

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE MINISTERE DE
L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED BOUDIAF - M'SILA

FACULTE DES MATHÉMATIQUES

ET DE L'INFORMATIQUE

DÉPARTEMENT D'INFORMATIQUE

**MEMOIRE de fin d'études Présenté pour l'obtention du
diplôme de MASTER**

Domaine : Mathématiques et Informatique Filière : Informatique

Spécialité : Informatique Décisionnel et Optimisation

Par: OULD MEHIEDDINE , HERIZI

Sarra , Khadidja

Thème:

**Développement d'un système intelligent pour
prédire la satisfaction envers une agence touristique
en se basant sur les avis des clients.**

Soutenu devant le jury composé de :

Mr. TARRAFI A	Université de Msila	Président
Mr. KADRI Said	Université de Msila	Superviseur
Ms. MELIOUHE Amel	Université de Msila	Examineur

Promotion : 2020 /2021

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTÈRE DE
L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED BOUDIAF - M'SILA

FACULTE DES MATHÉMATIQUES

ET DE L'INFORMATIQUE

DÉPARTEMENT D'INFORMATIQUE

**MEMOIRE de fin d'études Présenté pour l'obtention du
diplôme de MASTER**

Domaine : Mathématiques et Informatique Filière : Informatique

Spécialité : Informatique Décisionnel et Optimisation

Par: OULD MEHIEDDINE , HERIZI

Sarra , Khadidja

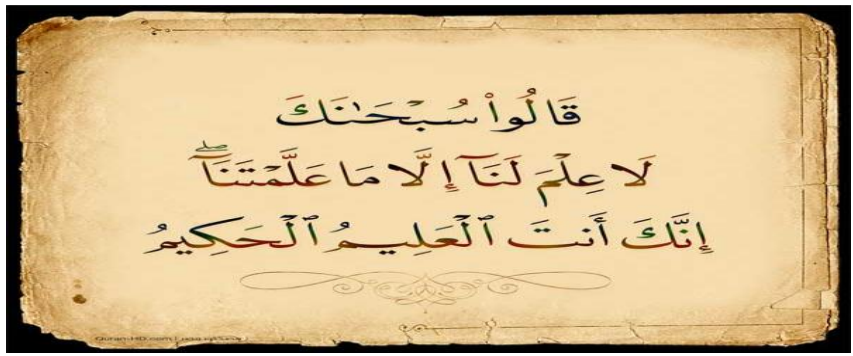
Thème:

**Développement d'un système intelligent pour
prédire la satisfaction envers une agence touristique
en se basant sur les avis des clients.**

Soutenu devant le jury composé de :

Mr. TARRAFI A	Université de Msila	Président
Mr. KADRI Said	Université de Msila	Superviseur
Ms. MELIOUHE Amel	Université de Msila	Examineur

Promotion : 2020 /2021



Dédicaces

À notre Noble Messenger, notre Maître Muhammad, que les prières et la paix de Dieu soient sur lui.....Au phare de la connaissance et à l'Imam choisi.

A ceux qui m'ont appris à gravir les échelons de la vie avec sagesse et patience

A ceux qui cherchent à me pousser sur la voie de la réussite à mon cher père.

A qui tu es resté jusqu'à ce que j'atteigne ma force

A celle au grand cœur qui m'influence sur elle A..... ma chère maman.

A ceux qui portent fièrement mon nom avec moi A ceux qui me touchent en les mentionnant à mes frères et sœurs.

A qui vous avez marché ensemble sur le chemin

A qui nous nous sommes joints main dans la main pour la réussite A tous les collègues (à mes mères....à mes colombes.....à nos amis en particulier notre amie Khadîdja Ferhati sur Tous les encouragements pour toute aide)

A ceux qui nous ont appris que la connaissance est une lumière A ceux qui ont formulé pour nous leur connaissance en lettres d'or A nos honorables professeurs.



Remerciements

*En premier lieu je remercie Le Bon Dieu
le tout-puissant de nous avoir donné
la force ainsi
que la patience d'accomplir ce modeste
travail.*

A Notre Encadreur

Mr KADRI SAID

Votre compétence, votre supervision ont toujours suscité notre plus profond respect.

*Merci pour votre accueil et vos conseils, ce fut un honneur de faire partie de vos
étudiants et de bénéficier de votre riche éducation.*

*Vos qualités éducatives et humanitaires sont pour nous un modèle. Nous sommes
toujours impressionnés par votre gentillesse et votre présence.*

*Veillez accepter nos remerciements, monsieur, pour le grand honneur que vous
nous avez accordé en acceptant de superviser ce travail.*

Aux membres du jury

Président du Jury : Tarafi A, Meliouh Amel.

*Chers membres du jury, nous sommes très honorés de votre consentement à juger ce
travail. Merci pour votre accueil et vos conseils.*

Table des matières

<i>Dédicaces</i>	I
<i>Remerciements</i>	II
Liste des figures	V
Liste des tableaux	VI
Introduction générale	2

Chapitre1:Analyse de sentiments sur le tourisme

1.1. Introduction	5
1.2. Définition de l'analyse des sentiments	5
1.3. Taches de l'analyse des sentiments	6
1.4. Les Caractéristiques de l'analyse des sentiments	7
1.4.1. Catégorisation des sentiments : phrases objectives versus phrases subjectives	7
1.4.2. Les niveaux de l'analyse des sentiments	7
1.4.2.1. Niveau du document	8
1.4.2.2. Niveau de phrase	8
1.4.2.3. Niveau de la substance et de l'apparence	8
1.5. Méthodes de l'analyse de sentiment	8
1.5.1. Analyse lexicale	8
1.5.2. Apprentissage automatique	9
1.6. Applications et perspectives	9
1.6.1. prédiction	9
1.6.1.1. Résultat d'une élection	9
1.6.1.2. Succès d'un film	10
1.6.1.3. Nouveau produit	10
1.6.2. Supervision	10
1.6.2.1. Utilisation en entreprise	10
1.7. Conclusion	10

Chapitre2: Intelligence artificielle, l'apprentissage automatique et l'apprentissage profond

2.1. Introduction	12
2.2. L'intelligence artificielle	12
2.3.1. Définition	13
2.3.2. Types d'apprentissage automatique	13
2.3.2.1. Apprentissage Supervisé	14
2.3.2.2. Apprentissage par renforcement (Reinforcement Learning)	15
2.3.3. Algorithmes utilisés	15
2.3.3.1. Les machines à vecteur de support SVM	15
2.3.3.2. La régression logistique	15
2.3.3.3. La méthode des k plus proches voisins pour un apprentissage supervisé	16
2.3.3.4. Analyse en composantes principales	16
2.4. L'apprentissage profond	17
2.4.1. Définition	17
2.4.2. Un classement en trois catégories du DL	17
2.4.2.1. Architectures profondes génératives	17
2.4.2.2. Architectures profondes discriminative	17
2.4.2.3. Architectures hybrides profondes	17
2.4.3. Modèles d'apprentissage en profond	18
2.4.3.1. Réseaux de neurones artificiels (ANN) :	18
2.4.3.2. Auto-Encodeurs empilés	18
2.5. Techniques d'apprentissage automatique et d'apprentissage en profondeur	19
2.6. Conclusion	20

Chapitre3: Environnement de Programmation

3.1. Introduction	22
3.2. Présentation des outils et des bibliothèques	22
3.2.1. Le hardware	22
3.2.2. Les outils et les bibliothèques utilisés	23
3.2.2.1. Python	23
3.2.2.1. Top des meilleures bibliothèques et packages	25
3.3. IDLE	29
3.4. C3.2.2.3. Anaconda	27
Conclusion	29

Chapitre: Application réalisée et Résultats obtenus

4.1. Introduction	32
4.2. Problématique de notre travail	32
4.3. Prétraitement du dataset utilisé	32
4.4. Le modèle proposé	33
4.5. Evaluation	34
4.6. Interface de l'application	39
4.7. Page de démarrage	39
4.8. Page de d'accueil	40
4.10. Conclusion	41
Résumé: 45	
BIBLIOGRAPHIE	45

Liste des figures

Fig 1.1: Classes Analyse du sentiment.....	6
Fig 1.2: Flux des tâches de l'analyse des sentiments.....	9
Fig 2.1: L'intelligence artificielle et l'apprentissage automatique et l'apprentissage profond.....	9
Fig 2.2 : classification et régression	11
Fig 2.3: Les « vecteurs de support » SVM.....	11
Fig 2.4: La régression logistique	11
Fig 2.5: Méthode des k plus proches voisins.....	11
Fig 2.6: Analyse en composantes principales.....	11
Fig 2.7: Architecture de réseau de neurones convolutifs.....	12
Fig 2.8 : Auto-Encoder structure	13
Fig 3.1: Le hardware... ..	11
Fig 3.2 : Le python	12
Fig 3.3 : Le Anaconda	13
Fig 4.1 : Architecture du modèle ANN proposé	13
Fig 4.2: Training loss Vs Validation loss of the ANN modele	12
Fig 4.3: Training accuracy Vs Validation accuracy of the ANN model	13
Fig4.4: Accuracy VS Val Accuracy & Loss VS Val loss.....	13
Fig4.5: page de démarrage.....	13
Fig4.6: page d'accueil.....	12
Fig4.7:page À propos.....	12

Liste des tableaux

Table 2.1: Un apprentissage automatique Vs L'apprentissage profond.....	24
Table 2.1: Un apprentissage automatique Vs L'apprentissage profond.....	24
Table 4.1: Loss et accuracy obtenus.....	50
Table 4.2: Description du modèle ANN proposé.....	50



Introduction générale

Introduction générale

Au cours de la dernière décennie, l'intelligence artificielle (IA) est devenue un sujet populaire à la fois à l'intérieur et à l'extérieur de la communauté scientifique.

En 1956, un groupe d'informaticiens a suggéré que les ordinateurs pourraient être programmés pour penser et raisonner, "que chaque aspect de l'apprentissage ou toute autre caractéristique de l'intelligence [peut], en principe, être décrite avec une telle précision qu'une machine [peut] être conçue pour la simuler. " Ils ont décrit ce principe comme «l'intelligence artificielle».

De nombreux articles dans des revues technologiques et non technologiques ont couvert les sujets de l'apprentissage automatique (ML), de l'apprentissage profond (DL) et de l'IA [1–6]. Pourtant, la confusion persiste autour de l'IA, du ML et de la DL. Les termes sont fortement associés, mais ne sont pas interchangeables. Apprentissage profond appelé apprentissage neuronal profond ou réseau neuronal profond. Il s'agit d'un sous-ensemble de l'apprentissage automatique qui utilise des réseaux de neurones avec de nombreuses couches et qui a gagné une immense, en particulier au cours des dernières années pour résoudre des tâches complexes avec de grandes quantités de données., tandis que l'apprentissage automatique est basé sur l'idée que les machines devraient pouvoir apprendre et s'adapter par l'expérience, l'IA fait référence à une idée plus large où les machines peuvent exécuter des tâches «intelligemment». L'intelligence artificielle applique l'apprentissage automatique, l'apprentissage en profondeur et d'autres techniques pour résoudre des problèmes réels. En termes simples, l'intelligence artificielle est un domaine qui se concentre sur l'automatisation des tâches intellectuelles que les humains exécutent généralement, et ML et DL sont deux moyens spécifiques d'atteindre cet objectif .Dans un autre sens, tout l'apprentissage automatique est de l'IA, mais toute l'IA n'est pas de l'apprentissage automatique et tout l'apprentissage profond est de l'apprentissage automatique, mais tout l'apprentissage automatique n'est pas un apprentissage profond. Les gens comptent sur les opinions des autres pour prendre des décisions, surtout s'ils font partie de leur cercle de confiance. De plus, il existe des tonnes de sites Web très respectés qui donnent l'opinion des gens sur les lieux qu'ils ont visités en termes de services et d'hôtels, et que des millions de personnes lisent avant de prendre une décision. C'est pourquoi les systèmes d'analyse des sentiments sont de plus en plus importants pour traiter automatiquement les informations et identifier les sentiments des utilisateurs. Ils analysent leurs mots-clés saisis, qui sont généralement conditionnés par les caractéristiques des plates-formes de microblogging, car un grand nombre

de messages sont postés chaque jour, fournissant une excellente source d'informations qui ne peuvent pas être gérées manuellement. Par conséquent, nous utilisons une méthode d'apprentissage automatique basée sur un algorithme pour analyser les données, en tirer des leçons et prendre des décisions en fonction de vos informations. Parmi les algorithmes figurent: Support Vector Machine, K-Nearest Neighbor (KNN) et Naive Bayes.

