

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique  
Université A.MIRA-BEJAIA



Faculté des Sciences Exactes  
Département d'informatique

## Mémoire de Fin De Cycle

En vue de l'obtention du diplôme de MASTER Recherche en  
Informatique

Option : Intelligence Artificielle

### Thème

---

*L'aide à la décision basée sur les outils  
d'intelligence artificielle*

---

Réalisé par :

*M<sup>lle</sup>* Benhellal Melaz

Soutenu le 2 Juillet 2023 devant le jury composé de :

Présidente	Mme EL BOUHISSI Houda	MCA	U. A/Mira Béjaïa.
Examinatrice	Mme Zemouche Djamila	MCB	U. A/Mira Béjaïa.
Promotrice	Mme Ait Hacene Souhila	MAA	U. A/Mira Béjaïa.
Promotrice	Mme Chibani Samia	MCA	U. A/Mira Béjaïa.

Promotion : 2022/2023

---

## Remerciements

---

Avant tout nous tenons à remercier notre Dieu , le miséricordieux et tout-puissant, qui nous a donné le courage pour accomplir brillamment notre parcours universitaire. Que sa guidance nous accompagne tout au long de notre vie future.

Tout d'abord, ce travail n'aurait pas pu voir le jour et être aussi riche sans l'aide et l'encadrement de Madame Ait Hacen Souhila et Madame Chibani Samia. Nous les remercions de nous avoir encouragés dans cette démarche et d'avoir mis à notre disposition le temps nécessaire pour sa réalisation. Nous exprimons notre gratitude pour leur engagement et leur soutien, ainsi que pour la pertinence de leurs remarques, de leurs retours et leur souci du détail, qui ont contribué à la réalisation de ce mémoire.

Nos remerciements vont également aux membres du jury d'avoir accepté d'examiner notre travail et de l'enrichir par leurs propositions.

Si nous en sommes là aujourd'hui, c'est également grâce à la qualité de l'encadrement pédagogique et à la présence d'esprit remarquable des différents enseignants qui ont assuré notre formation. Nous leur adressons nos remerciements les plus sincères.

---

## Dédicace

---

Je suis honoré de dédier ce humble travail à tous ceux qui ont été ma source d'inspiration et de soutien tout au long de ce parcours. Vos encouragements, votre amour et votre soutien inconditionnel ont été les fondements de ma réussite.

Je souhaite tout particulièrement dédier ce travail à ma mère, qui m'a toujours encouragé à poursuivre mes rêves et à travailler dur pour atteindre mes objectifs. Je dédie également ce travail à la mémoire de mon père, qui continue de vivre dans mon cœur.

Je tiens à exprimer ma gratitude à mes frères Aissa, Razik et Omar, qui ont toujours été à mes côtés. Votre confiance en moi m'a donné la détermination nécessaire pour surmonter tous les défis.

Je souhaite également dédier ce travail à mes chères amies Wissam, Lahna, Salima et Nina, qui ont été présentes tout au long de ce parcours. Votre soutien inconditionnel et votre amitié précieuse ont été d'une valeur inestimable.

Ce mémoire vous est dédié à tous, car c'est grâce à votre amour, votre soutien et votre confiance en moi que j'ai pu réaliser ce travail. Je vous suis éternellement reconnaissant et je dédie ces réussites à vous tous.

Melaz Benhellal

# Table des matières

<b>1 La prise de décision</b>	<b>7</b>
1.1 Introduction	7
1.2 La prise de décision	7
1.2.1 Types de décisions	7
1.2.2 Les approches de prise de décision	8
1.2.3 Les limites de la prise de décision humaine	9
1.3 L'intelligence artificielle	9
1.3.1 L'intelligence artificielle pour la prise de décision	9
1.3.2 Domaines d'applications de l'IA pour la prise de décision	10
1.3.3 Les avantages de l'utilisation de l'intelligence artificielle pour la prise de décision	10
1.4 Conclusion	10
<b>2 L'intelligence Artificielle</b>	<b>12</b>
2.1 Introduction	12
2.2 Définition de l'ntelligence artificielle	12
2.3 Apprentissage automatique (Le machine learning(ML))	13
2.3.1 A quoi sert le Machine Learning	13
2.3.2 Different types de l'apprentissage automatique	14
2.4 Apprentissage profond (deep learning(DL))	16
2.4.1 Réseaux de neurones artificielle :	17
2.4.2 L'architecture des réseaux de neurones profonds :	17
2.4.3 Les différents algorithmes de l'Apprentissage profond	18
2.5 Entraînement des réseaux de neurones	19
2.5.1 L'objectif de l'entraînement d'un réseau de neurones	20
2.5.2 Les fonctions d'activation	20
2.6 Systèmes de recommandation	21
2.6.1 Comment fonctionnent les systèmes de recommandation	22
2.7 Conclusion	24
<b>3 Etat de l'art</b>	<b>25</b>
3.1 Introduction	25
3.2 Etat de l'art	25

