ci. Il apprend à s'adapter à la parole de l'utilisateur et améliore ainsi la reconnaissance vocale au fil du temps. Il essaie également de converser avec l'utilisateur lorsqu'il n'identifie pas la demande de l'utilisateur.

#### **Tâches prises en charge:**

Siri permet d'effectuer de nombreuses actions en commandant l'outil uniquement par la voix : appel téléphonique, SMS, recherches sur le Web, configuration du téléphone, Siri peut aussi taper des textes pour vous, rechercher vos requêtes sur les moteurs de recherche, composer un contact, avertir sur la météo, et lire les titres des journaux sélectionnés, etc.

- **Google assistant** : Comme siri est la première assistante personnelle dans les smartphones, Google assistant est le plus populaire car est intégré à la plupart des smartphones Android actuels

Google assistant est un programme d'assistance virtuelle fonctionnant grâce à **l'intelligence artificielle** disponible principalement sur les appareils mobiles Android et les accessoires de maison connectée.

### **Les taches de Google assistant :**

Google assistant peut effectuer des recherches sur Internet, planifier des événements et des alarmes, ajuster les paramètres matériels de l'appareil de l'utilisateur et afficher des informations à partir du compte Google de l'utilisateur.

- **Alexa (Amazon) :** Alexa est le nom de la technologie d'assistant virtuel créé par Amazon. Elle est présente dans différents appareils connectés à Internet. Basée sur des commandes vocales, cette technologie peut être utilisée pour piloter plus simplement les équipements compatibles et avoir accès à différentes informations rapidement, en posant directement la question à l'assistant vocal.

### **Les taches d'Alexa :**

Il est capable d'interaction vocale, de lire de la musique, faire des listes de tâches, régler des alarmes, lire des podcasts et des livres audio, et donner la météo, le trafic et d'autres informations en temps réel. Alexa peut également contrôler plusieurs appareils intelligents en faisant office de hub domotique.

## **3.4 Comparaison entre Google, Alexa et siri :**

Pour faire la comparaison entre les AVI, Loupfunds<sup>12</sup> a posé 800 questions aux trois assistants numériques les plus populaires : Google Assistant, Siri et Alexa.

- **Méthodologie [26] :**

Le test est composé de 800 questions à chaque assistant, et elles sont classées selon deux paramètres :

Comprenait-il ce qui était demandé ?

Va-t-il donner la bonne réponse ?

---

<sup>12</sup> [www.loupfunds.com](http://www.loupfunds.com)

L'ensemble de questions, conçues pour tester de manière exhaustive la capacité et l'utilité d'un assistant numérique, est divisé en 5 catégories :

**Local** – Où se trouve le café le plus proche ?

**Commerce** – Commandez-moi d'autres serviettes en papier.

**Navigation** – Comment puis-je me rendre à Uptown dans l'autobus ?

**Information** – Qui joue les Twins ce soir ?

**Commandement** – Rappelez-moi d'appeler Jerome à 14 h aujourd'hui.

- **Les résultats :**

Google Assistant a répondu à 93% d'entre eux correctement, et comprenant correctement toutes les 800 questions, Siri était le suivant, répondant à 83% correctement et seulement mal compris deux questions. Alexa a répondu correctement à 80% et seulement mal compris une. (Figure 2.2)[26]

Query Results		
	Answered Correctly	Understood Query
Google Assistant	92.9%	100%
Siri	83.1%	99.8%
Alexa	79.8%	99.9%

**Figure 2.2 : les résultats d'IQ test [26]**

Par rapport à il y a un an, chaque plateforme a connu des avancées significatives. En juillet 2018, Google Assistant a répondu à 86 % des questions correctement, par rapport à 79% pour Siri et 61% pour Alexa. [26]

Voici notre comparaison entre ces trois applications :

	<b>Google assistant</b>	<b>siri</b>	<b>Alexa</b>
<b>Nombre des langues</b>	30	13	4
<b>Taux de réponses IQ test(2019) [19]</b>	92%	83.1%	79.8%
<b>Entrée textuelle</b>	oui	non	non
<b>Entrée vocal</b>	oui	oui	non
<b>OS supporté</b>	Android, Chrome OS IOS, iPad OS, Linux Ka iOS	IOS, WatchOS, TvOS iPadOS	Amazon Echo, Fire OS, iOS, Android, Microsoft Windows
<b>Supporte l'arabe</b>	oui	oui	non

**Tableau 2.1 étude comparative entre les AVI**

### **Développer plusieurs langues :**

Dès lors qu'il est adapté à une langue donnée, un assistant vocal est difficilement exportable dans de nouveaux marchés. Il faut alors adapter l'assistant à la langue locale pour chaque nouveau marché géographique et tenir compte des accents locaux, des dialectes régionaux [27]

L'adaptation d'un assistant vocal à différentes langues est l'un des gros enjeux de ce type de produits. En particulier, la langue arabe constitue une vraie difficulté, À ce titre, Google assistant, assistant vocal qui existe depuis le plus longtemps et connaît le plus de langues différentes, est le seul assistant vocal américain à comprendre parfaitement l'arabe standard.

## **4. Les assistants personnels en Arabe :**

L'arabe est parlé par près de 300 millions de personnes à travers le monde, Lorsqu'il s'agit de comprendre la cinquième langue la plus parlée au monde, la

technologie du XXI<sup>e</sup> siècle est en retard. **Alexa** ne parle pas arabe. **Siri** et Google traduction peuvent comprendre l'arabe standard, mais pas les dialectes. Il existe deux applications qu'il peut comprendre l'arabe standard et dialecte.

- **Salma (سلمى)** : Une application en arabe basée sur les technologies modernes intelligence artificielle.

« Ce projet ambitieux est développé par la Jordan Theme Company, qui a affecté une équipe de programmeurs, de data scientists et d'intelligence artificielle à la construction de cet assistant numérique arabe. Une des bonnes possibilités de Salma est de transformer le texte en lecture audio, une technique très utile où cette technologie peut plus tard lire des livres et des articles en douceur et avec joie. L'entreprise a travaillé si dur que la voix de Salma est cohérente et proche de prononcer la bonne langue, ce qui n'est pas une étape facile. » [28]

- **Rafiq** : en tant qu'application est capable de différentes choses : Contrôle téléphonique, comme envoyer des messages SMS et WhatsApp, Passer un appel téléphonique, Régler les alarmes, Traduire un texte en 6 langues, Équilibrer la recharge via la caméra, Ouvrir des applications, Contrôler l'état du téléphone.

Intégration de services tiers, comme demander un Uber Ride, Obtenir les résultats et l'heure des matches de football, Recevoir les dernières actualités, obtenir les prix des produits de souq.com & Jumia, Obtenir les taux de change de la devise, Obtenir la météo actuelle, Rechercher des endroits sur Google maps, Faites une recherche Google. [29]  
Rafiq est disponible sous la forme d'un ChatBot voix peut d'être intégré comme un chatbot dans les applications mobiles.

**Principales caractéristiques :**

- Commandes vocales de soutien
- Soutenir la langue arabe
- Soutenir le dialecte local (égyptien et saoudien)
- Cette Solution est toujours en développement,

## **5. Les assistants virtuels pour maisons intelligentes :**

L'assistant de maison intelligente vous permet de contrôler une gamme d'appareils connectés dans votre maison à partir de votre smartphone. Des lumières aux alarmes, l'idée est que tant que votre maison et votre téléphone sont connectés à Internet, vous avez le contrôle. Connecter les appareils dans votre maison afin qu'ils puissent communiquer avec vous, les uns avec les autres, est beaucoup plus facile que vous ne le pensez.[30]

Il faut distinguer entre l'Assistant Personnel Virtuel (Apple Siri, Google Now,...) et l'Assistant Personnel Virtuel de maison (Home d'Apple, Google Home de Google et Amazon Echo – Alexa d'Amazon). Ce dernier est destiné à un usage domestique.

### **5.1 Exemple des assistants virtuels pour maison intelligente :**

**Google home :** est une application mobile qui vous permet de contrôler, organiser et gérer des dispositifs compatibles (lumières, caméras, téléviseurs, etc.) depuis une seule et même interface (l'application Google Home fonctionnant comme une télécommande avec les appareils connectés des plus grandes marques) ; gérer les paramètres et les préférences de l'Assistant Google ; gérer les paramètres de vos appareils.[31]

**Alexa** : est un assistant virtuel présenté dans différents appareils, y compris certaines TV. Basé sur des commandes vocales, Il est par exemple possible de demander à l'assistant d'ajuster le volume, de changer de chaîne, d'éteindre la TV ou encore de l'interroger à propos de la météo à venir.

Les assistants virtuels d'Amazon ou de Google trouvent une utilisation quotidienne à la maison. Ils fonctionnent avec le Wi-Fi notamment pour le speaker dont le dispositif vocal est relié à un haut-parleur (figure2.3), car l'utilisateur a besoin d'une réponse orale à ses demandes. Au niveau de son utilisation, vous pouvez utiliser la commande vocale pour envoyer des mails avec Google Home, vous pouvez régler la température de votre maison avec une simple commande vocale. Pour conclure, Amazon propose des mises à jours périodiques pour améliorer l'efficacité de leurs produits. Mais dans un futur proche, la plupart des bureaux, maisons et voitures auront leur propre système d'intelligence virtuelle par commande vocale dont l'un des leaders étant Alexa d'Amazon et Google home



Figure 2.3: speakers d'alexia et google home [32]

## **5.2 Exemple des assistants virtuels pour maison intelligente en arabe :**

Les assistants intelligents type **Google Home et Alexa** sont encore privés de la prise en charge de la langue arabe, ils doivent probablement encore avoir du travail à effectuer pour parfaire la reconnaissance vocale pour plus de 300 millions de personnes.

En conclusion, après une recherche profonde, il n'existe plus d'assistant de maison intelligente qui peut parler l'arabe, sauf Google assistants qui est un assistant personnel et un assistant aussi pour maison intelligentes.

## **6. Conclusion :**

Dans ce chapitre nous avons abordé toutes les notions de TAL, les assistants personnels et les assistants pour des maisons intelligentes. Nous avons vu aussi des exemples des assistants virtuels, Dans le prochain chapitre, on va détailler la conception de notre système et son l'architecture générales.