Dans ce travail, nous avons choisi d'articuler notre étude autour de quatre chapitres principaux:

Le premier chapitre intitulé "Analyse des Sentiments sur le Tourisme " définit Analyser des sentiments et leur relation avec les touristes.

Le chapitre 2 est intitulé " Intelligence Artificielle, Apprentissage automatique et apprentissage profond" définit les Types d'apprentissage automatique et Algorithmes utilisés et les catégories du DL et modèles d'apprentissage en profond.

Le chapitre 3 décrit "l'environnement de programmation " notamment la plate-forme Anaconda qui travaille avec un langage python et les bibliothèques les plus importants que nous devons installer.

Enfin, dans le dernier chapitre, nous avons présenté l'application réalisée.



Chapitre1:
**Analyse des
sentiments pour
le tourisme**

1.1. Introduction

Les logiciels d'analyse des sentiments est un élément clé de la recherche sur les mégadonnées touristiques pour sa capacité à détecter les opinions positives et négatives dans le texte. L'analyse des sentiments est raisonnablement rapide et précise, permettant d'extraire des modèles d'un grand nombre de textes qui ne seraient pas évidents pour les experts lisant ces mots et il est donc devenu nécessaire de l'utiliser dans le domaine du tourisme pour prendre la bonne décision nous passons en revue le sujet de l'analyse des sentiments dans le tourisme en étudiant les points principaux.

1.2. Définition de l'analyse des sentiments

L'analyse des sentiments est une technologie courante qui utilise des stratégies d'analyse des médias sociaux pour analyser les réactions et les commentaires des clients. Il est souhaitable de réaliser une analyse des sentiments à partir de sites Web tels que Trip Advisor, car un grand nombre de datasets gratuits peuvent être obtenus à partir de ces sites Web pour des recherches à grande échelle, alors que ces données à grande échelle ne peuvent pas être obtenues facilement par les méthodes de recherche traditionnelles. Le big data fournit un nouveau type de données à utiliser dans la recherche sur le tourisme, et pose également des exigences plus élevées en matière de traitement des données. Actuellement, peu d'études ont été menées sur l'applicabilité et la précision des méthodes d'analyse des sentiments dans la littérature de recherche sur le tourisme. En outre, les recherches contemporaines ignorent la possibilité d'intégrer des connaissances humaines, telles que les graphes de connaissances, dans les méthodes existantes afin d'améliorer les performances de l'analyse des sentiments des textes. Le Big Data se caractérise par un énorme volume de données, et les exigences en matière de vitesse et de précision de l'analyse des sentiments sont de plus en plus élevées [8]. Par conséquent, la perspective de développer des méthodes d'analyse des sentiments adaptées et efficaces pour des types spécifiques de big data dans le contexte du tourisme est une proposition très intéressante. [1] [2] [3] [4]



Fig 1.1: Les Classes de l'analyse des sentiments.

1.3. Taches de l'analyse des sentiments

L'analyse des sentiments est l'un des développements les plus intéressants de ces dernières années. De nos jours, l'analyse des sentiments a gagné en valeur avec l'émergence des réseaux sociaux, qui attirent l'intérêt des chercheurs, des journalistes, des entreprises et des gouvernements. Elle est devenue l'un des sujets les plus intrigants de ces dernières années et l'un des sujets qui suscitent l'intérêt des chercheurs. Les développements technologiques actuels permettent à un fichier de stocker et de récupérer une énorme quantité de données, l'accent est désormais mis sur les méthodes d'extraction. Créer des informations et des connaissances à partir de sources brutes. En effet, la mise en réseau des réseaux sociaux représente un nouveau secteur catalytique dans le contexte du big data: les expressions en langage naturel des personnes peuvent être facilement signalées par des messages texte, un contenu unique est rapidement créé avec des dimensions massives qui doivent être analysées à partir d'une méthode efficace pour créer des actions exploitables. Connaissances pour les processus décisionnels résolution.

Cependant, l'analyse des sentiments est souvent utilisée de manière incorrecte lorsque l'on parle de la classification de la polarité, et il s'agit plutôt d'une sous-tâche visant à découvrir la polarité des textes à partir de sentiments positifs, négatifs ou neutres. Elle peut également avoir une polarité neutre comme la plupart des travaux de l'analyse des sentiments n'est généralement pas supposée. Seulement des émotions positives et négatives pour des raisons de simplicité.

1.4. Les Caractéristiques de l'analyse des sentiments

L'analyse des sentiments est un domaine de recherche vaste et complexe, et les principales caractéristiques qui composent le processus d'analyse des sentiments sont décrites et discutées en détail.

1.4.1. Catégorisation des sentiments : phrases objectives versus phrases subjectives

La distinction entre les phrases subjectives et les phrases objectives est le premier objectif de l'analyse des sentiments. Aucune autre tâche essentielle n'est requise, et lorsqu'elle est classée comme subjective, sa polarité (positive, négative ou neutre) doit être estimée (voir la fig 1.2).

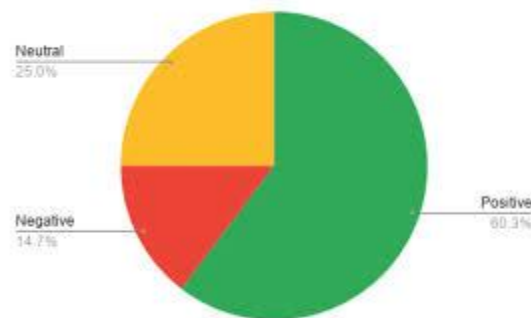


Fig 1.2:Flux des tâches de l'analyse des sentiments.

1.4.2. Les niveaux de l'analyse des sentiments

Le but de l'analyse des sentiments est de «définir» la première option pour appliquer l'analyse des sentiments est de définir le sens du texte que nous analysons

dans le cas étudié. L'analyse des sentiments peut être étudiée dans les réseaux sociaux Il est divisé en trois niveaux.

1.4.2.1. Niveau du document

Détermine l'opinion générale de l'ensemble du document. Cette analyse fonctionne bien pour des documents qui présentent un point de vue précis, mais moins pour des comparaisons car elle ne fera pas la différence entre les sujets abordés.

1.4.2.2. Niveau de phrase

A travers cela, cette analyse peut donner une mesure de la « neutralité » du texte, au niveau de la phrase: détermine l'opinion générale de la phrase (positive, négative ou neutre), par exemple pour analyser l'entrée de Wikipédia. Les méthodes utilisées sont celles d'auto-analyse.

1.4.2.3. Niveau de la substance et de l'apparence

Basé sur le point de vue suivant: L'opinion est composée de sentiment et de but (opinion). Effectuez une analyse plus approfondie des niveaux de texte et de phrase. Par exemple, la phrase "Les i Phones sont bons, mais ils doivent encore travailler dur en termes d'autonomie et de sécurité de la batterie" évalue trois aspects: i Phone (neutre), autonomie de la batterie (négatif), sécurité (négatif)

1.5. Méthodes de l'analyse de sentiment

Il existe deux grandes catégories d'analyse : l'analyse lexicale et l'analyse par apprentissage automatique. Cependant, il existe des outils qui tirent profit de ces deux méthodes. [5]

1.5.1. Analyse lexicale

L'approche de l'analyse lexicale consiste à extrapoler l'émotion qui émane de la phrase par l'analyse sémantique des mots. Nous utilisons des dictionnaires qui désignent les mots annotés avec une polarité et le contexte dans lequel ils conviennent. Parce que cette approche consiste à classer la phrase par des exemples de phrases qui existent déjà et pour lesquelles les émotions sont déjà identifiées. [6]

1.5.2. Apprentissage automatique

Les principales méthodes de classification de mots sont basées sur les algorithmes suivants:

- Classification naïve bayésienne ;
- Principe d'entropie maximale ;
- Machine à vecteurs de support.

L'algorithme le plus efficace semble être la machine à vecteurs de support. [7]

1.6. Applications et perspectives

L'analyse de sentiments peut trouver de nombreuses applications dans le domaine de la prédiction et de la supervision.

1.6.1. prédiction

Evolution des indices boursiers tels que le NASDAQ ou le DOW Jones : Toutes les études sur ce sujet semblent montrer une forte corrélation entre l'analyse et l'évolution des valeurs boursières, car une étude montre que le taux d'émotion sur les sites (espoir, peur, joie) était proportionnelle au mouvement des indices boursiers (plus les gens devenaient plus silencieux, plus les indices boursiers augmentaient) dans l'ensemble. [8]

1.6.1.1. Résultat d'une élection

Plusieurs études ont été réalisées Une étude a montré une forte corrélation entre les estimations basées sur les données de Google Trends et les résultats de plusieurs élections. [8]

D'une part, d'autres études offrent un bilan plus mitigé: c'est-à-dire la présence de tentatives infructueuses de prédiction qui montre que les mesures, bien que stables, sont souvent trop imprécises pour être fiables. [10]

La représentation démographique disproportionnée des internautes par rapport au nombre total d'électeurs, les jeunes étant plus actifs sur les réseaux sociaux que les adultes L'âge conduit en particulier à des différences. [11]

1.6.1.2. Succès d'un film

Dans son modèle de prédiction Les analyses montrent un haut degré de précision, qui peut surpasser les modèles actuels. [12]

1.6.1.3. Nouveau produit

Enfin, la prospection peut être utilisée pour prédire le succès d'un produit avant son lancement. [13]

1.6.2. Supervision

1.6.2.1. Utilisation en entreprise

Des entreprises ont montré qu'elles mettaient en œuvre des outils permettant de récupérer des informations sur leurs réputations en exploitant les flux de données publics exposés sur les forums et réseaux sociaux . Certaines proposent notamment des outils plus avancés de façon à superviser le sentiment dégagé par leurs clients pour ainsi améliorer leur communication. . [14] [15]

1.7. Conclusion

Dans ce chapitre, nous avons présenté la définition de l'analyse des sentiments et ses caractéristiques ainsi que ses niveaux et comment utiliser l'analyse des sentiments dans le domaine du tourisme.

Chapitre2:
**Intelligence
artificielle,
l'apprentissage
automatique et
l'apprentissage
profond**

2.1. Introduction

Au cours de la dernière décennie, l'intelligence artificielle (IA) est devenue un sujet brûlant au sein et en dehors de la communauté scientifique; Plusieurs articles dans des revues techniques et non techniques ont abordé les sujets de l'apprentissage automatique (ML), de l'apprentissage profond (DL) et de l'intelligence artificielle. [2-1] Malgré cela, il existe de nombreuses idées fausses sur le concept d'intelligence artificielle, d'apprentissage automatique et de DL. Les termes sont étroitement liés, mais ils ne sont pas interchangeables. En d'autres termes, tout le machine learning est de l'intelligence artificielle, mais pas toute l'intelligence artificielle est un apprentissage automatique et tout l'apprentissage profond est un apprentissage automatique, mais tout l'apprentissage automatique est pas du l'apprentissage profond. Dans ce chapitre (nous essayons) de lever le conflit entre ces termes techniques pour mieux expliquer ces concepts et faire circuler un groupe d'idées sur ces concepts.