<b>3.3 Conclusion</b> . . . . .	31
<b>4 Approche proposée SR-DNN</b>	<b>35</b>
<b>4.1 Introduction</b> . . . . .	35
<b>4.2 Approche proposée SR-DNN</b> . . . . .	35
<b>4.2.1 Les étapes de SR-DNN</b> . . . . .	35
<b>5 Implémentation</b>	<b>44</b>
<b>5.1 Introduction</b> . . . . .	44
<b>5.2 Etude de cas</b> . . . . .	44
<b>5.3 Conclusion</b> . . . . .	46

# Table des figures

2.1 Intelligence artificielle, machine learning et deep learning . . . . .	13
2.2 Algorithme de Classification . . . . .	15
2.3 Algorithme de Régression . . . . .	15
2.4 Algorithme de Clustering . . . . .	16
2.5 Architecture des réseaux de neurones . . . . .	17
2.6 Réseau neuronal convolutif . . . . .	18
2.7 Réseaux de neurones récurrents . . . . .	19
2.8 Recommandation basées sur la connaissance . . . . .	22
2.9 Recommandation basées sur le contenu . . . . .	23
2.10 Recommandation basées sur le filtrage collaboratif . . . . .	23
2.11 Recommandation hybrid . . . . .	24
4.1 Les étapes de SR-DNN . . . . .	36
4.2 Le dataset . . . . .	37
4.3 Le dataset après la normalisation et l'encodage . . . . .	38
4.4 L'architecture de RS-DNN . . . . .	39
4.5 L'entraînement du RS-DNN . . . . .	40
4.6 Évaluation des métriques d'erreur . . . . .	42
4.7 Comparaison des valeurs réelles 'etoile' et prédites avec coefficient de détermination . . . . .	43
5.1 Les informations fournit par l'utilisateur . . . . .	44
5.2 Un nouveau DataFrame . . . . .	45
5.3 Une destinations et un types d'hébergement encoder . . . . .	45
5.4 Résultats d'exécution de notre approche . . . . .	45

# Liste des tableaux

3.1 Étude comparative des travaux connexes 1 . . . . .	34
--------------------------------------------------------	----