# **Chapitre III :**

## **Conception de notre plateforme**

## 1. Introduction :

Il y a beaucoup d'applications dans le domaine domestique, mais aucune d'entre elles n'est dédiée à la langue Arabe, sauf Google assistant.

L'objectif de notre travail est la mise en œuvre d'un système de compréhension des requêtes textuelles/parlées des utilisateurs afin de fournir les informations demandées ou réaliser l'action requise pour la gestion d'une maison intelligente.

Après un tour d'horizon sur les assistants intelligents, ce chapitre présente la conception de la solution que nous proposons, et détaille les modèles utilisés, ainsi que le processus de développement.

Mais d'abord nous allons présenter l'architecture globale de notre système.

## 2. Architecture proposée :

Notre système est divisé en deux grandes parties : la première concerne le chargement et le formatage de l'ensemble de données, tandis que la seconde représente notre système statistique. La sortie de notre système est un modèle qualifié.

Dans cette section nous présentons l'architecture de notre plateforme (Figure3.1) adopté pour chacun des modèles proposés.

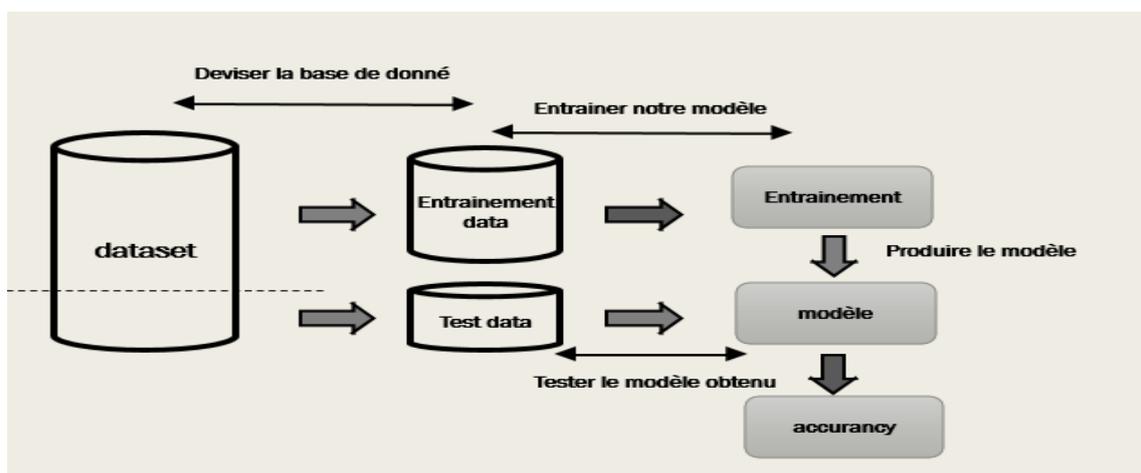


Figure 3.1 : l'architecture de notre model

### 3. Collection du dataset :

Tout système intelligent doit apprendre les connaissances pertinentes. Pour un meilleur fonctionnement et une meilleure prise de décision. Pour ce faire, il faut un corpus de données qui sert d'entrée au système.

Nous avons mené nos expériences en arabe et en anglais en parallèle. Par conséquent, nous avons utilisé deux bases de données pour tous nos modèles.

#### 3.1. Corpus d'anglais :

Nous avons mené nos études sur « Fluent Speech Commands dataset 2020(2.8GO) », pour l'anglais. La base de données Fluent Speech Commands contient 30 043 énoncés provenant de 97 haut-parleurs. Il est enregistré en tant que 16 kHz mono-canal[33]. Les fichiers contenant chacun un seul énoncé utilisé pour contrôler les appareils ménagers intelligents ou l'assistant virtuel, par exemple, « mettre de la musique » ou « augmenter la chaleur dans la cuisine ».

Chaque phrase est étiquetée avec trois emplacements : action, objet et localisation, (figure 3.2).

	A	B	C	D
1	transcription	action	object	location
2	Resume	activate	music	somewhere
3	recover	activate	music	somewhere
4	recuperate	activate	music	somewhere
5	re-start this music	activate	music	somewhere
6	re-start the music	activate	music	somewhere
7	re-start the song	activate	music	somewhere
8	this melody should be played again	activate	music	somewhere
9	this song should be restarted	activate	music	somewhere

Figure3.2 : reprenste la base de donnée anglaise.

Un créneau prend l'une des multiples valeurs: par exemple le créneau « localisation » peut prendre les valeurs « aucun », « cuisine », « chambre à coucher » ou « salle de bain ».

Nous nous référons à la combinaison des valeurs de fente comme l'intention de l'énoncé. Pour chaque intention, il y a plusieurs libellés possibles : par exemple, l'intention {action : "activate", objet : "lights", location : "none"} peut être exprimée comme "allumer les lumières" ou "allumer les lampes".

L'ensemble de données a un total de 248 phrases. Après nettoyage la base original fluent, Des 248 phrases, nous avons les paraphrasé à 2000 phrases avec quilbot[34], et pour les résultats sera précise, nous avons ajouté une autre base de données qui contient 13000 phrases[35], nous avons dans la première BD 2000 phrases et dans la deuxième 13000 phrases, ce qui nous laisse dans une base de donnée de 15000 phrases, figure 3.3 représente la collection du data anglaise.

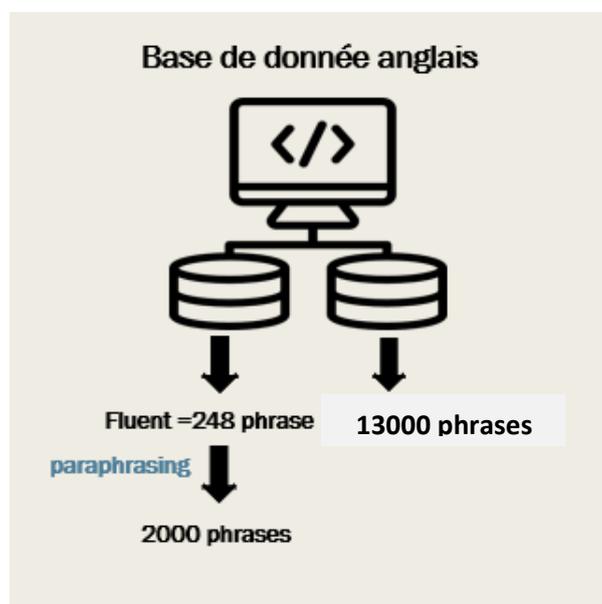


Figure 3.3: schéma de collection du dataset anglaise

### 3.2. Corpus arabes :

Pour la deuxième base de données, nous avons traduit la base de données de 2000 phrases avec Reverso [36] et Google traduction [37] pour obtenir une base d'étude de 1500 phrases. (figure3.4).

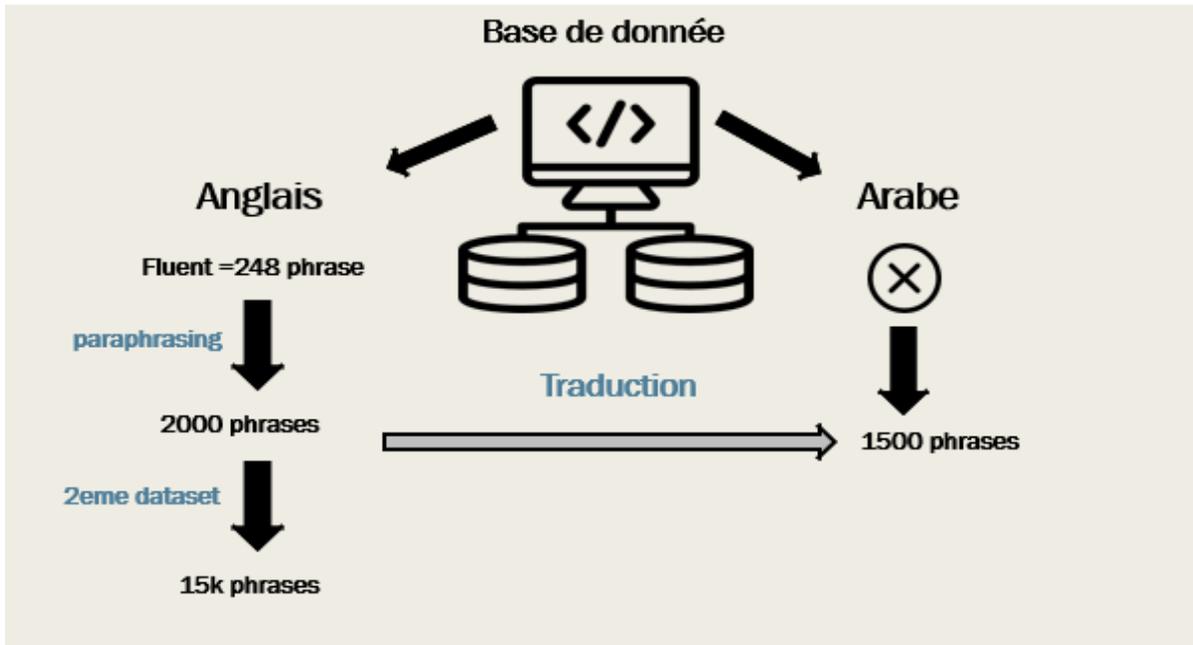


Figure 3.4 : processus de chargement dataset arabe

	A	B	C	D
1	الامر	الفعل	أغراض	الموقع
2	عَبَّرَ اللُّغَةَ	تغيير اللغة	اللغة	لا شيء
3	إِسْتَأْنَفَ	تنشيط	موسيقى	لا شيء
4	شغَلَ الأضواء	تنشيط	الضوء	لا شيء
5	أشغَلَ الأضواء	تنشيط	الضوء	لا شيء
6	أطفئ الأَنْوَارَ	تعطيل التشغيل	الضوء	لا شيء
7	الحجم أعلى	زيادة	موسيقى	لا شيء
8	ارفع الحجم	زيادة	موسيقى	لا شيء
9	خَفِّضَ الحجم	نقصان	موسيقى	لا شيء
10	ارفع درجة الحرارة	زيادة	الحرارة	لا شيء

Figure 3.5 : représente base de données arabe.