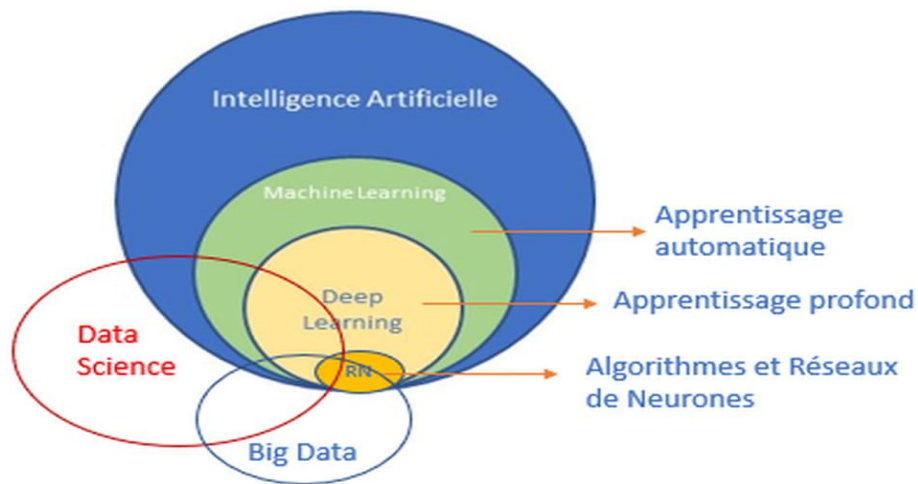


Fig 2. 1: Intelligence artificielle, l'apprentissage automatique et l'apprentissage profond

2.2. L'intelligence artificielle

L'intelligence artificielle est née dans les années 1950 et était une discipline universitaire créée en 1956.

L'intelligence artificielle est «le groupe de théories et de techniques appliquées pour produire des machines capables de simuler l'intelligence humaine», et donc la capacité d'un ordinateur ou un robot contrôlé par un ordinateur pour effectuer des tâches, généralement parce qu'il nécessite de l'intelligence et de l'excellence. Bien que l'intelligence artificielle ne puisse pas effectuer la variété de tâches qu'un humain normal peut effectuer, certains systèmes d'IA peuvent simuler des humains dans des tâches spécifiques. [16] [17] [18]

L'IA a trois niveaux différents :

IA étroite : Quand une machine peut faire une certaine tâche mieux qu'un humain. [18]

B) Intelligence Générale Artificielle : Quand elle peut effectuer n'importe quelle tâche intellectuelle avec le même degré de précision qu'un être humain. [18]

C) Intelligence artificielle active : quand elle peut vaincre les humains dans de nombreuses tâches. [18]

2.3. Un apprentissage automatique

2.3.1. Définition

Un apprentissage automatique (*Machine learning* en anglais ML) est un sous-ensemble de l'intelligence artificielle. Où à Travers elle l'étude des algorithmes informatiques qui s'améliorent automatiquement par l'expérience et entraînent ainsi l'ordinateur à accomplir des tâches qui peuvent être impossibles pour une personne. Lorsque l'appareil a fini d'analyser les données et apprend de ces données, il peut prédire la valeur ou la catégorie de le nouveau point de données et devient ainsi capable de prendre des décisions avec une intervention humaine minimale.

De plus, Arthur Samuel a défini l'apprentissage automatique comme «un domaine d'étude qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés». [18] [19][20][21][22] [23]

2.3.2. Types d'apprentissage automatique

Il est intéressant de comprendre et de définir les différents types d'apprentissage automatique que nous pouvons rencontrer. , Afin que l'apprentissage automatique puisse être catégorisé en fonction du type de données disponibles en apprentissage supervisé, apprentissage non supervisé et apprentissage par renforcement.

Supervisé et non supervisé sont principalement utilisés par de nombreux ingénieurs en apprentissage automatique et les données, l'apprentissage par renforcement est vraiment puissant et compliqué à fournir des problèmes. [18][19][24]

2.3.2.1. Apprentissage Supervisé

C'est le type d'algorithme ML qui suppose initialement des données d'entrée et de sortie. Son rôle est de former le modèle ou l'appareil, car le but de ce problème est d'approximer la fonction de cartographie (et) aussi précisément que possible afin que vous puissiez commencer à prendre une décision ou prédire quand de nouvelles données viendront. L'algorithme est utilisé pour apprendre la fonction de mappage de sorte que lorsqu'il y a de nouvelles données d'entrée (x), la variable de sortie (y) peut être attendue pour l'ensemble de données; Il concerne $y = f(x)$. Les algorithmes de ML supervisés appartiennent aux types suivants [18] [19] [20] [25].

- **Classification** : Chaque apprentissage est associé à une valeur cible qualitative, qui correspond à une classe . IL peut y avoir deux classes (classification binaire) ou plus (classification multiclasse). [18]
- **Regression**: C'est le processus de recherche de corrélations entre les variables dépendantes et indépendantes de sorte que le but du modèle est d'estimer le résultat correct, étant donné le vecteur de caractéristiques. [26]



Fig 2.2: classification et regression

Les données d'entraînement n'incluent pas les cibles ici, nous ne disons donc pas au système . Les données d'entraînement ne sont pas structurées (contiennent des données bruyantes, des données inconnues, etc. Le système lui-même doit comprendre les données que nous fournissons et les extraire automatiquement significatif pour les données. Il existe également différents types d'apprentissage non supervisé tels que détection d'agglomération et d'anomalies (le clustering est très populaire).

- **Regroupement**: C'est similaire à la classification multi-classes mais ici nous ne donnons pas les étiquettes, le système comprend à partir des données elles-mêmes et du regroupement des données et c'est une sorte de problème où nous regroupons les choses similaires. [27]

2.3.2.2. Apprentissage par renforcement (Reinforcement Learning)

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

Ex : L'algorithme de Q-learning est un exemple classique. [28] [29]

2.3.3. Algorithmes utilisés

De nombreux algorithmes d'apprentissage existent aujourd'hui et peuvent être efficaces. On note par exemple :

2.3.3.1. Les machines à vecteur de support SVM

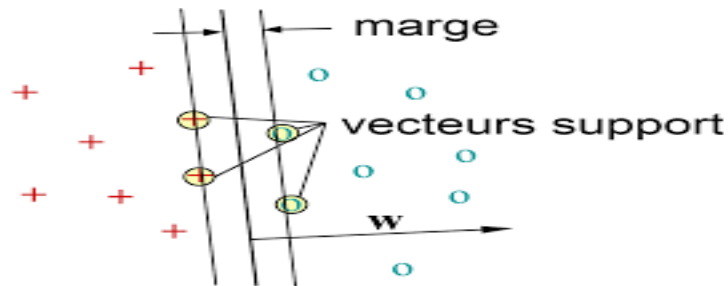


Fig 2.3: Les « vecteurs de support » SVM.

2.3.3.2. La régression logistique

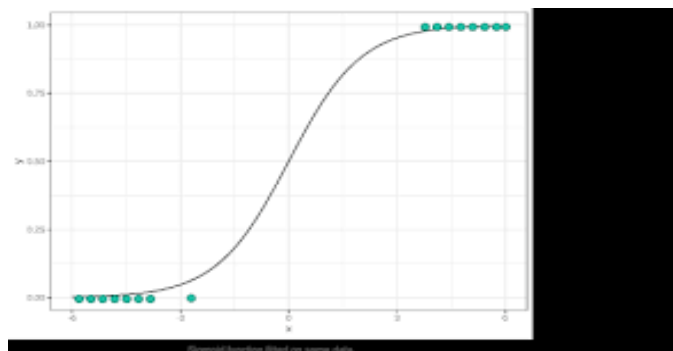


Fig 2.4: La régression logistique.

2.3.3.3. La méthode des k plus proches voisins pour un apprentissage supervisé

- **Exemple:**

Exemple de classification k-NN. L'échantillon de test (cercle vert) pourrait être classé soit dans la première classe de carrés bleus ou la seconde classe de triangles rouges. Si $k=3$ (cercle en ligne pleine) il est affecté à la seconde classe car il y a deux triangles et seulement un carré dans le cercle considéré. Si $k=5$ (cercle en ligne pointillée) il est affecté à la première classe (3 carrés face à deux triangles dans le cercle externe).

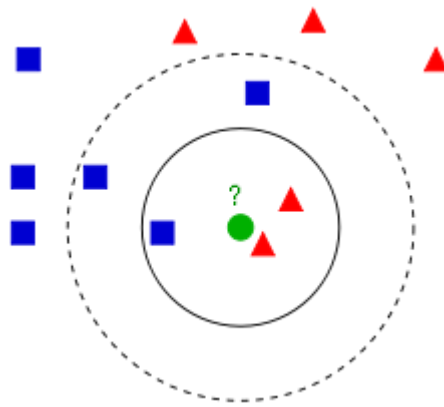


Fig 2.5: Méthode des k plus proches voisins

2.3.3.4. Analyse en composantes principales

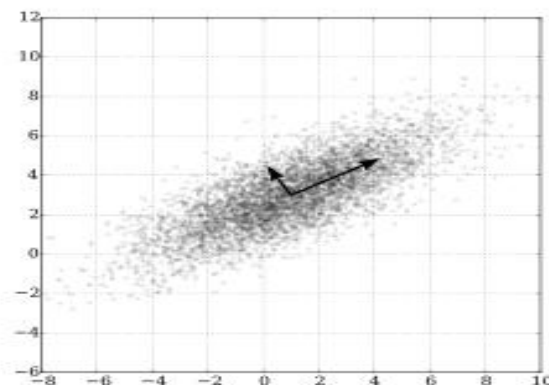


Fig 2.6: Analyse en composantes principales.

2.4. L'apprentissage profond

2.4.1. Définition

L'apprentissage profond est l'incarnation de l'idée de représentations en cascade où l'apprentissage profond s'entraîne sur d'énormes ensembles de données. Parce qu'il y a de nombreux paramètres qu'un algorithme d'apprentissage doit comprendre. Il contient des dizaines, voire des centaines de couches successives de représentations. Il traite de nombreux réseaux de neurones, dont le plus important est le réseau de neurones artificiels, et constitue donc un sous-ensemble et un complément d'algorithmes d'apprentissage automatique. Ces technologies ont permis des avancées significatives et rapides dans les domaines de l'analyse des signaux audio ou visuels, en particulier la reconnaissance vocale, la reconnaissance faciale, le traitement du langage machine et la vision par ordinateur. Dans les années 2000, ces progrès ont stimulé d'importants investissements privés, universitaires et publics, notamment de la part du GAFAM (Google, Apple, Facebook, Amazon et Microsoft). [31] [32] [33] [34]

2.4.2. Un classement en trois catégories du DL

On peut classer en gros la plupart des travaux dans ce domaine en trois classes :

2.4.2.1. Architectures profondes génératives

Qui vise à capturer des images de haute qualité puis la corrélation des données observées ou visuelles pour l'analyse ou la synthèse de modèles.

Et / ou décrire les distributions statistiques courantes des données visuelles

et les catégories associées. Dans ce dernier cas, l'utilisation de la règle de Bayes peut transformer cette commande

Type d'architecture en une seule distinction.

2.4.2.2. Architectures profondes discriminative

Qui visent à fournir directement un pouvoir discriminant à des fins de classification des modèles, ouvent en caractérisant les distributions postérieures des classes conditionnées sur les données visibles et cachées des données.

2.4.2.3. Architectures hybrides profondes

Là où l'objectif est la discrimination assistée (Souvent d'une manière importante) avec les résultats des constructions obstétricales amélioration et / ou meilleure réglementation, ou lorsque des normes discriminatoires

Utilisé pour connaître les paramètres dans l'un des modèles génératifs profonds de la classe 1 dessus.