# Liste des Abréviations

**AE** Auto-Encodeur. [19](#)

**CNN** Réseaux de neurones convolutionnel. [18](#)

**DL** Deep Learning. [1](#), [12](#), [16](#)

**DNN** Réseau de neurones profond. [18](#)

**IA** Intelligence Artificielle. [6](#)

**LSTM** (Long Short-Term Memory. [19](#)

**MAE** Mean Absolute Error). [41](#), [42](#)

**ML** Machine learning. [12](#)

**MSE** Mean Squared Error. [40](#)

**ReLU** Rectified Linear Unit. [21](#), [38](#), [39](#)

**RMSE** Root Mean Squared Erro. [41](#), [42](#)

**RNN** Réseaux de neurones récurrent. [18](#)

**SR-DNN** système de recommandation basée sur les Réseaux de neurones profond.  
[35](#)

**Tanh** Tangente Hyperbolique. [21](#)

# Introduction

La prise de décision joue un rôle fondamental dans notre vie quotidienne, que ce soit pour des choix personnels ou professionnels. Cependant, prendre des décisions éclairées peut s'avérer complexe face à la multitude de facteurs à considérer et à la quantité croissante d'informations disponibles. C'est dans ce contexte que l'intelligence artificielle **IA** a fait son apparition pour offrir de nouvelles perspectives pour améliorer le processus de prise de décision. Grâce à ses algorithmes sophistiqués, l'IA peut rapidement traiter d'énormes quantités de données et fournir des recommandations basées sur des preuves solides. L'un des outils clés pour aider à la prise de décision est la recommandation basée sur l'intelligence artificielle. Ce domaine de recherche vise à développer des systèmes de recommandation capables de proposer des suggestions pertinentes et personnalisées en fonction des préférences et des besoins de chaque utilisateur. Ces systèmes tiennent compte de multiples facteurs, tels que l'historique de navigation, les interactions précédentes et les données démographiques, pour offrir des recommandations précises et adaptées. Parmi les techniques d'IA utilisées dans la recommandation, le Deep Learning, ou apprentissage profond, a suscité un intérêt considérable. Le Deep Learning s'inspire du fonctionnement du cerveau humain pour développer des algorithmes d'apprentissage automatique complexes. Dans le cadre de notre étude, nous avons exploré différents articles scientifiques qui abordent le sujet de la recommandation basée sur le Deep Learning. Nous nous sommes intéressés à son application dans divers domaines tels que le tourisme, la recommandation de cours, de films, de musique, et bien d'autres. Cette analyse approfondie nous permettra de mieux comprendre l'impact de la recommandation basée sur le Deep Learning dans notre société moderne, ainsi que ses avantages, ses limites et les défis auxquels elle fait face. De plus, nous avons proposé une approche novatrice de recommandation basée sur le Deep Learning spécifiquement adaptée au domaine du tourisme. Nous avons mis en œuvre cette approche et évalué son efficacité dans la recommandation d'un séjour touristique. En intégrant les techniques avancées du Deep Learning, notre approche vise à fournir des recommandations personnalisées et pertinentes pour les voyageurs, améliorant ainsi leur expérience touristique globale.

# Chapitre 1

## La prise de décision

### 1.1 Introduction

Dans ce premier chapitre, nous allons présenter plusieurs définitions et concepts clés liés à la prise de décision. Nous examinerons les différents types de décisions et mettrons en évidence les limites et les défis auxquels elles font face. De plus, nous aborderons l'intégration de l'intelligence artificielle comme une solution prometteuse pour surmonter ces limitations et améliorer le processus de prise de décision.

### 1.2 La prise de décision

La prise de décision peut être définie comme le processus cognitif par lequel un individu ou un groupe d'individus évalue diverses options ou alternatives et sélectionne celle jugée la plus appropriée ou la plus bénéfique pour atteindre un objectif particulier. Cela comprend l'analyse des informations disponibles, la pesée des résultats possibles, la prise en compte des préférences individuelles ou de groupe et l'évaluation des risques et des avantages associés à chaque option. Les décisions sont prises dans de nombreux domaines de la vie des décisions personnelles quotidiennes aux décisions stratégiques complexes dans les affaires et d'autres professions. ces décisions sont souvent influencés par des facteurs cognitifs, émotionnels, sociaux et contextuels et varient en termes de rationalité, de degré d'incertitude et de complexité. [1]

#### 1.2.1 Types de décisions

Dans cette section, nous allons aborder les différents types de décisions. [2]

- **Décisions programmées** Ce sont des décisions qui sont prises de manière répétitive et suivent des règles ou des procédures établies. Elles sont souvent basées sur des schémas préétablis et peuvent être automatisées. Les décisions programmées sont prises dans des situations familières et courantes, où les solutions

optimales sont connues à l'avance. Par exemple, un gestionnaire peut avoir une procédure établie pour approuver les demandes de congé des employés, en suivant des critères prédéfinis.

- **Décisions stratégiques** Ce sont des décisions prises au plus haut niveau de l'organisation qui ont un impact significatif sur la direction et l'orientation de l'entreprise. Les décisions stratégiques sont souvent prises par la direction ou les cadres supérieurs et peuvent affecter la vision, les objectifs, les partenariats, les investissements majeurs, etc.
- **Décisions tactiques** ces décisions se situent au niveau intermédiaire de l'organisation et sont généralement prises par des responsables de différents départements ou divisions. Ils sont liés à la mise en œuvre des décisions stratégiques et visent à atteindre des objectifs opérationnels à court ou moyen terme. Par exemple, l'allocation des ressources, la planification de la main-d'œuvre ou la gestion des stocks.
- **Décisions opérationnelles** il s'agit de décisions prises au niveau le plus bas de l'organisation, souvent par le personnel et les supérieurs opérationnels. Ils concernent les tâches et processus quotidiens nécessaires au bon fonctionnement de l'entreprise. Les décisions opérationnelles peuvent inclure la planification des horaires de travail, l'allocation des ressources, la résolution des problèmes techniques, etc.