Nos bases de données ne sont pas trop volumineuses par rapport aux exigences de l'apprentissage profond, Notre choix a été imposé par le manque des ressources (base de donnée arabe).

### 3.3. Nettoyage de la base de données :

La première phase est le nettoyage de l'ensemble des données. Nous avons supprimé toutes les données qui n'étaient pas nécessaires à notre travail tel que les données redondantes.

### 3.3.1 Éliminer les mots vides :

Un des filtres les plus courants consiste à retirer les mots porteurs d'une information faible (« *stop Word* » ou « mots vides ») afin de limiter le stockage. [38]

Le choix d'éliminer les mots vides est délicat car, d'une part, ne pas supprimer les mots vides augmente l'espace de stockage nécessaire, d'autre part les éliminer peut diminuer la pertinence des recherches.

Dans notre corpus l'élimination des mots vides, comme « on », « off » pose une ambiguïté, car ces deux derniers mots là sont des **mots clé** (action) dans nos phrases par exemple (tableau3.1)

Avant	après
Turn the lights on	Turn lights
It's too loud turn it down	Loud turn

Tableau 3.1 exemple d'élimination des mots vides en anglais

Dans la première phrase notre modèle peut la comprendre, par « activer » ou « désactiver » les lampes, tant qu'il existe **un objet** et **une action**, il va prendre la location comme « every where » par contre dans la deuxième phrase il ne va pas la comprendre, car il manque l'action et l'objet

Avant	après
أوقف الموسيقى من فضلك	أوقف الموسيقى
ارفع درجة الحرارة في الغرفة	ارفع حرارة الغرفة

Tableau 3.2 exemple d'élimination les mots vides en arabe

### 3.3.2 Normalisation :

La normalisation est une technique souvent appliquée dans le cadre de la préparation des données pour le Machine Learning. [39] Certains algorithmes ont également besoin d'une normalisation pour modéliser correctement les données.

Et dans notre cas on ne va pas l'utilisée pour une simple raison, le vocabulaire de notre corpus est limité et des phrases sont simples et courtes.

## 4. Entraînement :

Lors de cette étape, nous avons transformé les données brutes en caractéristiques, on a utilisé trois proposés au cours de cette phase. Il s'agit d'un Modèle Naive Bayes (MNB), Le Plus Proche Voisin (KNN), et d'un réseau de neurones récurrent Bi-LSTM.

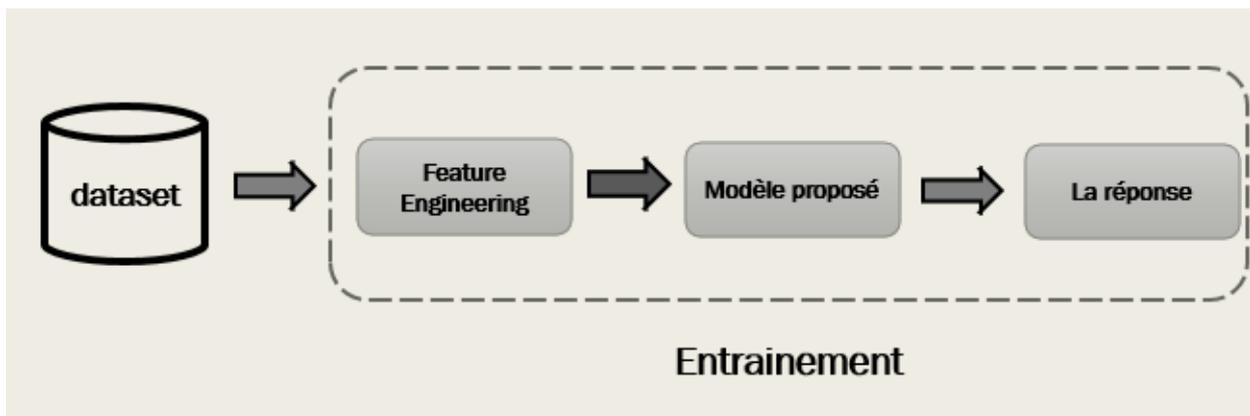


Figure 3.6 : processus d'entraînement

### 4.1 Feature Engineering :

Le Feature Engineering est un processus qui consiste à **transformer les données brutes en caractéristiques** représentant plus précisément le problème sous-jacent au modèle prédictif. Pour faire simple, il s'agit d'appliquer une connaissance du domaine pour extraire des représentations analytiques à partir des données brutes et de les préparer pour le Machine Learning. [40]

Nous mettrons en œuvre les différentes idées suivantes afin d'obtenir des fonctionnalités pertinentes de notre ensemble de données.

- **Les vecteurs TF-IDF** : est une méthode pour implémenter des outils de recherche d'information. L'exemple typique est le moteur de recherche, qui doit sélectionner les documents les plus pertinents pour une requête donnée. Une

simple recherche booléenne dans les index donne trop de résultats, qui ne sont pas ordonnés [41].

Les TF-IDF peuvent être générés à différents niveaux de jetons d'entrée (words, characters, n-grams).

- **Word Level TF-IDF** : Matrice représentant les scores TF-IDF de chaque terme dans différents documents
  - **N-gram Level TF-IDF** : N-grammes sont la combinaison de N termes ensemble.
  - **Character Level TF-IDF** Matrice représentant les scores TF-IDF de n-grammes de niveau Caractère dans le corpus
- 
- **Word Embeddings** : est une forme de représentation de mots et de documents à l'aide d'une représentation vectorielle. (**Glove, FastText, and Word2Vec**)[42]

## 4.2. Les modèles proposés :

La dernière étape de la méthode de classification des textes consiste à former le classificateur avec les caractéristiques qui ont été construites à l'étape précédente. Il existe de nombreux modèles d'apprentissage automatique qui peuvent être utilisés pour former un modèle final. Nous allons tester notre data avec 3 modèles, pour comparer les résultats, et créer un modèle performant.

### • 4.2.1 Naïve bayes model :

Naive Bayes est un algorithme de classification pour les problèmes de classification binaires (deux classes) et multi-classes. Il est appelé Bayes naïf parce que les calculs des probabilités pour chaque classe sont simplifiés pour rendre leurs calculs réalisables.

Permet de calculer la probabilité d'une donnée appartenant à une classe donnée, en prenant en considération nos connaissances précédentes. Le théorème de Bayes est énoncé comme suit:

$$P(\text{class}|\text{data}) = (P(\text{data}|\text{class}) * P(\text{class})) / P(\text{data})$$

Où:

P (class) : est la probabilité de classe compte.

P (data) : est la probabilité des données.

P (class|data) est la probabilité de classe compte tenu des données fournies.

Nous devons calculer la probabilité des données par la classe à laquelle elles appartiennent. Cela signifie que nous devons d'abord séparer nos données de formation par classe. Après calculer la probabilité de chaque classe.

	precision	recall	f1-score	support
Book	1.0000	1.0000	1.0000	657
CreativeWork	0.9225	0.9504	0.9362	564
Music	0.9975	0.9540	0.9753	1260
Restaurant	0.9984	0.9875	0.9929	638
ScreeningEvent	0.9626	0.9785	0.9705	605
Weather	0.9919	0.9777	0.9848	628
heat	0.9146	0.9677	0.9404	155
lamp	0.0000	0.0000	0.0000	0
light	1.0000	0.9529	0.9759	425
music	0.0000	0.0000	0.0000	0
volume	0.7215	1.0000	0.8382	57
accuracy			0.9707	4989
macro avg	0.7736	0.7972	0.7831	4989
weighted avg	0.9790	0.9707	0.9744	4989

0.9707356183603929

Figure 3.7 : exemple de probabilité avec Word Level

### 4.2.2 KNN :

L'algorithme K-NN (K-Nearest Neighbors) est une méthode d'apprentissage supervisé. Elle peut être utilisée aussi bien pour la régression que pour la classification [24]

Pour effectuer une prédiction, l'algorithme K-NN va **se baser sur le jeu de données en entier**. En effet, l'algorithme va chercher les **K instances du jeu de données les plus proches de notre observation**. Ensuite pour ces voisins, l'algorithme se basera sur leurs variables de sortie (output variable) pour calculer la valeur de la variable de l'observation qu'on souhaite prédire.

Les étapes de l'algorithme sont les suivantes, A partir des données d'entrée :

1. On choisit une fonction de définition pour la distance (on parle aussi de fonction de similarité) entre observations.
2. On fixe une valeur pour de plus proches voisins dans notre cas  $K=3$ .
3. Calculer toutes les distances entre cette observation en entrée et les autres observations du jeu de données
4. Conserver les k observations du jeu de données qui sont les plus « proches » de l'observation à prédire.
5. Retourner la valeur calculée dans l'étape précédente comme étant la valeur qui a été prédite par l'algorithme pour l'observation en entrée qui était inconnue.

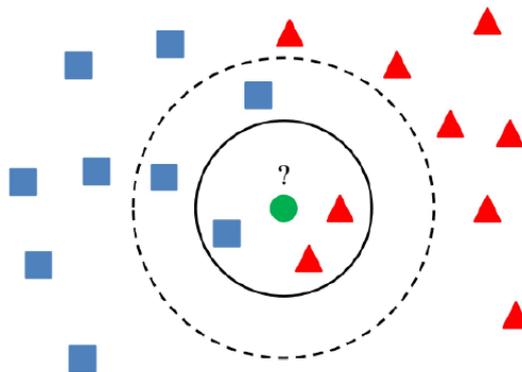


Figure 3.8 : l'architecture générale du KNN

### 4.2.3 RNN\_LSTM :

Contrairement aux réseaux neuronaux d'alimentation dans lesquels les sorties d'activation ne sont propagées que dans une seule direction, les sorties d'activation des neurones se propagent dans les deux directions (des entrées aux sorties et des sorties aux entrées) dans les réseaux neuronaux récurrents. Cela crée des boucles dans l'architecture du réseau neuronal qui agit comme un « état de mémoire » des neurones. Cet état permet aux neurones de se souvenir de ce qui a été appris jusqu'à présent. [25]

Dans cette section, nous allons expliquer en détails notre premier modèle Bi-LSTM, que nous utiliser comme point de départ.