2.4.3. Modèles d'apprentissage en profond

2.4.3.1. Réseaux de neurones artificiels (ANN) :

Ce sont les formes de traitement d'image et de vision par ordinateur les plus couramment utilisées. C'est une puissante technologie d'apprentissage automatique du domaine de l'apprentissage profond, un réseau de neurones convolutifs prend une photo exprime sous forme de tableau de nombres, il applique une série d'opérations sur cette matrice et, à la fin, renvoie la probabilité qu'un objet de l'image appartienne à une certaine classe ou bien savoir que l'image que vous avez pris contient un bâtiment ou une voiture ou autre objet .un ANN peut être utilisé pour distinguer des états très similaires d'un objet. Un réseau de neurones convolutifs contient une séquence de couches (couche de convolution, couche élu (unités linéaires modifiées), couche d'échantillonnage, couche entièrement connectée) à son architecture de construction, comme le montre la figure [2.7].

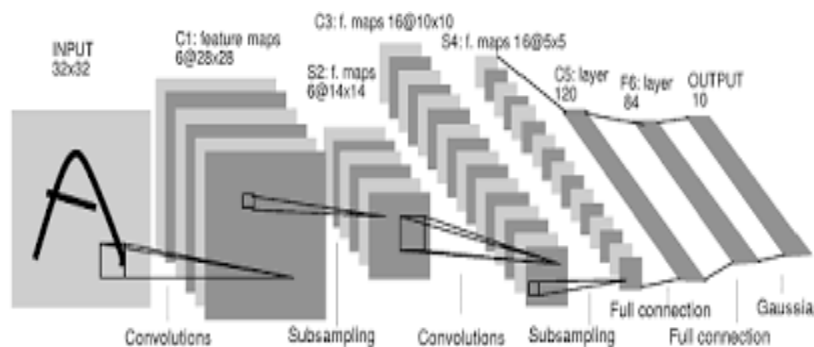


Fig 2.7: Architecture de réseau de neurones convolutifs

2.4.3.2. Auto-Encodeurs empilés

Auto-Encodeurs empilés (SAE) est l'un des types les plus importants du Web profond. La figure ci-dessous montre la structure de SAE, et il se compose de trois

Calques, appelés: couche d'entrée, couche de code et couche de reconstruction. Entrée native X. Il entre dans la couche d'entrée et subit un chiffrement de propagation avant pour le retour Y. Ensuite, l'entrée Y codée est encore décodée pour donner X '. X est une reconstruction de l'entrée originale X et a les mêmes dimensions que X. [20]

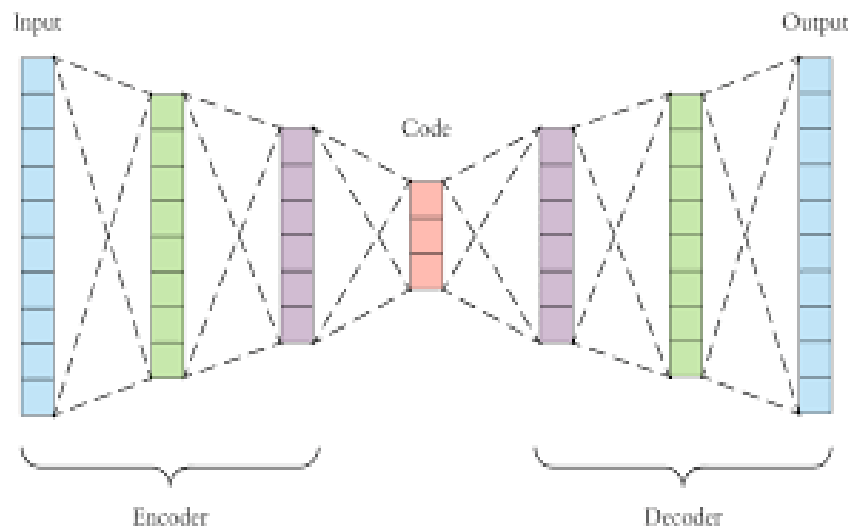


Fig 2.8: Auto-Encoder structure

2.5. Techniques d'apprentissage automatique et d'apprentissage en profondeur

À présent que vous comprenez les concepts de l'apprentissage automatique et de l'apprentissage profond, comparons les deux techniques. Dans l'apprentissage automatique, l'algorithme doit être informé de la façon d'effectuer une prédiction précise en utilisant plus d'informations (par exemple, en effectuant une extraction de caractéristiques). Dans l'apprentissage profond, l'algorithme peut apprendre à effectuer une prédiction précise par le biais de son propre traitement de données, grâce à la structure du réseau neuronal artificiel. Le tableau suivant compare les deux techniques de manière plus détaillée.

Table 2.1: Un apprentissage automatique Vs L'apprentissage profond.

Éléments de comparaison	L'apprentissage automatique	L'apprentissage profond
Nombre de points de données	Peut utiliser de petites quantités de données pour faire des prédictions.	A besoin de grandes quantités de données d'entraînement pour effectuer des prédictions.
Dépendances matérielles	Peut fonctionner sur des machines bas de gamme. Ne nécessite pas beaucoup de	Nécessite des machines haut de gamme. Effectue fondamentalement un grand nombre d'opérations de multiplication de matrices. Un GPU peut optimiser efficacement ces

	puissance de calcul.	opérations.
Processus de personnalisation	Nécessite que les caractéristiques soient identifiées et créées avec précision par les utilisateurs.	Apprend des caractéristiques de haut niveau à partir de données, et crée de nouvelles caractéristiques de façon autonome.
Approche d'apprentissage	Scinde le processus d'apprentissage en étapes plus petites. Combine ensuite les résultats de chaque étape dans une seule sortie.	Parcourt le processus d'apprentissage en résolvant le problème de bout en bout.
Temps d'exécution	Nécessite relativement peu de temps pour apprendre, de quelques secondes à quelques heures.	Nécessite généralement un temps d'entraînement assez long, car un algorithme d'apprentissage profond implique de nombreuses couches.
Sortie	La sortie est généralement une valeur numérique telle qu'une note ou une classification.	La sortie peut avoir plusieurs formats, comme un texte, un score ou un son.

2.6. Conclusion

Dans ce chapitre, nous avons appris la relation et la différence entre l'apprentissage en profondeur et l'apprentissage automatique et l'intelligence artificielle. Nous avons également mentionné les types d'apprentissage automatique et les algorithmes qui sont utilisés, puis nous avons abordé les classes uniques dans les modèles de DL et d'apprentissage en profondeur. Enfin, nous avons établi une petite comparaison entre L'apprentissage profond et l'apprentissage automatique.



Chapitre3:
Environnement de
Programmation

3.1. Introduction

Après avoir expliqué et étudié dans la deuxième partie l'IA, ML et DL, nous allons présenter dans cette partie l'importance et la méthode de téléchargement et d'installation de la plate-forme de travail sur laquelle nous allons travailler afin de s'adresser à plus fortes plate-forme Anaconda qui travaille avec le langage python.

Ensuite, nous présentons les différentes bibliothèques les plus importants que nous devons installer, par exemple Keras et TensorFlow. Afin de démarrer le projet en utilisant DL.

3.2. Présentation des outils et des bibliothèques

3.2.1. Le hardware

- ❖ Un PC portable Dell i5-5300U CPU @ 2.30GHz 2.29 GHz
- ❖ RAM de taille 8 GO.
- ❖ Disque dur de taille 118 GO.
- ❖ Carte graphique Intel(R) HD Graphics 5500.
- ❖ Système d'exploitation 64 bits Windows 10.



Fig 3. 2: Le hardware.

3.2.2. Les outils et les bibliothèques utilisés

3.2.2.1. Python

- **Définition :**

Python est un langage de programmation interprété, multi-paradigm et multiplateformes. il favorise la programmation impérative structurée, fonctionnelle et orientée objet. le langage de programmation python a été conçu dans les années 1980 en retard, et sa mise en œuvre a débuté en décembre 1989, Guido van Rossum (CWI). [37][38] [39] [40]



Fig 3. 2: Le Python.

- Langage Python :

Table 3.1: Les versions Python.

<i>Les versions</i>	<i>La dernière micro</i>	<i>La date de sortie</i>
. Python 0.9	0, 9,9. [39]	20/02/1991. [39]
. Python 1,0	1, 0,4. [39]	26/01/1994. [39]
. Python 1.1:	1, 1,1. [39]	11/10/1994. [39]
. Python 1.2:		13/04/1995. [39]
. Python 1.3:		13/10/1995. [39]
. Python 1.4:		25/10/1996. [39]
. Python 1.5:	1, 5,2. [41]	03/01/1998. [39]
. Python 1.6:	1.6.1. [41]	05/09/2000. [39]
. Python 2.0:	2.0.1.[42]	16/10/2000. [41]
. Python 2.1:	2.1.3.[42]	15/04/2001. [42]
. Python 2.2:	2.2.3 [42]	21/12/2001. [43]
. Python 2,3:	2, 3,7. [42]	29/06/2003. [44]
. Python 2,4:	2, 4,6. [42]	30/11/2004. [45]
. Python 2,5:	2, 5,6. [42]	19/09/2006. [46]
. Python 2,7	2, 7,18. [47]	03/07/2010. [47]
. Python 3,0:	3, 0,1. [42]	03/12/2008. [42]
. Python 3,1:	3, 1,5. [50]	27/06/2009. [50]
. Python 3,2:	3, 2,6. [51]	20/02/2011. [51]
. Python 3,3:	3, 3,7. [52]	29/09/2012. [52]
. Python 3,4:	3, 4,10. [53]	16/03/2014. [53]
. Python 3,5:	3, 5,10. [54]	13/09/2015. [54]
. Python 3,6:	3, 6,13. [55]	23/12/2016. [56]
. Python 3,7:	3, 7,10. [56]	27/06/2018. [57]
. Python 3,8:	3, 8,8. [58]	14/10/2019. [58]

On distingue deux versions de python: python 2 et python 3. Les différences entre ces deux versions sont multiples. Python 2.x est l'ancienne version, qui continuera d'être supportée et

donc de recevoir des mises à jour officielles jusqu'en 2020. Après cette date, elle continuera d'ailleurs sans doute de subsister de façon non officielle.

Python 3.x est la version actuelle du langage. Elle apporte de nombreuses fonctionnalités nouvelles et très utiles, telles qu'un meilleur contrôle de concurrence et un interpréteur plus efficace. Cependant, l'adoption de Python 3 a été long temps ralentie par le manque de bibliothèque tierces prises en charge. Un grand nombre d'entre elles n'étaient compatibles qu'avec Python 2, ce qui rendait la transition compliquée. Cependant, ce problème est aujourd'hui pratiquement résolu et il reste peu de raisons valables de continuer à utiliser *Python* 2.

Mais pour moi, j'ai utilisé la version 3.6.

- **Les principaux avantages:**

Le langage Python doit sa popularité à plusieurs avantages qui profitent aussi bien aux débutants qu'aux experts. Tout d'abord, il est facile à apprendre et à utiliser. Ses caractéristiques sont peu nombreuses, ce qui permet de créer des programmes directs.

Un autre avantage du Python est sa popularité. Ce langage fonctionne sur tous les principaux systèmes d'exploitation et plateformes informatiques. De plus, même s'il ne s'agit clairement pas du langage le plus rapide, il compense sa lenteur par sa versatilité.

Enfin, même s'il est principalement utilisé pour le Scripting et l'automatisation, Ce langage est aussi utilisé pour créer des logiciels de qualité professionnelle. Qu'il s'agisse d'applications ou de services Web, le Python est utilisé par un grand nombre de développeurs pour créer des logiciels.

3.2.2.1. Top des meilleures bibliothèques et packages

Si le Python est considéré comme le meilleur langage de programmation pour le Big Data, c'est grâce à ses différents packages et bibliothèque de science des données. Voici les plus popularisés.