## 1.2.2 Les approches de prise de décision

Dans cette partie, nous allons présenter les principales approches de prise de décisions [3](#)

- **Prise de décision rationnelle** : Cette approche est basée sur une analyse systématique et logique des informations disponibles. Cela implique de rassembler et d'évaluer toutes les données pertinentes, d'identifier les alternatives possibles, de peser le pour et le contre de chaque option, puis de sélectionner la meilleure solution en fonction de critères prédéfinis. La prise de décision rationnelle vise à maximiser les résultats en utilisant une approche objective basée sur le calcul.
- **Prise de décision intuitive** : Cette approche repose sur l'utilisation de l'expérience, de l'intuition et du jugement personnel. Les décisions intuitives sont prises rapidement et souvent automatiquement sans avoir à passer par une analyse détaillée des informations. Cette approche est souvent utilisée dans des situations familières où les décideurs ont développé une expertise et une connaissance approfondie du domaine concerné.

- **Prise de décision basée sur les données** : cette approche est basée sur l'utilisation de données et de modèles analytiques pour prendre des décisions. Cela comprend la collecte de données pertinentes, l'analyse statistique et l'identification de modèles ou de tendances. Cette approche permet de prendre des décisions objectives et factuelles en utilisant des preuves concrètes plutôt que l'intuition subjective.

### 1.2.3 Les limites de la prise de décision humaine

Il est essentiel de prendre en compte certaines limites inhérentes à la prise de décision humaine. Nous allons vous présenter certaines de ces limitations :

- Limitations de la mémoire et de l'attention : les êtres humains ont une capacité de mémoire et d'attention limitée. Lors de la prise de décision, il est possible de négliger certaines informations importantes ou de ne pas retenir toutes les données
- Émotions et états mentaux : Les émotions et les états mentaux peuvent influencer la prise de décision. Des facteurs tels que le stress, la fatigue, l'humeur
- Risque d'erreur : Même avec les meilleures intentions et les efforts les plus sincères, les êtres humains sont sujets à l'erreur.[\[4\]](#)

## 1.3 L'intelligence artificielle

L'intelligence artificielle (IA) est un domaine informatique passionnant qui a connu d'énormes progrès ces dernières années. Grâce à des techniques avancées telles que l'apprentissage automatique et les réseaux de neurones. L'IA aspire à révolutionner la prise de décision en surmontant les limites inhérentes à la prise de décision humaine.

### 1.3.1 L'intelligence artificielle pour la prise de décision

L'intelligence artificielle (IA) est utilisée pour améliorer la prise de décision de diverses manières. En utilisant des techniques avancées telles que l'apprentissage automatique et l'analyse de données, l'IA peut traiter de grandes quantités d'informations et détecter des modèles qui échappent souvent à l'observation humaine. Cela permet de prendre des décisions plus éclairées et plus précises, réduisant ainsi les erreurs humaines et les préjugés. De plus, l'IA peut automatiser certaines tâches de prise de décision, permettant aux professionnels de se concentrer sur des aspects plus stratégiques et complexes. En combinant les capacités de calcul rapide de l'IA avec l'expertise humaine, des décisions plus rapides, plus précises et fondées sur des preuves peuvent être prises qui peuvent avoir un impact significatif dans de nombreux domaines tels que la médecine, la finance, la logistique et bien d'autres.[\[5\]](#)

### 1.3.2 Domaines d'applications de l'IA pour la prise de décision

- Finance : l'IA est utilisée pour l'analyse des marchés financiers, la détection des fraudes, la gestion des risques et la prise de décision en matière d'investissement.
- Santé : l'IA est utilisée pour analyser des données médicales, diagnostiquer des maladies, recommander des traitements et prédire des résultats médicaux.
- Logistique : l'IA est utilisée pour optimiser la gestion de la chaîne d'approvisionnement, la planification des itinéraires, l'optimisation des stocks et la gestion des entrepôts.
- Industrie : l'IA est utilisée pour la maintenance prédictive des équipements, l'optimisation des processus de fabrication, la gestion des ressources et la prise de décision en temps réel.
- E-commerce : l'IA est utilisée pour personnaliser les recommandations, la segmentation des clients, la tarification dynamique et la gestion des stocks.