L'architecture générale de ce modèle est illustrée dans la Figure ou :

- $x_1, x_2, x_3 \dots, x_{30}$  : représenter la séquence de mots d'une phrase passée comme un input.
- $Y_1, y_2, y_3, \dots, y_{30}$  : représentent la séquence de mots de la paraphrase produite.

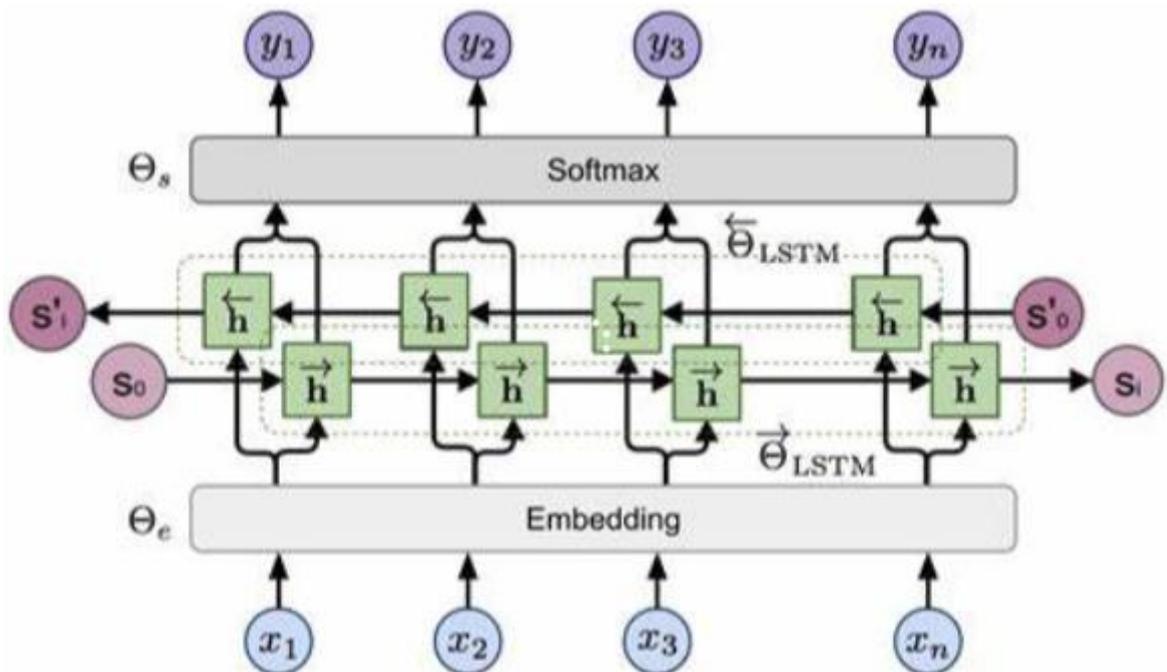


Figure 3.9 : l'architecture générale du modèle bi-lstm

**Feature Engineering:** Le Bi-LSTM utilisé nécessite des données de mapping aux vecteurs de nombres, pour lesquels plusieurs méthodes existent, nous avons

choisi le **Word embedding**. Il existe plusieurs techniques de Word Embedding. Nous nous intéressons au FastText.

**Couche Bi-LSTM :** Les réseaux de neurones Bi-LSTM sont une famille de réseaux de neurones approfondis qui fonctionnent sur des données séquentielles. Chaque modèle Deep Learning y compris le Bi-LSTM, comporte plusieurs paramètres et hyper paramètres considérés comme des éléments caractéristiques à définir, ils peuvent rendre le modèle beaucoup plus puissant selon les valeurs attribuées.

Les paramètres sont les coefficients que le modèle choisit lui-même pendant l'étude, les optimise à mesure que la formation progresse, et retourne un tableau de paramètres de minimisation.

D'autre part, les paramètres hypers, doivent être définis par l'administrateur, ils ne seront pas mis à jour par le modèle.

- **Fonction de perte (Loss) :** processus d'optimisation nécessite une fonction de perte dans le but de calculer l'erreur de modèle. En utilisant les réseaux de neurones, nous cherchons à minimiser les erreurs, Plus précisément, si la prédiction donnée s'écarte trop du résultat réel, alors La fonction de perte renvoie un très grand nombre. Progressivement, à l'aide d'une fonction optimisation, la fonction de perte apprend comment réduire l'erreur de prévision. Lors du calcul de l'erreur du modèle pendant le processus d'entraînement, une fonction de perte doit être choisie, dans notre cas, nous avons opté pour la fonction «CategoricalCrossentropy». C'est pour mesurer la distance probabilités des différentes classes et étiquettes des distributions observées, la formule mathématique de Categorical Cross Entropy est la suivante :

$$-\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C \mathbf{1}_{y_i \in C_c} \log p_{model}[y_i \in C_c]$$

Tel que :

- $N$ : est le nombre d'observations.

-  $C$ : nombre de classes.

-  $P$ : la probabilité de l'observation  $i$  par rapport à la classe  $c$ .

- **Taille des Lots (Batch size)** : Les ensembles de données sont également de petits exemples du dataset, sur lesquels l'algorithme s'entraîne, Le modèle pourra ainsi mieux s'adapter aux nouvelles données en estimant le gradient d'erreur avant d'actualiser les poids du modèle.
- **Nombre des nœuds LSTM** : Représente le nombre de cellules LSTM utilisées dans une couche cachée, L'idée est d'avoir un réseau neurone aussi simple que possible tout en classifiant correctement les données d'entrée.
- **Taux d'apprentissage (Learning rate)** : Le taux d'apprentissage est un hyper paramètre qui détermine la mesure dans laquelle les poids du réseau doivent être ajustés en fonction du gradient de perte.
- **Taille de l'embedding** : C'est la taille de l'espace vectoriel dans lequel les mots seront insérés.
- **Dropout** : cette technique consiste à éliminer une partie des nœuds qui alimentent une couche, pour que le NN ne soit pas trop lourd. Cela s'applique à la phase d'entraînement. Par exemple, si nous appliquons un dropout égal à 20% sur la couche d'entrées, cela signifie que 1 entrée sur 5 sera exclue de chaque cycle de mise à jour.

Hyper-paramètre	Fonction d'optimisation	Fonction loss	Nombre des nœuds lstm	Learning rate	Beatch size	Embedding size	Dropout
Définition	keras	Categorical Cross Entropy	248	0.0085	20	300	0.3

**Tableau 3.3 : configuration du hyper-paramètre du modèle bi-lstm**

**couche Softmax :** Les résultats de la couche « Bi-LSTM » passent par cette dernière couche, elle agit comme « classifieur », la fonction d'activation Softmax permet d'obtenir une distribution de probabilité sur les mots du vocabulaire. Le mot généré en sortie est celui qui possède la plus grande valeur de probabilité.

## 6. Conclusion

Dans ce chapitre Nous avons présenté l'architecture de notre assistant virtuel, commençant par l'architecture générale, **ensuite**, nous avons proposé trois modèles basés sur les approches du Deep Learning, à savoir un modèle LSTM Bidirectionnel (Bi-LSTM), un modèle Naive Bayes (MNV) et un troisième modèle le plus proches voisins (KNN) avec TF-IDF et FastText, en suite nous avons détaillé les modèles proposés pour générer les modèles finaux pour deux datasets distincts, l'un en arabe et l'autre en anglais.

Les résultats expérimentaux évalués sur des métriques automatiques démontrent que le modèle proposé Bi-LSTM apporte les meilleurs scores par rapport aux deux modèles proposés. Des évaluations humaines ont vérifié également l'efficacité de notre modèle.

Dans le chapitre suivant, nous allons présenter la mise en œuvre de notre système, et l'environnement lié à chaque étape de notre processus de développement.

# **Chapitre IV : Implémentation et évaluation**

## 1. Introduction :

Dans les chapitres précédents, nous avons parlé théoriquement des assistants virtuels en général, et notre assistant en particulier.

L'objectif principal de notre travail, est de proposer une assistante virtuelle, qui peut indiquer dans une phrase l'action, l'objet et la localisation à partir d'une phrase donnée par l'utilisateur.

Dans ce chapitre nous allons présenter le modèle final. Tout d'abord nous commençant par les langages utilisés ainsi les bibliothèques, ensuite nous allons détailler les démarches du travail, Par la suite, nous passons à la partie expérimentale où, nous illustrons nos jeux d'essai utilisés dans nos tests.

## 2. Environnement :

### 2.1. Le langage de programmation :

Python est un langage de programmation qui peut s'utiliser dans de nombreux contextes et s'adapter à tout type d'utilisation grâce à des bibliothèques spécialisées. Il est cependant particulièrement utilisé comme langage de script pour automatiser des tâches simples mais fastidieuses. On l'utilise également comme langage de développement de prototype lorsqu'on a besoin d'une application fonctionnelle avant de l'optimiser avec un langage de plus bas niveau. Il est particulièrement répandu dans le monde scientifique [43].

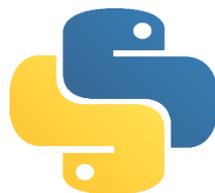


Figure 4.1 : logo de langage python[44]

## 2.2. Google drive :

Pour ce qui est du stockage des données, nous avons opté pour Google Drive. C'est un stockage en ligne, capable de partager le service qui offre 15 Go d'espace de stockage gratuit et une variété de paquets pour différents besoins de stockage[45]. En conséquence, l'ensemble de données est plus facilement accessible pour l'environnement d'exécution choisi, Google Colaboratory.



Figure 4.2 : logo de Google drive [46]

## 2.3. Google Colaboratory :

Colaboratory, souvent raccourci en "Colab", est un produit de Google Research. Colab permet à n'importe qui d'écrire et d'exécuter le code Python de son choix par le biais du navigateur. C'est un environnement particulièrement adapté à la machine Learning, à l'analyse de données et à l'éducation. En termes plus techniques, Colab est un service hébergé de notebooks Jupyter qui ne nécessite aucune configuration et permet d'accéder gratuitement à des ressources informatiques, dont des GPU.[47], Colab offre jusqu'à 35 Go de RAM et 107 Go d'espace disque, et un k80 TESLA GPU.