Pandas :

Le bibliothèque Panda est développé par des scientifiques de données familiers avec R et python est l'une des bibliothèques de science des données les plus populaires. Il fournit différentes structures et outils pour l'analyse des données ainsi que plusieurs techniques intégrées pour toucher, fusionner et filtrer les données, dont le cas le panda est utile car il a été développé spécifiquement pour l'extraction et la préparation. [59] [60]

- **NumPy:**

NumPy est une bibliothèque Python très populaire pour la gestion de grands tableaux et tableaux multidimensionnels, s'intègre facilement à de nombreuses bases de données différentes et est très utile pour les calculs scientifiques de base en apprentissage automatique. IL est particulièrement utile et idéal pour les opérations liées à l'algèbre linéaire, aux transformations instantanées et aux capacités de nombres aléatoires.

Les bibliothèques sophistiquées comme TensorFlow utilisent NumPy en interne pour gérer les tenseurs. [58]

- **Scipy**

Scipy est une bibliothèque de calculs techniques et scientifiques très populaire parmi les passionnés d'apprentissage automatique car il contient de nombreux modules d'optimisation, d'algèbre linéaire, d'intégration et de statistiques. Il y a une différence entre Scipy Library et Scipy Stack. Scipy est l'un des packages de base qui composent la pile Scipy. Scipy est également utile pour le traitement d'images. [61]

- **TensorFlow**

TensorFlow est une API d'apprentissage automatique et une bibliothèque open source fonctionnant en Python conçue par l'équipe Brain de Google en 2015.

Avec cette bibliothèque, les développeurs Python peuvent désormais utiliser facilement l'intelligence artificielle et l'apprentissage automatique, car elle vise à faciliter son utilisation. TensorFlow peut être utilisé à partir de nombreux langages logiciels tels que : Python, C++, Java, Scala et R ce qui rend c'est très rapide. Fournit des alternatives pour définir automatiquement des fonctions sur les tenseurs et calculer leurs dérivées

Les temps sont des constructions mathématiques de base dans des domaines tels que la physique et l'ingénierie. Cependant, historiquement. Ces tenseurs avaient moins progressé en informatique, qui était plus liée aux mathématiques discrètes et à la logique. Cette situation a commencé à changer radicalement avec l'avènement de la machine. [62] [63][64]

- **Keras**

Keras est l'API officielle de premier niveau pour TensorFlow Il s'agit d'une API et d'une bibliothèque de réseau neuronal open source écrites en Python et l'API la plus largement utilisée en

apprentissage en profondeur, car Keras et l'ensemble de sa bibliothèque sont plus évolutifs et facilitent l'exécution de nouvelles expériences. [65]

- **Tkinter**

Tkinter, ou "Tk Interface", est un module Python qui fournit une interface GUI (Graphical User Interface). Il est célèbre pour sa simplicité et est open source. Le module tkinter fait partie du standard Python, fourni avec l'interpréteur, et la documentation Python documente le tkinter comme s'il s'agissait d'un module intégré. [66]

3.2.2.3. Anaconda

- **Définition:**

Anaconda est une distribution libre et open source. des langages de programmation python et R appliqué au développement d'applications dédiées à la science des données et à l'apprentissage automatique (traitement de données à grande échelle, analyse prédictive, calcul scientifique), qui vise à simplifier la gestion des paquets et de déploiement. Les versions de paquets sont gérées par le système de gestion de paquets Conda. La distribution Anaconda est utilisée par plus de 6 millions d'utilisateurs et comprend plus de 250 paquets populaires en science des données adaptés pour Windows, Linux et MacOS. [67] [68] [69] [70]



Fig 3. 3: Le Anaconda.

- **Installation de Anaconda (Python 3.6) :**

⇒ Anaconda est facile à installer, il suffit de télécharger Anaconda 64 bits, d'exécuter le programme d'installation et de suivre les instructions.

1. Recherchez "Anaconda python" sur Google → Python V 3.6 pour Windows 64 bits.
2. Installez le fichier d'installation Anaconda 64 bits pour Windows.
3. Installez pour → option: Juste pour moi (option: Just for me).

⇒ Modifiez les packages anaconda en tapant les deux commandes:

> Conda update Conda

> Conda update -all

- **Installer des packages python Conda:**

⇒ C:\>conda install jupyter → Jupiter 1.0

⇒ C:\>conda install scipy → The Scipy table library

⇒ C:\>pip install sklearn → The machine learning library

⇒ C:\>pip install pandas → The machine learning library

⇒ C:\>pip install pandas-datareader → ML pandas library reading tools

⇒ C:\>pip install matplotlib → Install the curve plotting tool

⇒ C:\>pip install pillow → ML complementary tools

⇒ C:\>pip install requests → ML complementary tools

⇒ C:\>pip install h5py → ML complementary tools

⇒ C:\>pip install TensorFlow/TensorFlow==1.5.1 → Install the machine learning library and the Google TensorFlow interface.

⇒ C:\>pip install Keras/Keras==2.0.0 → Install the machine learning library and the Keras interface.

3.3. IDLE

IDLE est un programme open source qui vous permet d'écrire et d'exécuter vos propres programmes. Il s'agit d'un simple environnement de développement intégré (IDE) livré avec Python. IDLE n'est qu'un simple éditeur.

Mais pour l'instant, je suggère de s'en tenir à IDLE car il est facile à utiliser. Vous pouvez trouver IDLE dans le dossier anaconda 3 sur votre ordinateur. [64] [71]

3.4. Conclusion

Dans ce chapitre, nous avons présenté tous les outils nécessaires pour développer notre application en expliquant les espaces de travail et les bibliothèques comme Pandas, NumPy, Scipy, TensorFlow et Keras. Nous avons essayé d'expliquer comment installer tous ces outils pour commencer notre application, et comment les exploiter dans la phrase de programmation proprement.

**Chapitre 04 :
Application
réalisée et
Résultats
obtenus**

4.1. Introduction

Ce chapitre présente la partie expérimentale, il contient donc les éléments suivants :

- Le prétraitement du dataset utilisé.
- Le modèle proposé
- L'interface de notre application.

4.2. Problématique de notre travail

Nous nous concentrerons sur le problème de l'analyse des sentiments dans le domaine du tourisme. Cela exige l'analyse des commentaires des clients stockés dans un dataset et d'en tirer les différentes valeurs (positive, négative, neutre) cachées dans ces commentaires.

4.3. Prétraitement du dataset utilisé

- Chargement du dataset : par la fonction "pd.read_csv".

```
#read the datasets
```

```
dataset = pd.read_csv ("train.csv" )
```

```
testset = pd.read_csv ("./test.csv" )
```

- Créer un dictionnaire de mots et leurs indices

```
# Create Dictionary of words and their indices  
max_words = 10000  
tokenizer = Tokenizer(num_words=max_words)  
tokenizer.fit_on_texts(reviews)  
dictionary = tokenizer.word_index  
|
```

- Remplacer les mots de chaque revue de texte par des indices

```
# Replace words of each text review to indices  
allWordIndices = []  
for num,text in enumerate(reviews):  
    wordIndices = convert_text_to_index_array(text)  
    allWordIndices.append(wordIndices)
```

- Convertir les séquences d'index en sac "TF-IDF" de vecteur de mots

```
# Convert the index sequences into tfidf bag of words vector (one hot encoding)
allWordIndices = np.asarray(allWordIndices)
train_X=tokenizer.sequences_to_matrix(allWordIndices,mode='tfidf')
labels = keras.utils.to_categorical(labels,num_classes=2)
```

4.4. Le modèle proposé

Nous avons d'abord essayé de développer un classificateur binaire automatisé qui quantifiait le bonheur des touristes en fonction de certaines caractéristiques. C'est donc une classification binaire de deux (02) classes données (le bonheur a l'apparence de 1 ou pas de 0), puis on a essayé de développer par Tf_Idf, l'algorithme pour calculer la fréquence du mot-fréquence du document inverse comme suit :

$$Tf - IDF = tf_{ij} \times \log\left(\frac{N}{df_i}\right) \quad (4.1)$$

Toutes les variables d'entrée décrivant chaque visiteur sont des commentaires. Nous avons également noté que pour construire le meilleur modèle prédictif, il existe deux méthodes, l'approche ML classique et l'approche DL :

Tout d'abord, avant nous, ils ont utilisé l'approche ML classique et sont passés à la classification à l'aide de plusieurs algorithmes de ML, notamment LR, LDA, KNN, CART, NB et SVM. Quant à l'approche DL, nous proposons un modèle ANN composé de plusieurs couches simples : une seule couche d'entrées (256 neurones), cinq couches cachées (128, 64, 16, 32 neurones), et enfin une couche résultante (02 neurones). Comme nos outils de programmation nous avons utilisé Python, Tensorflow et Keras qui sont les plus utilisés dans ce domaine. La figure 4.5 présente un schéma détaillé du modèle ANN proposé pour son amélioration. Pour effectuer la tâche de classification.

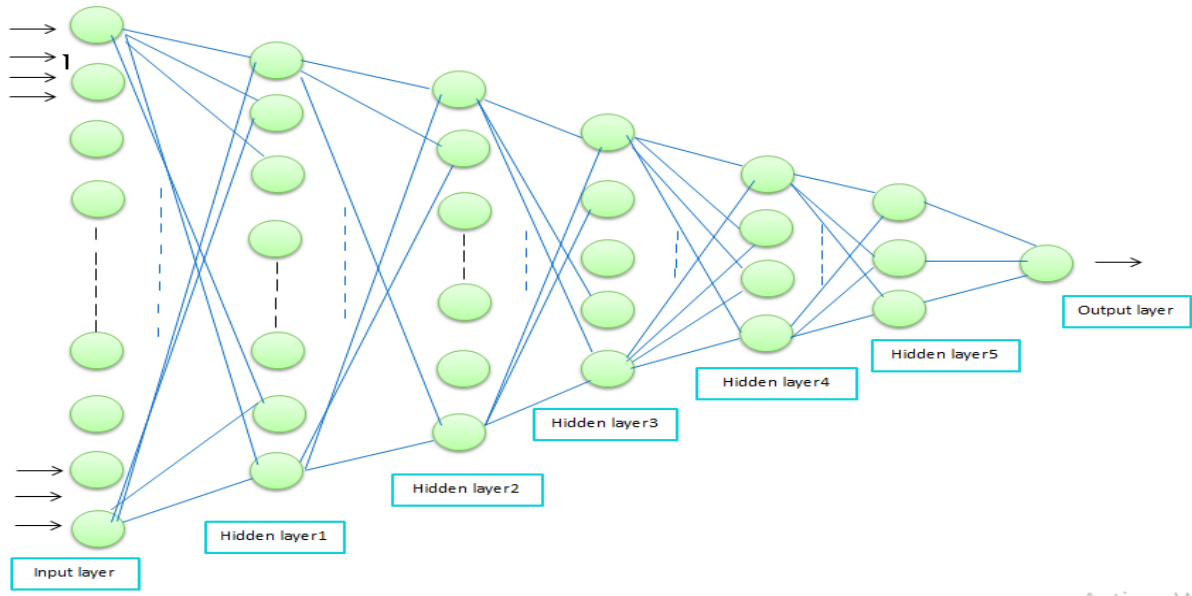


Fig 4.1 : Architecture du modèle ANN proposé.