Figure 4.3 : logo de google colab [48]

## 2.4. Les bibliothèques :

Un autre point positif pour Google Colaboratory est la disponibilité de tous les pythons nécessaires bibliothèques utilisées pour le traitement audio/texte et les expériences d'apprentissage profond.

Voici quelques-unes des bibliothèques les plus utiles que nous avons utilisées dans le cadre de notre travail :

**Pandas** : Pandas est une librairie python qui permet de manipuler facilement des données à analyser manipuler des tableaux de données avec des étiquettes de variables (colonnes) et d'individus (lignes).ces tableaux sont appelés Data Frames. On peut facilement lire et écrire ces datagrammes à partir ou vers un fichier tabulé. On peut tracer des graphes à partir de ces Data Frames grâce à matplotlib [49].

**Matplotlib** est une bibliothèque du langage de programmation Python destinée à tracer et visualiser des données sous formes de graphique [50]. Il s'agit d'une bibliothèque très riche qui génère des figures que vous pouvez enregistrer sous les formats .png (fichier image) ou .pdf (portable document format) [51].

**NLTK** : Natural Language Toolkit (NLTK) est une boîte-à-outil permettant la création de programmes pour l'analyse de texte [52], NLTK est une bibliothèque open source modules, tutoriels et ensembles de problèmes, fournir des données informatiques prêtes à l'emploi dans des cours de linguistique[53], NLTK est écrit en Python et distribué sous GPL open licence source. Au cours des trois dernières années, NLTK est devenu populaire dans l'enseignement et la recherche [54].

**Tensorflow** est une bibliothèque open-source développée par l'équipe Google Brain qui l'utilisait initialement en interne. Elle implémente des méthodes d'apprentissage automatique basées sur le principe des réseaux de neurones profonds (Deep Learning). Une API Python est disponible. Nous pouvons l'exploiter directement dans un programme rédigé en Python. C'est faisable, il existe des tutoriels et des ouvrages à ce sujet. Pourtant, j'ai préféré passer par Keras parce que le formalisme

imposé par Tensorflow est déroutant au possible pour un néophyte. Découvrir de nouveaux algorithmes devient vite rédhibitoire si on a du mal à se dépatouiller avec un outil que nous sommes censés utiliser pour les mettre en application [55].

**Keras** est une librairie Python qui encapsule l'accès aux fonctions proposées par plusieurs librairies de machine learning, en particulier Tensorflow. De fait, Keras n'implémente pas nativement les méthodes. Elle sert d'interface avec Tensorflow simplement. Keras facilite grandement la vie en proposant des fonctions et procédures relativement simples à mettre en œuvre. Un apprenant qui a déjà assimilé les démarches types du machine learning, qui a pu par ailleurs utiliser des librairies qui font référence telles que scikit-learn, ne sera pas dépaysé lorsqu'il aura à travailler avec Keras[56].

**Speech recognition** : Bibliothèque pour la reconnaissance vocale, avec prise en charge de plusieurs moteurs et API, en ligne et hors ligne.[57], il est utilisé pour identifier les mots dans la langue parlée.

**Flask** : est un petit cadre selon la plupart des normes, assez petit pour être appelé un « micro », mais le fait d'être petit ne signifie pas qu'il fait moins que les autres cadres. Flask a été conçu comme un cadre extensible à partir de la terre; il fournit un noyau solide avec les services de base, tandis que les extensions fournissent le reste. Parce que vous pouvez choisir les paquets d'extension que vous voulez, vous vous retrouvez avec une pile maigre qui n'a pas de bloat et fait exactement ce dont vous avez besoin [58].



Figure 4.4 : logo Flask [59]

En plus des bibliothèques d'apprentissage profond mentionnées ci-dessus, nous avons utilisé le Numpy pour effectuer des opérations de manipulation sur nos données [39].

### 3. La démarche de travail :

Nous pouvons résumer les démarches de la réalisation montrées dans la figure 4.8 comme suit :

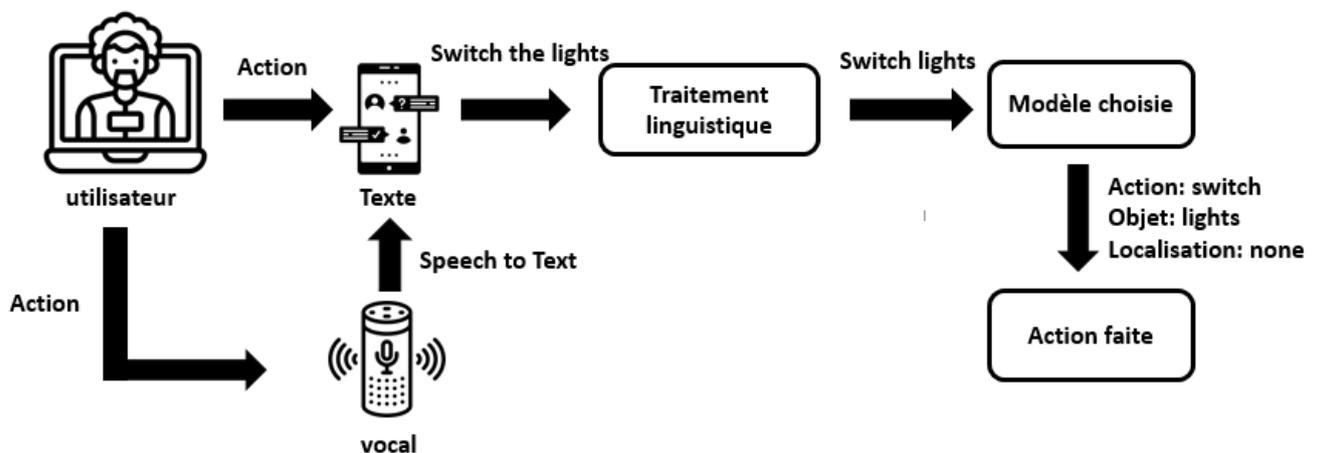


Figure 4.8 les démarches du travail

**3.1 Action écrit** : donner une action écrit pour notre assistant virtuel est la première étape du processus. Cette phase est très importante, car la compréhension de notre assistant dépend des mots. La reconnaissance du langage donne à l'assistant virtuel la capacité de reconnaître le sens global du texte saisi.

**3.2 Action vocal** : Pour activer son assistant vocal, l'utilisateur doit appuyer sur le bouton microphone  .

**3.3 Transformation de l'audio en texte** : "speech to Text" en anglais, cette étape transforme l'audio en texte. En effet, l'audio est décomposé sous forme de texte ce qui permet une analyse plus poussée.

**3.4 Traitement linguistique** : après avoir entré les actions (des mots, ou transformer les sons en mots) ici devient la deuxième étape, le traitement linguistique est de nettoyer la phrases et éliminer les mots vides.

**3.5 modèle choisis** : le but de notre modèle est d'extraire les 3 valeurs principales dans une phrase donnée :

- L'action : comment devrait faire l'action
- L'objet : sur quoi l'action devrait faire
- Localisation : où l'action devrait être faite

activate the kitchen lights		
Action	Objet	Location
Activate	lights	kitchen

**Tableau4.2 : exemple d'extraction l'action, l'objet et localisation**

Après avoir extrait ces 3 valeurs, l'action finale sera faite.

Après avoir exécuté notre algorithme nous avons obtenu ce modèle (figure4.5) qui schématise l'étape de prédiction.