4.5. Evaluation

Pour vérifier notre modèle proposé, nous avons utilisé plusieurs mesures de performance, à savoir :

- **La précision (Accuracy) :**

La précision est une échelle de notation utilisée pour évaluer les modèles de notation, elle montre donc à quel point le modèle est précis, qui est une proportion du nombre d'instances correctement prédites divisé par le nombre total d'instances dans l'ensemble de données multiplié par 100 pour donner un pourcentage. Par exemple, la précision est de 0.97, soit 97 % (97 prédictions correctes sur un total de 100 exemples). [72][75] [60]

- **La Perte(loss) :**

La fonction de perte est utilisée pour optimiser les valeurs des paramètres dans le modèle de réseau neuronal. La fonction de perte fait correspondre un ensemble de valeurs de paramètres du réseau à des valeurs scalaires, qui indiquent le degré auquel ces paramètres accomplissent les tâches que le réseau entend accomplir. Elle calcule l'erreur pour un seul exemple de formation. [73] [60]

- **Valeur de perte (lossvalue) :**

Donne la somme des erreurs commises pour chaque exemple dans l'ensemble d'apprentissage ou de validation. Elle doit être calculée sur les ensembles de données de formation et de validation.

- **Optimiseur(Optimizer):**

Un optimiseur est un algorithme ou une méthode utilisée pour modifier les propriétés du réseau neuronal (telles que les poids et le taux d'apprentissage). Les algorithmes ou stratégies d'optimisation sont chargés de réduire les pertes et de fournir les résultats les plus précis. Il existe de nombreux types d'optimiseurs[74] [60].

- **Batch_Size :**

Spécifie le nombre d'échantillons que seront propagés sur le réseau.

- **Normalisation par lots (Batch_Normalization) :**

Peut être utilisé dans un modèle. De plus, cette technique peut grandement améliorer la convergence pendant l'apprentissage et permettre à chaque couche du réseau d'apprendre plus d'informations indépendamment des autres couches.

- **Epoche :**

C'est le nombre maximum d'itérations d'entraînement.

- **Dropout :**

Il s'agit d'une technique de régularisation (pour lutter contre le surapprentissage).

- ❖ **model = models.Sequential ()**

Cette commande crée une pile pour stocker les différentes couches.

- ❖ **model.add (Batch_Normalization())**

Cette commande permet d'ajouter une couche batch_normalization.

- ❖ **model.add (Activation ('softmax'))**

Cette commande ajoute une couche d'activation pour notre couche (bidirectionnelle Lstm).

- ❖ **model.add (Dense (256, activation='elu'))**

Cette couche permet d'ajouter une couche de sortie composée de 256 neurones (nombre de notre classe), la fonction « elu » est utilisée pour calculer la probabilité de chaque classe.

❖ **model.compile (loss='categorical_crossentropy', optimizer='sgd', metrics= ['accuracy'])**

Cette commande permet de compiler notre modèle, elle prend trois paramètres : la fonction Loss et Optimizer, Metrics. On a choisi la fonction categorical_crossentropy comme fonction loss et adam comme optimizer .et accuracy comme metrics.

❖ **model.fit (train_X, labels,epochs=EPOCHS, batch_size=BATCH_SIZE, verbose=1, validation split=0.1, shuffle=True)**

Cette commande permet de lancer l'apprentissage en a comme paramètre notre description de train avec leur polarités, le nombre d'épochè ainsi les descriptions pour la validation et un point de restauration de meilleur résultat dans l'entraînement.

❖ **model.evaluate (input, pred)**

Cette commande permet d'évaluer notre modèle sur la base de test.

Tableau 4.1 : Description du modèle ANN proposé

Layer (type)	Output Shape	Param #
dense_7 (Dense)	(None, 256)	2560256
batch_normalization_6	(Batch (None, 256)	1024
dropout_6 (Dropout)	(None, 256)	0
dense_8 (Dense)	(None, 128)	32896
batch_normalization_7	(Batch (None, 128)	512
dropout_7 (Dropout)	(None, 128)	0
dense_9 (Dense)	(None, 64)	8256
batch_normalization_8	(Batch (None, 64)	256
dropout_8 (Dropout)	(None, 64)	0
dense_10 (Dense)	(None, 32)	2080
batch_normalization_9	(Batch (None, 32)	128
dropout_9 (Dropout)	(None, 32)	0
dense_11 (Dense)	(None, 16)	528
batch_normalization_10	(Batch (None, 16)	64
dropout_10 (Dropout)	(None, 16)	0
dense_12 (Dense)	(None, 2)	34
Total params: 2, 606,034.0		
Trainableparams: 2, 605,042.0		
Non-Trainableparams: 992.0		

Tableau 4.2: Perte et précision obtenues.

	Perte	Précision
Train	0.06	0.986
Test	0.1082	1.0

Fig4.2 montre l'évaluation de train perte et de perte de validation dans le temps en fonction du nombre, d'époques. **Fig4.3** montre l'évolution de précision du train et de précision de la validation en fonction du nombre d'époques. **Fig4.4** tracer des courbes et imaginez la perte et la précision

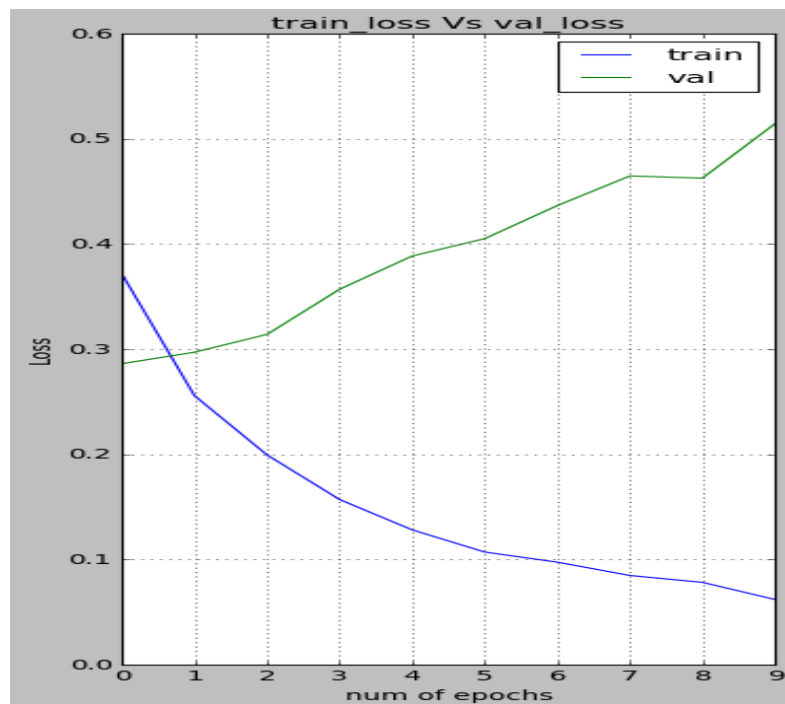


Fig4.2: Training loss Vs Validation loss of the ANN model

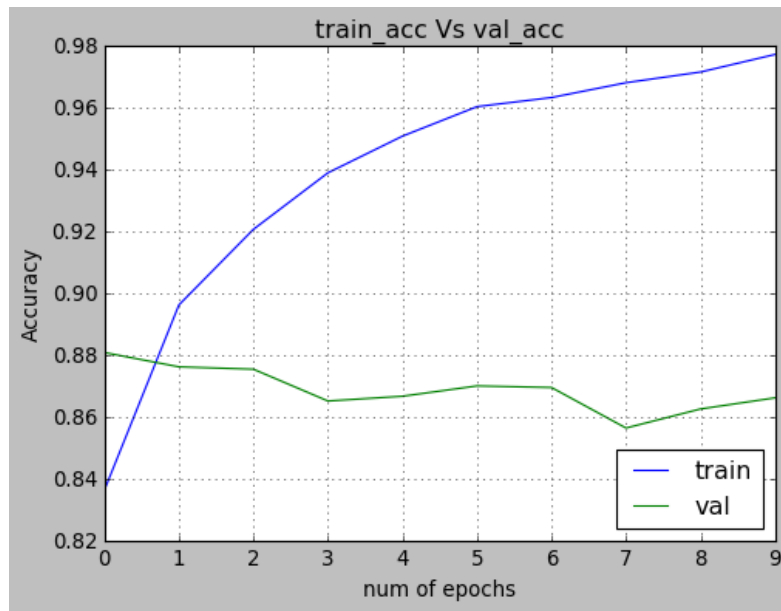


Fig4.3: Training accuracy Vs Validation accuracy of the ANN model

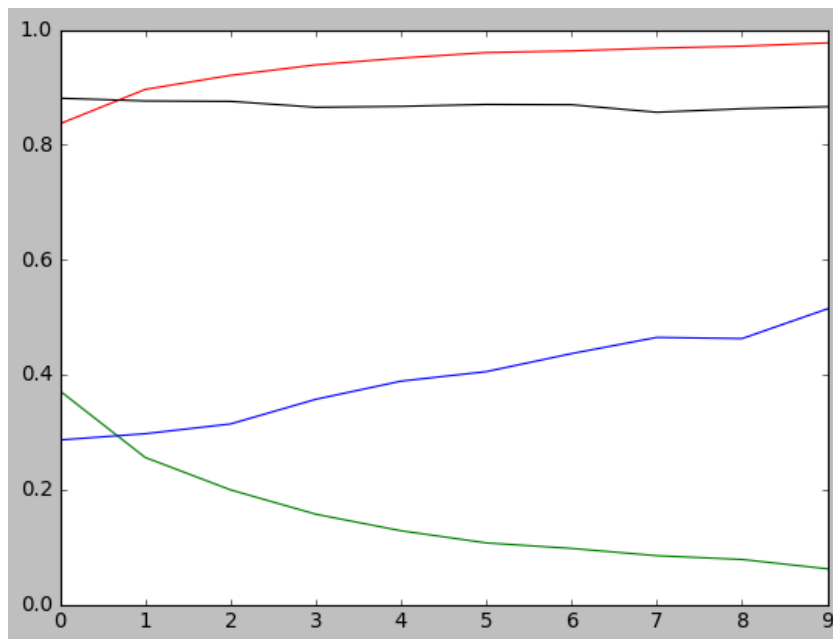


Fig4.4: Accuracy VS Val Accuracy & Loss VS Val loss

4.6. Interface de l'application

Notre interface se compose de trois pages : la page de démarrage, la page d'accueil et la page à propos. La page de démarrage contient le nom de l'application, ainsi que deux boutons, dont l'un nous amène à la page d'accueil et le second à la page à propos. La page d'accueil à une catégorisation des pages de commentaires des clients, la dernière page "À propos de la page" donne également une brève définition des commentaires des clients, il y en a trois pour chaque page et un bouton de sortie pour quitter la fenêtre.

4.7. Page de démarrage

Sur la page de démarrage, il y a deux boutons A propos et Entrée.



Fig4.5:page de démarrage

4.8. Page de d'accueil

Sur la page principale, il y a trois boutons : télécharger, évaluer et savoir Le bouton "Télécharger" lorsque vous cliquez dessus vous donne accès au dataset dans votre ordinateur, lorsque vous choisissez le fichier, il apparaîtra sur l'appareil sous forme de fichier csv

Après avoir sélectionné le dataset classé avec le bouton Charger, vous devez cliquer sur le bouton Entraîner. Le bouton "Train" imprimera un pourcentage de taux d'apprentissage. Ensuite, nous pouvons appuyer sur le bouton de test pour imprimer le taux de reconnaissance des sentiments dans les commentaires des clients

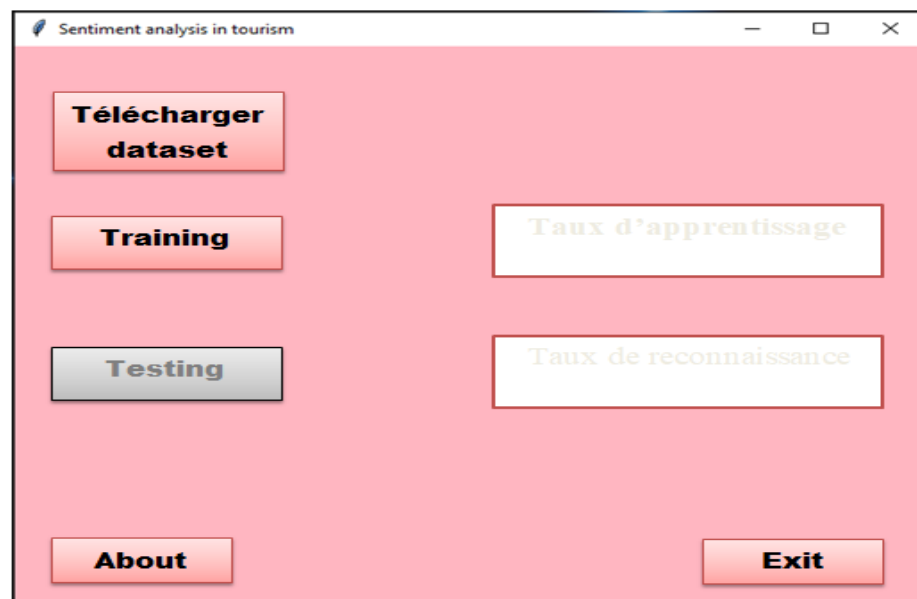


Fig4.6:page d'accueil

. À propos de la page :

La page A propos est la dernière page de notre interface et n'est qu'une étiquette qui donne une brève définition de notre interface.

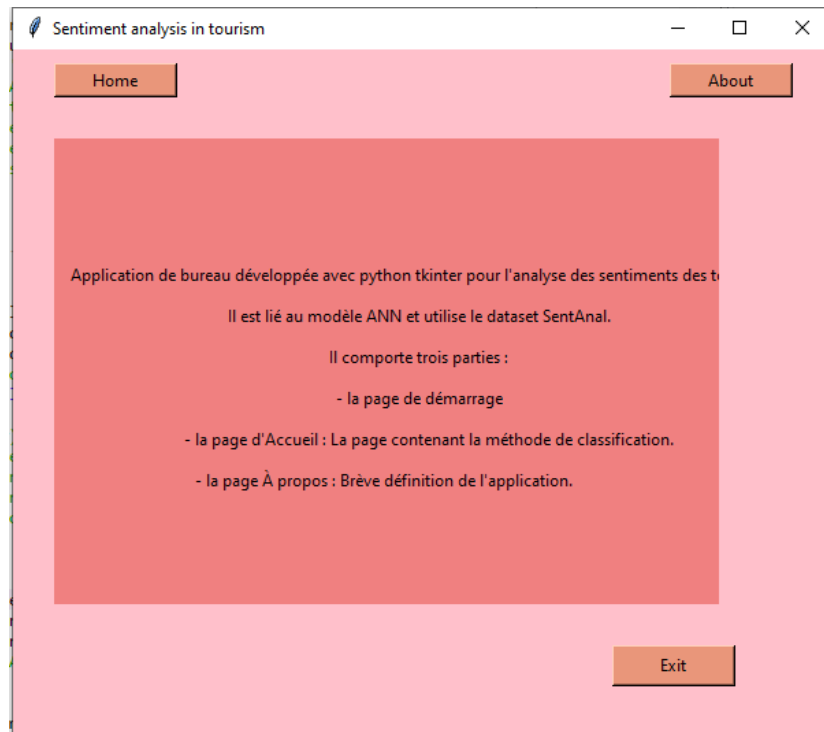


Fig4.7:page À propos

4.10. Conclusion

Notre travail peut être divisé en deux parties principales : le modèle ANN construit et notre interface d'application, donc dans ce chapitre nous avons présenté le travail expérimental, est incarné dans ce qui suit : La description du modèle proposé, l'estimation du modèle ANN proposé, l'interprétation des résultats obtenus, nous avons également défini Quelques concepts, dans le dernier nous avons présenté notre interface d'application.



CONCLUSION GENERALE



Conclusion générale

Au cours des dernières années, des progrès importants ont été réalisés dans le domaine du ML, en particulier avec l'émergence d'un nouveau sous-domaine appelé DL Deep Learning. Elle repose principalement sur l'utilisation de plusieurs réseaux de neurones. Par conséquent, la reconnaissance d'objets est principalement basée sur l'approche ML qui offre des performances élevées.

Par exemple, l'analyse des sentiments et les retours clients envers les services des agences de tourisme deviennent un domaine important en raison de ses applications dans de nombreux domaines (hôtels, agences de tourisme, etc.).

L'analyse des sentiments est un domaine de recherche actif qui doit toujours être amélioré en termes de précision. Dans le présent travail, nous avons mis en œuvre une tâche de classification sur un ensemble de données pour obtenir les commentaires des clients.

Le traitement des données vise à supprimer des signes et des phrases des textes, à supprimer des signes, des répétitions, des mots et des prépositions exclus et à remplacer des lettres et des expressions sans importance dans le texte.

Le formulaire exige que les données soient clairement organisées et parce que la longueur des textes et des mots et le nombre de mots qu'ils contiennent peuvent varier et ne rendent pas le texte clair ou ordonné, nous utilisons des techniques d'extraction de caractéristiques pour chaque texte afin de saisir clairement le formulaire. La technique d'extraction de caractéristiques la plus importante est peut-être la technologie de vectorisation de comptage : dans ce processus, nous convertissons les mots en nombres et donnons à chaque mot un nombre, c'est-à-dire créons un dictionnaire qui représente chaque mot, son nombre et le nombre de répétitions, et les données sont donc valides pour entrer dans le modèle d'apprentissage en profondeur.

Nous construisons un modèle ANN pour effectuer la même tâche de classification. Les résultats obtenus étaient plus précis et très surprenants lors de l'application de l'approche DL à travers le modèle ANN que nous avons construit.

En tant que résultats empiriques, la méthode donne la meilleure précision avec une grande dataset avec un taux d'erreur de classification de 6% dans l'apprentissage et de 10% dans les tests d'erreur de classification.

Nous proposons un modèle ANN composé de plusieurs couches simples : une seule couche entrées (256 neurones), cinq couches cachées (128, 64, 16, 32 neurones), et enfin couche résultante (02 neurones). Comme nos outils de programmation nous avons utilisé Python, Tensorflow et Keras qui sont les plus utilisés dans ce domaine. La figure 4.5 présente un schéma détaillé du modèle ANN proposé pour son amélioration Pour effectuer la tâche de classification.

Nous avons construit un modèle ANN pour effectuer la tâche de classification. Les résultats obtenus étaient plus précis et très surprenants lors de l'application de l'approche DL via le modèle ANN que nous avons construit. En tant que résultats expérimentaux, l'approche donne la meilleure précision avec une grande base de données avec un taux d'erreur de classification d'entraînement de 0,6 % et un taux d'erreur de classification d'échec de test de 1 %.

BIBIOGRAPHIE

- [1] Ravi, K. and Ravi, V. (2015). A survey on opinion mining and sentiment analysis : tasks, approaches and applications. Knowledge-Based Systems journal, Elsevier. 89, pp. 14-16
- [2] Tanaya, H. (2017). Survey on Sentiment Analysis : A Comparative Study. 159 (6), pp. 4–7.
- [3] Assiri, A., Emam, A. and Aldossari, H. (2015). Arabic Sentiment Analysis : A Survey. (IJACSA) International Journal of Advanced Computer Science and Applications, 6 (12), pp. 75–85.
- [4] Alaei, A.; Becken, S.; Stantic, B. Sentiment Analysis in Tourism: Capitalizing on Big Data. J. Travel Res. 2019, 58, 175–191. [Google Scholar] [CrossRef]
- [5] (en) Andrius Mudinas, Dell Zhang et Mark Levene, « Combining lexicon and learning based approaches for concept-level sentiment analysis », [WISDOM '12 Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining Article No. 5, 2012 (ISBN 978-1-4503-1543-2, lire en ligne [archive]).
- [6] (en) Prabu Palanisamy, Vineet Yadav et Harsha Elchuri, « Serendio: Simple and Practical lexicon based approach to Sentiment Analysis », Second Joint Conference on Lexical and Computational Semantics (*SEM), vol. 2`titre volume=Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), juin 2013 (lire en ligne [archive]).
- [7] (en) Bo Pang, Lillian Lee et Shivakumar Vaithyanathan, « Thumbs up?: sentiment classification using machine learning techniques », EMNLP '02 Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10, 2002 (lire en ligne [archive]).
- [8] (en) Xue Zhang, Hauke Fuehres et Peter A. Gloor, « Predicting Stock Market Indicators Through Twitter “I hope it is not as bad as I fear” », The 2nd Collaborative Innovation Networks Conference - COINs2010, 2011 (lire en ligne [archive]).
- [9] (en) Spyros E. Polykalas, George N. Prezerakos et Agisilaos Konidaris, « A General Purpose Model for Future Prediction Based on Web Search Data: Predicting Greek and Spanish Election », 27th International Conference on Advanced Information Networking and Applications Workshops, 2013 (lire en ligne [archive])
- [10] (en) Andranik Tumasjan,, Timm O. Sprenger, Philipp G. Sandner et Isabell M. Welpe, « Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment », Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 2010 (lire en ligne [archive])
- [11] (en) Panagiotis T. Metaxas, Eni Mustafaraj et Daniel Gayo-Avello, « How (Not) to Predict Elections », IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, 2011 (lire en ligne [archive]).
- [12] (en) Asur, S, « Predicting the Future with Social Media », Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 (lire en ligne [archive]).
- [13] (en) Alexander Pak et Patrick Paroubek, « Twitter as a Corpus for Sentiment Analysis and Opinion Mining », Language Resources and Evaluation Conference, 2010 (lire en ligne [archive]).
- [14] (en) Adam Funk, Yaoyong Li, Horacio Saggion, Kalina Bontcheva et Christian Leibold, « Opinion Analysis for Business Intelligence Applications », OBI '08 Proceedings of the

Bibliographie

- first international workshop on Ontology-supported business intelligence Article No. 3, 2008 (ISBN 978-1-60558-219-1, lire en ligne [archive]).
- [15] (en) Mari Carmen Rodríguez-Gancedo, Javier Caminero, José Relañó et Carlos Picazo, « T-OMT: A Novel Opinion Mining Tool for Improving Global Customer Relationship Management », CASFE'13, 2013 (lire en ligne [archive]).
- [16] Copeland, B.J."Artificial intelligence".Encyclopedia Britannica, 11 Aug. 2020, <https://www.britannica.com/technology/artificial-intelligence>. Accessed 4 May 2021.
- [16] « [Encyclopédie Larousse en ligne - intelligence artificielle](#) » [archive], sur Larousse.
- [18] Ali, M. A., Svetinovic, D., Aung, Z., & Lukman, S. (2017). Malware detection in android mobile platform using machine learning algorithms. International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS) Khalifa University of Science and Technology Abu Dhabi, United Arab Emirates.
- [19] Sazzadur Rahman, A.K.M., Javed Mehedi Shamrat, F. M., Tasnim, Z., Roy, J., Hossain, S. (2019). A Comparative Study on Liver Disease Prediction Using Supervised Machine Learning Algorithms. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 11.
- [20] Zain Amin, M., Ali, A.Performance Evaluation of Supervised Machine Learning Classifiers for Predicting Healthcare Operational Decisions.C-Section Classification Database Report, UCI Machine Learning Repository, University of California, Irvine, USA.Department of Computer Science & Engineering University of Engineering & Technology, Lahore, Pakistan.
- [21] Ravi Kumar, G.,Dr.Ramachandra, G. A. ,Nagamani, K.(2013).An Efficient Prediction of Breast Cancer Data using Data Mining Techniques.International Journal of Innovations in Engineering and Technology (IJET).Department of Computer Science & Technology S K University, Anantapur, AP, India, Rayalaseema University, Kurnool, AP, India.Vol. 2 Issue 4.
- [22] Ekinci,E. , İlhan Omurca,S. , Acun,N.(2018).A Comparative Study on Machine Learning Techniques using Titanic Dataset .7th International Conference on Advanced Technologies (ICAT'18), Antalya/TURKEY.Kocaeli University, Faculty of Engineering, Computer Engineering Department Umuttepe Campus, Kocaeli, Turkey.
- [23] Said G. Developping an efficient Predictive model based on ML and DL approaches to detect diabetes, an International Journal of Computing and Informatics INFORMATICA ISSN(p): 0350-5596, Volume-45, Issue-2, Juin 2021, <http://www.informatica.si/index.php/informatica/index>
- [24] Jain,N., Rahul,K. , Khamaru,I. , Jha,A.K. , Ghosh,A.,(2017),"Hand Written Digit Recognition using Convolutional Neural Network (CNN)",International Journal of Innovations & Advancement in Computer Science, Volume 6, Issue 5.
- /[25]Alwzway,H.A.,Albehadili,H.M.,Alwan,Y.S.,Islam,N.E.,(2016),"HandwrittenDigit Recognition Using Convolutional Neural Networks",International Journal of Innovative Research in Computer and Communication Engineering , Vol. 4, Issue 2.
- [26] <https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/> visited (28/4/2021)
- [27] <https://medium.com/deep-math-machine-learning-ai/different-types-of-machinelearning-and-their-types-34760b9128a2> visited (25/4/2018).
- [28] Voir Machine Learning , chap. 13 Reinforcement Learning, p. 367-390.
- [29] Voir Machine Learning, p. 373-380.