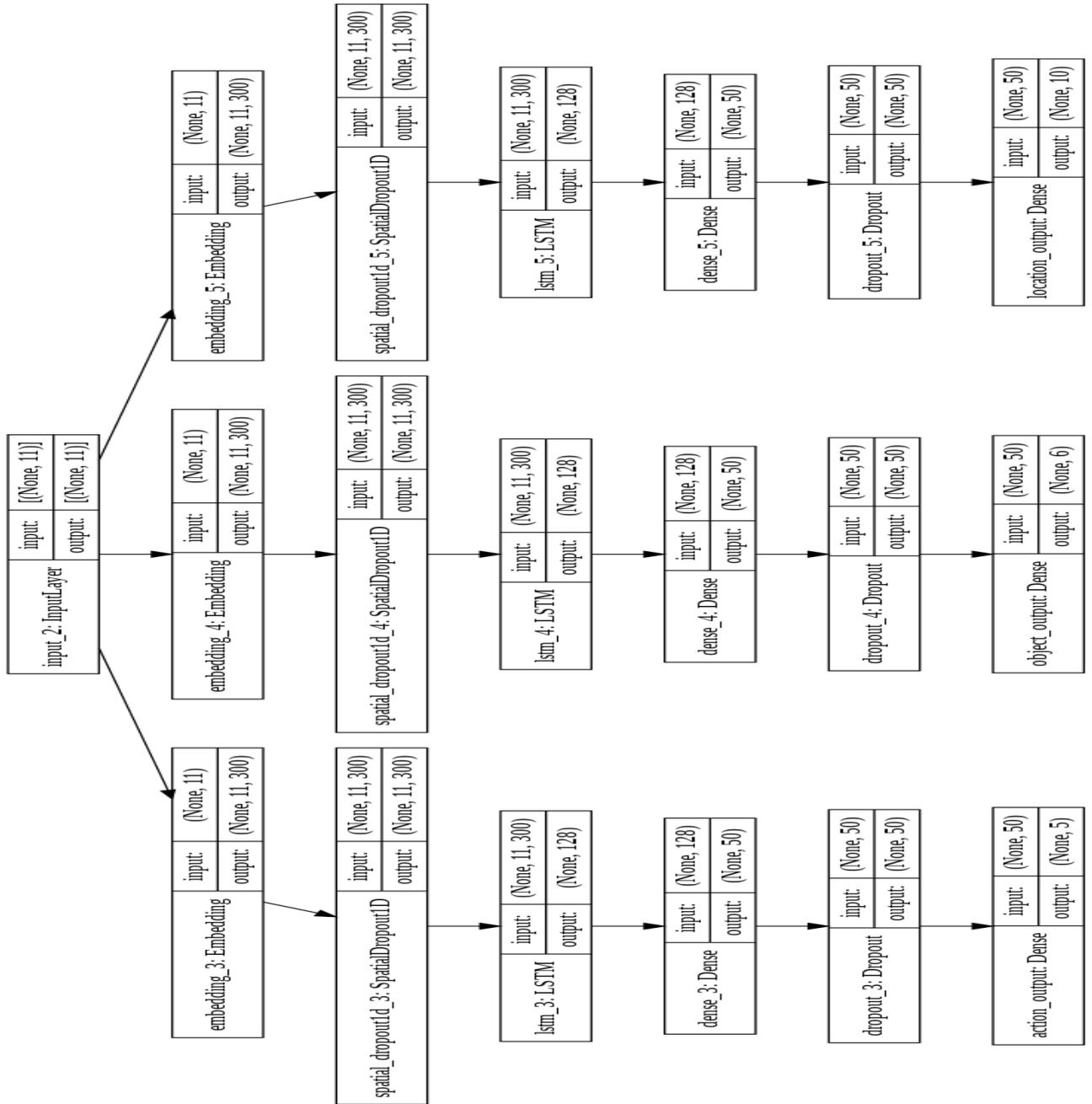
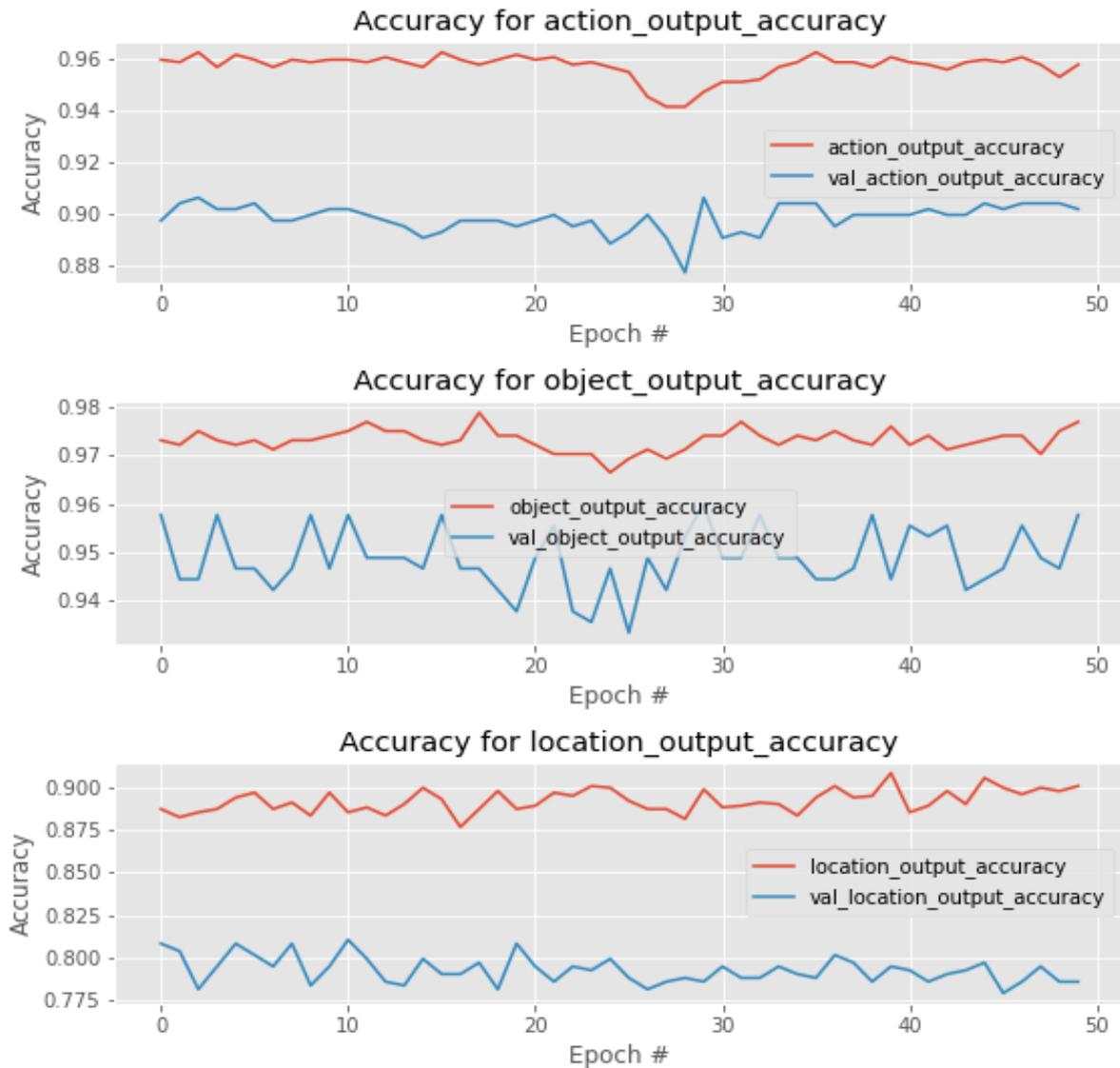


Figure4.5 schéma Bi-LSTM de notre modèle

Les résultats que nous avons obtenus de la validation de notre réseau de neurones LSTM sont présentés en figure 4.6



**Figure4.6 résultat accuracy**

Comme la montre la figure ci-dessus, nous avons testé la précision du Bi-LSTM et les résultats obtenus de l'évaluation sont :

	<b>validation</b>	<b>training</b>
<b>localisation</b>	<b>82%</b>	<b>91%</b>
<b>action</b>	<b>91%</b>	<b>96%</b>
<b>Object</b>	<b>96%</b>	<b>98%</b>

**Tableau 4.1 les résultats d'accuracy**

Figure 4.7 représente losses de chaque epoch.

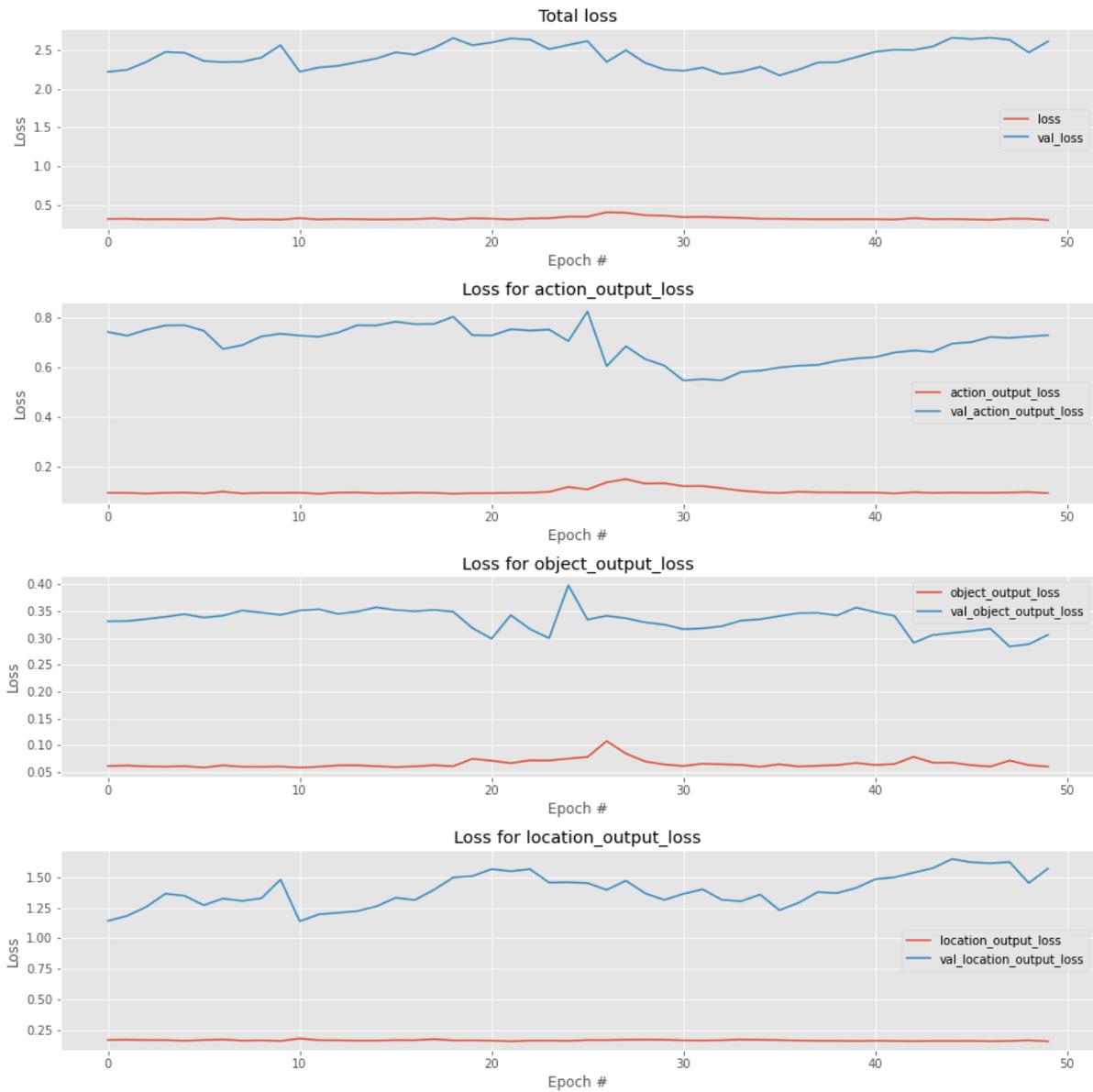


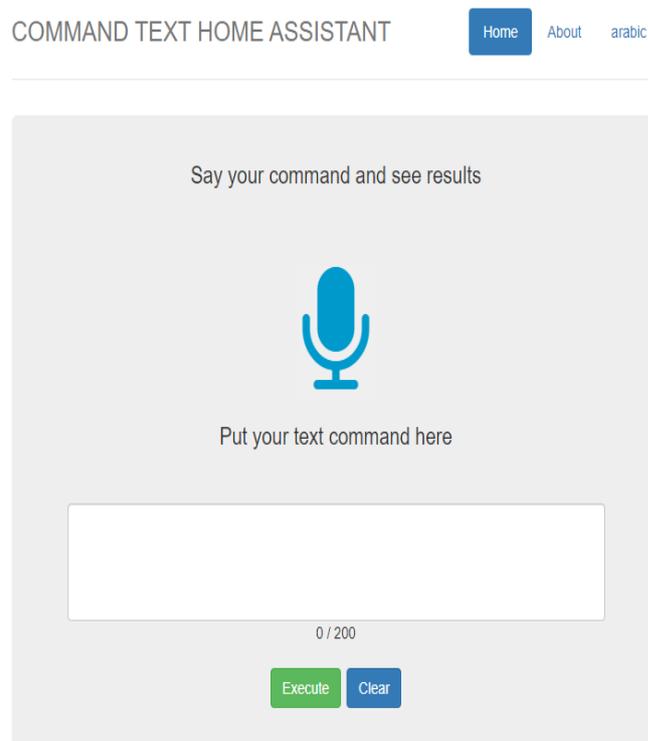
Figure 4.7 epoch losses

## 4. Réalisation du projet :

Nous assistant intelligent est présenté par les interfaces suivants :