Bibliographie

- [30] Gadri S, Neuhold E(2020) Building Best Predictive Models Using ML and DL [30Gadri S ,Neuhold E.Approchoaches to Categorize Fashion Clothes.In Rutkowski L,Scherer R,Korytkowski M,Pedrycz W (Computer Science, vol 12415.Springer ,cham.https://doi.org/10.1007/978-3-030-61401-0_9
- [31] Voir Machine Learning, chap. 4 Artificial Neural Networks, p. 81-127
- [32] Voir Machine, chap. 3 Decision Tree Learning, p. 52-80.
- [33] Voir Machine Learning, chap. 9 Genetic Algorithms, p. 249-273.
- [34] Saha, Roshni, (2019). "Classification of Parkinson's Disease Using MRI Data and Deep Learning Convolution Neural Networks" Creative Components. 241.
- [35] Chollet, F., (2018), Deep Learning with python, by Manning Publications Co. All rights reserved.
- [33] Shashank Mishra, D. Malathi, K. Senthilkumar De^\$pt,(2018), DIGIT RECOGNITION USING DEEP LEARNING, International Journal of Pure and Applied Mathematics Volume 118 No. 22.
- [36]- <http://www.trytoprogram.com>.
- [37]-"The Making of Python".Artima Developer. Retrieved March 22,2007.
- [38]-van Rossum, Guido (January 20,2009)."A Brief Timeline of Python" .Retrieved November 29,2019.
- [39]-"Why was Python created in the first place?".Python FAQ. Retrieved March 22,2007.
- [40]-"Releases / python .org".Retrieved November 29,2019.
- [41]-"Drake, Fred L, Jr. (July 25,2000)."PEP 160-Python 1.6 Release Schedule".Retrieved November 29,2019.
- [42]- "Download Python / Python. Org". Retrieved November 29 ,2019.
- [43]-Hylton, Jeremy."PEP 200-Python 2.0 Release Schedule".Retrieved November 29,2019.
- [44]-Hylton, Jeremy (October 16,2000)."PEP 226-Python 2.1 Release Schedule".Retrieved November 29,2019.
- [45]-Warsaw, Barry; van Rossum, Guido (April 17,2001)."PEP 251 –PYTHON 2.2 Release Schedule" .Retrieved November 29,2019.
- [46]-van Rossum, Guido (February 27,2002)." PEP 283-Python 2.3 Release Schedule".Retrieved November 29,2019.
- [47]-Warsaw, Barry;Hettinger, Raymond;Baxter, Anthony (July 29,2003)."PEP 320-Python 2.4 Release Schedule".Retrieved November 29,2019.
- [48]-Neorwitz, Neal; Van Rossum, Guido; Baxter, Anthony (February 7, 2006)."PEP 356-PYTHON 2.5 Release Schedule".Retrieved November 29,2019.
- [49]- Peterson, Benjamin (November 3, 2008)."PEP 373-python 2.7 Release Schedule".Retrieved April.
- [50]-"Peterson, Benjamin (February 8, 2009)."PEP 375- python 3.1 Realease Schedule".Retrieved November 29, 2019.
- [51]-Brandl, Georg (Décembre 30, 2009). "PEP 392-python 3.2 Release Schedule". Retrieved November 20,2019.
- [52]-Brandl, Georg (March 23, 2011). "PEP 398-Python 3.3 Release Schedule".Retrieved November 29,
- [53]-Hastings, Larry (October 17, 2012). "PEP 429 – Python 3.4 Release Schedule" Retrieved November 29, 2019.

Bibliographie

- [54]-Hastings, Larry (September 22,2014)."PEP 478- Python 3.5 Release Schedule".Retrieved March 17, 2020.
- [55]-Deily, Ned (May 30, 2015). "PEP 494- Python 3.6 Release Schedule".Retrieved April 20, 2020.
- [56]- Deily, Ned (December 23, 2016). "PEP 537- Python 3.7 Release Schedule". Retrieved April 20, 2020.
- [57]-Langa, Lukasz (January 27, 2018)."PEP 569-Python 3.8 Release Schedule" .Retrieved April 20,2020.
- [58]-Langa, Lukasz (October 13.2020). "PEP 596 – Python 3.9 Release Schedule" .Retrieved October 13, 2020.
- [59]-Pablo, Pablo (May 25, 2020)."PEP 619-Python 3.10 Release Schedule".Retrieved May 25,2020.
- [60]Nour El Houda A, said Gadri: *Traffic signs recognition for autonomous cars*, University Msila , 2020.
- [61]Hassina T, *Developing an Efficient CNN Model to Recognize Handwritten Digits*, **Msila** University, 2020
- [62] Said G. Efficient Arabic handwritten character recognition based on machine learning and deep learning approaches, The 2nd International Conference on Advances in Intelligent Systems, Soft Computing and Optimization Techniques 2020 (ICAISCO 2020) held in Penang, Malaysia, 01 - 02, April 2020 (**Scopus Indexed Conference**).
- [63]- Dalcín, L., Paz, R., & Storti, M. (2005). MPI for Python. Journal of Parallel and Distributed
- [64]- Bharath Ramsundar, TensorFlow Tutorial, 2015
- [65]- Rudraswamimath, V.R., Bhavanishankar, K.,(2019),"Handwritten Digit Recognition using CNN",International Journal of Innovative Science and Research Technology.
- [66] D. Beniz, and A. Espindola. using tkinter of python to create graphical user interface (gui) for scripts in lnls. in: pcapac2016. Brazil, p.2.
- [67]- <https://www.anaconda.com/what-is-anaconda/>
- [68]-Christine Doig, «Anaconda for R users: SparkR and rBokeh »[archive],sur Developer Blog,Continuum Analytics 1-février 2016
- [69]- <http://conda.pydata.org/docs/>.
- [70] LAOUES Baya, Said Gadri : *Développement d'un modèle ANN efficace pour détecter et classifier des patients diabétiques*, University **Msila** , 2020.
- [71] Brian Heinold ,A Practical Introduction to Python Programming, Department of Mathematics and Computer Science Mount St. Mary's University,2012.
- [72] 2007. International Vocabulary of Metrology: Basic and General Concepts and Associated Terms (VIM) =. 3rd ed. Genève: international organization of legal metrology, p.50.
- [73] Theanets.readthedocs.io. 2020. LossFunctions — Theanets 0.7.3 Documentation. URL: <https://theanets.readthedocs.io/en/stable/api/losses.html>
- [74] Medium. 2020. VariousOptimizationAlgorithms for Training Neural Network. URL: <https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6>.
- [75] Google Developers. 2020. Classification: Accuracy | Machine Learning Crash Course. URL: <https://developers.google.com/machine-learning/crash-course/classification/accuracy>.

Résumé:

Avec la propagation des réseaux sociaux, de nombreuses personnes ont commencé à faire confiance à l'opinion exprimée dans les publications sur ces réseaux et sites web. Comme l'une des références pour développer le tourisme, où le tourisme peut accéder à diverses sources d'information. Sur cette perspective vient notre étude. Elle a comme objectif principal de passer en revue certains des différents modèles antérieurs dans le domaine de l'analyse des sentiments et de l'exploration des opinions des gens sur les médias sociaux, et la suggestion d'un nouveau modèle qui prend en compte les opinions des gens sur le même endroit dans différentes régions ou pays, où l'étude a révélé un manque de modèles précédents car ils ne couvraient pas ce aspect important. Dans le présent travail, nous avons conçu un système intelligent pour reconnaître le degré de satisfaction des clients des agences de tourisme. Dans une première étape, nous avons utilisé des algorithmes d'apprentissage automatique, notamment : LR, LDA, KNN, CART, NB et SVM. Enfin, nous avons développé un modèle ANN multicouche pour améliorer la précision de l'analyse et d'exploration.

Mots clés: l'apprentissage automatique, reconnaître le degré de satisfaction des clients, le domaine de l'analyse des sentiments, un système intelligent.

ملخص : مع انتشار شبكات التواصل الاجتماعي بدأ الكثير من الأشخاص يعتمدون على الآراء الواردة في هذه الشبكات والمواقع كأحد المراجع الأساسية لتطوير السياحة، حيث يمكن للسياح الآن الوصول إلى مصادر مختلفة من المعلومات ويمكنهم إنشاء المحتوى الخاص بهم ومشاركة وجهات نظرهم وخبراتهم مما يساعدهم في اتخاذ القرار الصحيح. ولكن بسبب كثرة الآراء المتوفرة في شبكات التواصل الاجتماعي حول مختلف المواضيع فمن الصعب الرجوع إلى مئات وآلاف التعليقات والآراء مما يجعل المعالجة اليدوية شبه مستحيلة. وعلى هذا الأساس جاءت هذه الدراسة بهدف مراجعة بعض النماذج المختلفة السابقة في مجال تحليل المشاعر والتنقيب على آراء الناس داخل مواقع التواصل الاجتماعي، واقتراح نموذج مطور يأخذ بعين الاعتبار هذه الآراء حول نفس المكان باختلاف المنطقة والدولة، حيث وجدت الدراسة نقصاً واضحاً في النماذج السابقة كونها لم تغطي هذا الجانب المهم. في العمل الحالي، قمنا بتصميم نظام ذكي للتعرف على درجة الرضا لدى زبائن الوكالات السياحية، اطلعنا في المرحلة الأولى على خوارزميات التعلم الآلي، بما في ذلك: LR و LDA و KNN و CART و NB و SVM. في الأخير، قمنا بتطوير نموذج ANN يتكون من عدة طبقات لتحسين دقة الكشف.

الكلمات المفتاحية: التعلم الآلي، التعرف على درجة الرضا لدى زبائن، مجال تحليل المشاعر، نظام ذكي.

Abstract:

With the spread of social networks, many people have started to trust the opinions expressed in the posts on these networks and websites. As one of the references to develop tourism, where tourism can access various sources of information from this perspective comes our study. Its main purpose is to review some of the different past models in the field of sentiment analysis and exploring people's opinions on social media, and suggesting a new model that takes into account the opinions of people on the same place in different regions or countries, where the study found a lack of previous models as they did not cover this important aspect. In the present work, we have designed an intelligent system to recognize the degree of satisfaction of customers of tourism agencies. In a first step, we used machine learning algorithms, including: LR, LDA, KNN, CART, NB and SVM. Finally, we have developed a .multilayer ANN model to improve the precision of analysis and exploration.

Keywords: Apprentissage automatique, reconnaître le degré de satisfaction of customers, sentiment analysis, intelligent system.