### 4.1 Page d'accueil :

Toute personne utilisant notre application passe par page d'accueil (figure4.9)



**Figure 4.9 page d'accueil en anglais**

La figure 4.10 représente Page d'accueil en Arabe

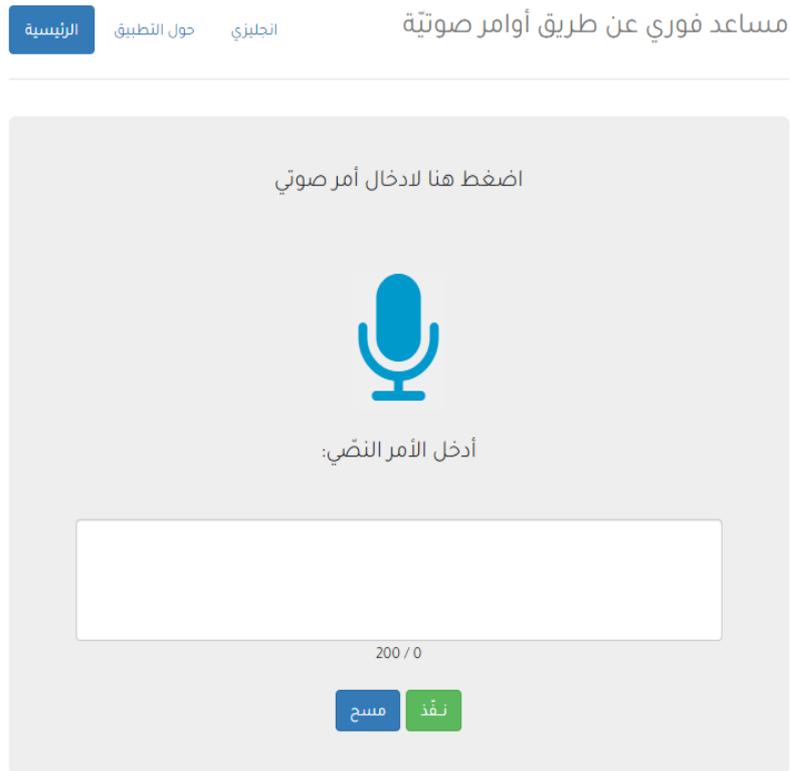


Figure 4.10 page d'accueil en arabe

Pour cette dernière section, nous allons montrer quelques exemples de prédiction que nous avons eue avec notre assistant virtuel dans les 2 langues.

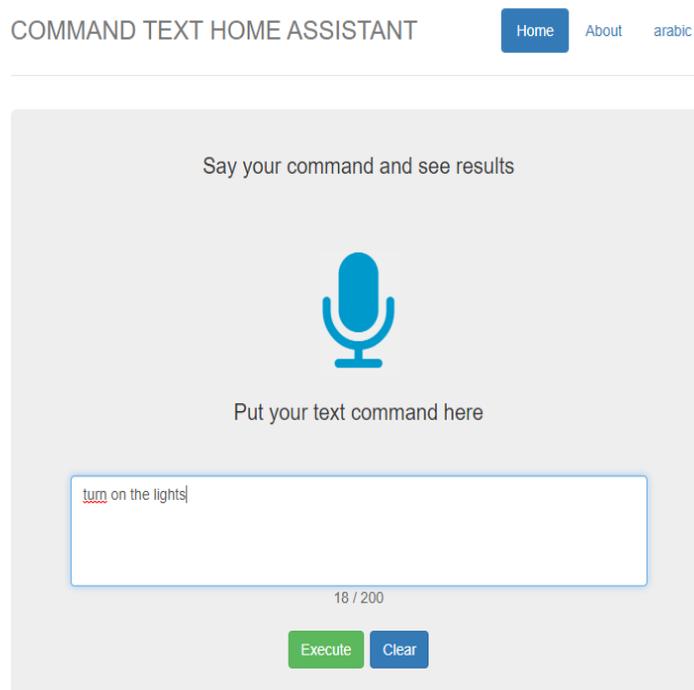
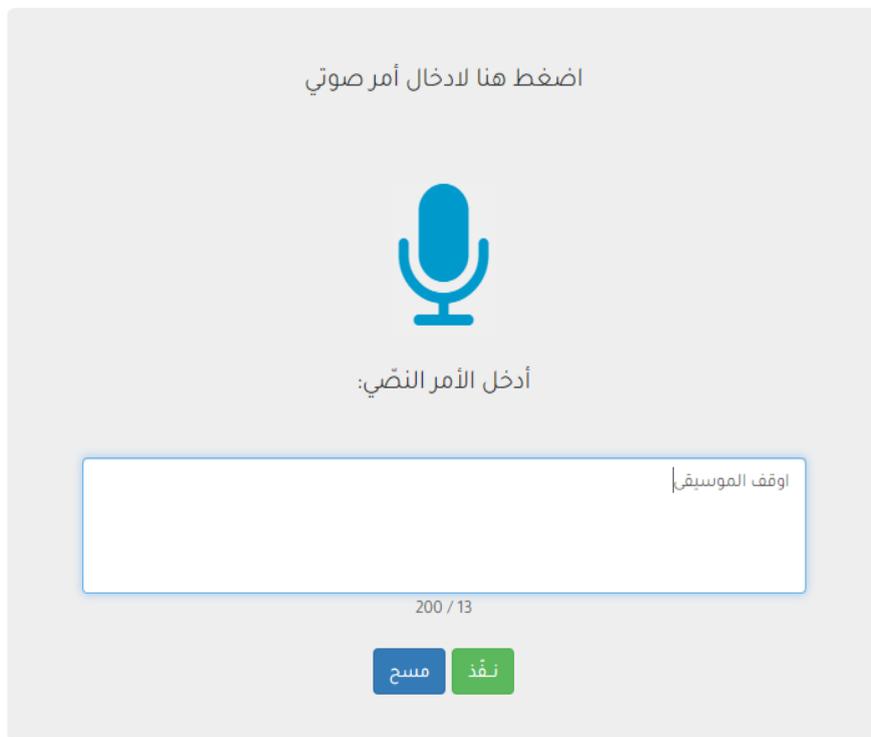


Figure 4.11: exécution d'une commande en anglais



Figure 4.12 résultat d'exécution en anglais



4.13 : exemple d'exécution en arabe



Figure4.14 résultat d'exécution en arabe

## 5. Test et résultat :

Les résultats que nous avons obtenus de la formation et de la validation de notre Bi-LSTM se résume par :

Notre travail est initialement dédié à la langue arabe, cependant nous avons utilisé l'Anglais comme langue de référence. Ceci nous permettra de comparer nos résultats avec les travaux existants et de valider notre modèle indépendamment de la langue arabe.

Chacun de nos trois modèles proposés est appliqué pour les deux langues arabe et anglais.

L'entraînement, la validation et les tests de chaque modèle ont été exécutés sur l'environnement Google Colaboratory (proposé par Google Research). Il permet aux développeurs d'implémenter des codes en python. Il est adapté beaucoup plus pour les tâches de Data Science. Cet environnement donne la possibilité d'utiliser un accélérateur graphique GPU afin d'optimiser le temps d'exécution.

Après avoir exécuter notre dataset avec MNB et KNN en utilisant **Word Level** , **N-gram Level** , **Character Level** et **FastText**. Nous avons obtenu les résultats suivant :

MNB	Word Level	Character Level	Character_wb Level	FastText
Action	96%	93%	93%	15%
objects	96%	94%	94%	15%
localisation	94%	90%	89%	9.7%

**Tableau 4.3: résultat mnb pour la langue anglais**

Pour la langue arabe :

MNB(Ar)	Word Level	Character Level	Character_wb Level	FastText
Action	89%	86%	87%	35%
objects	90%	94%	94%	47%
localisation	65%	68%	74%	26%

**Tableau 4.4: résultat mnb arabe**

KNN	Word Level	Character Level	Character_wb Level	FastText
Action	79%	94%	94%	60%
objects	79%	95%	95%	61%
localisation	72%	94%	94%	56%

**Tableau 4.5: résultat KNN anglais**

KNN(Ar)	Word Level	Character Level	Character_wb Level	FastText
Action	89%	90%	89%	66%
objects	94%	95%	94%	76%
localisation	47%	62%	68%	37%

**Tableau4.6 : resultat knn arabe**

Les tableaux ci-dessous représentent les résultats de notre base de données

LSTM	Word Level	Character Level	Character_wb Level	FastText
Action				91%
objects				96%
localisation				98%

**Tableau4.7: resultat lstm anglais**

LSTM(Ar)	Word Level	Character Level	Character_wb Level	FastText
Action				82%
objects				84%
localisation				81%

**Tableau 4.8: resultat lstm en arabe**

Pour le modèle KNN les résultats obtenus sont très mauvais pour toutes les métriques que ce soit sur l'arabe ou l'anglais par rapport les deux autres modèles proposés. Une amélioration est constatée pour le modèle MNB, Les valeurs arabes sont bien meilleures que les valeurs anglaises, mais cela n'appartient pas à l'intérêt de la recherche car, il est préférable de trouver des valeurs pour deux langues qui sont proches et élevées et c'est ce que nous trouvons dans le dernier modèle étudié.

Le modèle BI-LSTM a donné un bon score pour l'Arabe et l'anglais, Ceci peut être expliqué par rapport au corpus d'entraînement.

## **6. Conclusion :**

Dans ce chapitre, nous avons présenté notre mise en œuvre concernant la solution que nous avons proposée, y compris les outils que nous avons utilisés, ainsi que les différentes étapes de la mise en œuvre du nettoyage de l'ensemble de données jusqu'à la forme finale d'assistant. La partie suivante est une conclusion générale sur notre rétrospective de ce que nous avons réalisé dans ce tour de temps et ce que va être réalisé dans le futur pour améliorer et perfectionner notre système.

# **Conclusion générale**

## Conclusion générale

Le besoin d'assistant virtuel intelligent a considérablement augmenté ces dernières années, et on s'attend à ce qu'il augmente encore plus au cours des prochaines années. Cela est dû au fait que les assistants fournissent un moyen de communication rapide, entre les entreprises et les individus, ce qui profite aux deux parties.

Dans ce projet, notre objectif était de créer un assistant de maison intelligent fonctionnel capable d'avoir prédiqué le choix d'utilisateur, et assurer le bon fonctionnement de leur maison.

Notre travail consiste à proposer un système de classification. Ce système est conçu pour les maisons intelligentes dans les deux langues, anglais et arabe pour classer les prédictions selon l'action, l'objet et localisation de la commande.

Au début de cette recherche nous avons présenté les concepts fondamentaux d'intelligence artificielle, machine Learning et deep Learning, et la différence entre eux.

Dans l'état de l'art, nous avons fait une synthèse des approches des assistants intelligents existants et les travaux les plus populaires dans la domotique. Et nous avons montré le manque des assistants virtuels en arabe, En ce sens nous considérons que notre travail est d'explorer les techniques du deep Learning pour la création d'un assistant virtuel en arabe. Dans la troisième partie de notre approche, nous avons présenté l'architecture de notre système, et la conception de la solution que nous proposons, et détaillés des modèles utilisés, ainsi que le processus de développement.

A la fin de cette étude, nous avons démontré, l'environnement ainsi que les bibliothèques et les méthodes de création de notre modèle de prédiction. Nous avons choisi de développer le modèle Bi- LSTM comme une solution de base à laquelle nous pouvons nous référer pour apprécier l'amélioration avec un autre

modèle, Et nous avons présenté l'interface de notre application et quelques exemples d'exécution.

De nombreuses améliorations peuvent être apportées afin de rendre notre assistant plus précis

- Utiliser d'autres langues telque le dialecte Arabe par exemple.
- Agrandir la base de données, pour obtenir des résultats plus précis.
- Ajouter la classification multi-labeled pour qu'un seul message puisse contenir plusieurs intentions.

Il est important de noter que la réalisation de ce projet nous a été bénéfique sur tous les plans. Ce projet nous a été une bonne occasion pour découvrir et maîtriser les différents outils de programmation notamment python, c'était une introduction pour nous au domaine de l'apprentissage automatique, et cela ouvrira certainement des portes pour nous d'explorer.

## Bibliographie :

- [1] Abdelati.R, Ounayn.H, Kellili.R, El Abdi.N, Ismaili.N, Satar.M, Bouchouirbat.M, Goucheq.A. « Conception et réalisation d'une maison intelligente ». **Sep 2019.**
- [2] M. Tabaaa, K. Alamia, A. Dandacheb, Z. Benabboua, B. Chouria, K. Bousmara, Initiation à la conception et réalisation d'un réseau de capteurs intelligent : Maison Intelligente, 2017.
- [3] [online] [www.larousse.fr](http://www.larousse.fr)
- [4]– Stuart Russell, Peter Norvig ,Intelligence artificielle –, 2010 Pearson France.
- [5] Intelligence artificielle et société, mars 2019, [online]  
[https://www.academiesciences.fr/pdf/rapport/AI\\_G7\\_2019\\_FR.pdf](https://www.academiesciences.fr/pdf/rapport/AI_G7_2019_FR.pdf)
- [6] Yann LeCun, Qu'est-ce que l'intelligence artificielle ?, [online] [https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505\\_Intelligence\\_Artificielle\\_Y.\\_LeCun.pdf](https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505_Intelligence_Artificielle_Y._LeCun.pdf)
- [7] Ziad Mohamed, artificial intelligence definition, ethic and station 2018/2019, page2
- [8] Zara. I, L'intelligence artificielle principe, outils et objectifs, 2019,
- [9] comment fonctionne l'intelligence artificielle, zaion magazine électronique,
- [10] online: <https://www.docteurordinateur.com/l-intelligence-artificielle.html>
- [11]Valentin Blanchot, Histoire de l'intelligence artificielle, 2018
- [12] online: <https://www.freelance-provider.com/lintelligence-artificielle-ia-et-la-machine-learning/>
- [13]\_Kabir.S, BRB based Deep Learning Approach with Application in Sensor Data Streams, 2019.
- [14] Julien.A , Méthodes avancées en apprentissage supervisé et non-supervisé, Université M2 SISE 2020/2021
- [15] bastien L, 2021, data analytics ,IA,[online ]
- [16 ] online: <https://datakeen.co/3-deep-learning-architectures-explained-in-human-language/>

- [17] Sreelakshmi K, Vinayakumar R and Soman KP, Deep-Segregation of Plastic (DSP): Segregation of plastic and non-plastic using deep learning, juin 2019
- [18] online: <https://www.juripredis.com/fr/blog/id-19-demystifier-le-machine-learning-partie-2-les-reseaux-de-neurones-artificiels>
- [19] online: <https://towardsdatascience.com/an-introduction-to-recurrent-neural-networks-for-beginners-664d717adbd>
- [20] online: [https://fr.wikipedia.org/wiki/R%C3%A9seau\\_de\\_neurones\\_r%C3%A9currents](https://fr.wikipedia.org/wiki/R%C3%A9seau_de_neurones_r%C3%A9currents)
- [21] formation en ligne, Découvrez les cellules à mémoire interne : les LSTM, [online] : - <https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5814656-decouvrez-les-cellules-a-memoire-interne-les-lstm>
- [22] axel de Goursac, NLP au cœurs d'interaction humain- IA.
- [23] F. Fuchs & B. Habert, dir, le traitement automatique des langues : des modèles aux ressources, LXXII : 1, 2004
- [24] wikipedia, virtuel assistant [online] : [https://en.wikipedia.org/wiki/Virtual\\_assistant](https://en.wikipedia.org/wiki/Virtual_assistant)
- [25] Vala. A , Les assistants virtuels alimentés par l'IA et l'avenir du travail, 2021, [online] : <https://www.zdnet.fr/actualites/les-assistants-virtuels-alimentes-par-l-ia-et-l-avenir-du-travail-39920799.htm>
- [26] [Online] : <https://loupfunds.com/annual-digital-assistant-iq-test/>
- [27] Assistants Vocaux Et Enceintes Connectées l'impact de la voix sur l'offre et les usages culturels et médias, Mai 2019, [online] :
- [28] [online] : <https://itsolutions.center/tip/%D8%B3%D9%84%D9%85%D9%89%D8%A7%D9%84%D9%85%D8%B3%D8%A7%D8%B9%D8%AF-%D8%A7%D9%84%D8%B4%D8%AE%D8%B5%D9%8A/>
- [29] [online]: page official d'application rafiQ: <https://www.rafiq-assistant.com/>
- [30] A guide to home automation and smart home assistants, [online]: <https://www.clickatell.com/articles/technology/home-automation-smart-home-assistants/>

- [31] [online]: page official de google:  
<https://support.google.com/chromecast/answer/7071794?hl=fr&co=GENIE.Platform%3DAndroid>
- [32] himang.S, Topic and Intent Classifier from Scratch, 2018.
- [33] fluant.ai
- [34] [www.quilbot.com](http://www.quilbot.com), site de paraphrasing
- [35] [online dataset]: <https://arxiv.org/abs/1805.10190>
- [36] www.reverso.net, site de la traduction
- [37] translate.google.com
- [38] recherche d'information l'indexation, Philippe Rigaux, 2014
- [39] Blanca Li, Module Normaliser des données, 2020.
- [40] Understanding Word Embeddings, [online] : <https://towardsml.com/2018/06/12/understanding-word-embeddings/>
- [41] Recherche d'information cours 3ème année Licence Informatique – ISIL
- [42] Génération automatique de définitions pour le français, Timothee Mickus, Mathieu Constant, Denis Paperno, juin 2020
- [43] wikipedia, [online] : [https://fr.wikipedia.org/wiki/Python\\_\(langage\)](https://fr.wikipedia.org/wiki/Python_(langage))
- [44] <https://fr.wikipedia.org/wiki/Fichier:Python-logo-notext.svg>
- [45] <https://support.google.com/drive>
- [46] [https://fr.wikipedia.org/wiki/Fichier:Google\\_Drive\\_logo.png](https://fr.wikipedia.org/wiki/Fichier:Google_Drive_logo.png)
- [47] ] [site official google collab](https://research.google.com/colaboratory/faq.html?hl=fr), [online]:  
<https://research.google.com/colaboratory/faq.html?hl=fr>
- [48] [https://commons.wikimedia.org/wiki/File:Google\\_Colaboratory\\_SVG\\_Logo.svg](https://commons.wikimedia.org/wiki/File:Google_Colaboratory_SVG_Logo.svg)
- [49] [online]: <http://www.python-simple.com/python-pandas/panda-intro.php>
- [50] Sandro.T, Matplotlib for Python Developers, 2009, [online ]: « Matplotlib for Python Developers – Preface)

[51] Matplotlib, [online] : <https://cahier-de-prepa.fr/mp1-janson/download?id=526>

[52]: Abder.R, Introduction au Natural Language Toolkit (NLTK), 2017.

[53]: Edward.L, NLTK: the Natural Language Toolkit, 2002.

[54] Bird.S, NLTK: The natural language toolkit, 2006.

[55] Tanagra Data Mining, online pdf : [http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/fr\\_Tanagra\\_Tensorflow\\_Keras\\_Python.pdf](http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/fr_Tanagra_Tensorflow_Keras_Python.pdf)

[56] Deep Learning avec Tensorflow et Keras, Tanagra, avril 2013

[57] [online] <https://pypi.org/project/SpeechRecognition/>

[58] Grinberg.M, flask Web Développement, 2014.

[59] <https://foutucode.fr/les-frameworks-python-web>