### 1.3.3 Les avantages de l'utilisation de l'intelligence artificielle pour la prise de décision

L'intelligence artificielle présente de nombreux avantages dans la prise de décision, notamment :

- Améliorée la Précision : L'intelligence artificielle est capable d'analyser rapidement et avec précision de vastes quantités de données, ce qui permet de prendre des décisions plus précises .
- Apprentissage continu : Les systèmes d'IA ont la capacité d'apprendre en permanence à partir de nouvelles données, ce qui leur permet de s'améliorer au fil du temps et de prendre des décisions de plus en plus précises à mesure qu'ils accumulent de l'expérience.
- Rapidité de traitement : Grâce à leur capacité à traiter rapidement l'information, les systèmes d'IA peuvent prendre des décisions en temps réel, ce qui permet d'accélérer les processus décisionnels et d'obtenir des résultats plus rapidement.
- Gestion des informations complexes : L'intelligence artificielle est capable de traiter et de comprendre des informations complexes de manière plus efficace que les humains. Grâce à ses algorithmes avancés[6]

## 1.4 Conclusion

En conclusion, la prise de décision est un processus essentiel dans la vie quotidienne et dans le contexte professionnel. Cependant, les décideurs humains sont confrontés à des limitations et des défis inhérents, tels que les biais cognitifs et les erreurs de jugement. L'intelligence artificielle (IA) apporte une contribution cruciale dans ce contexte. L'IA offre de nouvelles possibilités passionnantes pour améliorer le processus de prise de décision. Grâce à des techniques avancées, tels que l'analyse de données avancée et

l'apprentissage automatique, l'IA peut fournir des analyses précises et automatiser certaines tâches décisionnelles. Elle peut également identifier des modèles ou des tendances difficiles à détecter pour les humains. Les applications de l'IA dans divers domaines tels que la finance, la santé, la logistique et le marketing démontrent son potentiel en matière de prise de décision. Cependant, l'utilisation de l'IA soulève également des défis et des considérations éthiques, notamment en ce qui concerne la confidentialité des données et la responsabilité en cas d'erreurs. Il est essentiel d'aborder ces questions de manière responsable et éthique tout en tirant parti des avantages que l'IA peut offrir. En somme, l'intégration de l'IA dans le processus de prise de décision présente de grandes opportunités pour améliorer la précision, la vitesse et l'efficacité, tout en garantissant une prise de décision éthique et responsable.

# Chapitre 2

## L'intelligence Artificielle

### 2.1 Introduction

Dans ce chapitre, nous plongeons dans les domaines passionnants de l'intelligence artificielle (IA), de l'apprentissage automatique **ML** et de l'apprentissage en profondeur. Nous commençons par définir l'intelligence artificielle, qui comprend les techniques et les systèmes capables de reproduire l'intelligence humaine. Ensuite, nous explorerons ML, une approche de l'IA qui permet aux machines d'apprendre à partir des données sans être explicitement programmées. Enfin, nous discutons de l'apprentissage en profondeur **DL**, une avancée du ML qui utilise des réseaux de neurones artificiels pour résoudre des problèmes complexes en apprenant une représentation hiérarchique des données. De plus, nous aborderons un outil de recommandation spécifique basé sur ces techniques qui vise à améliorer la prise de décision en fournissant des recommandations pertinentes.

### 2.2 Définition de l'intelligence artificielle

L'intelligence artificielle (IA) est une branche de l'informatique qui vise à développer des systèmes et des machines capables de simuler l'intelligence humaine. L'intelligence artificielle comprend des techniques et des algorithmes qui permettent aux machines d'apprendre, de raisonner, de percevoir et de résoudre des problèmes de manière autonome. L'IA vise à reproduire des capacités cognitives telles que la compréhension du langage, la prise de décision, la reconnaissance de formes et l'apprentissage pour effectuer des tâches complexes et améliorer les performances dans divers domaines **[7]**. La figure 2.1 représente le schéma de l'intelligence artificielle, qui inclut notamment le domaine de l'apprentissage automatique (machine learning), lui-même subdivisé en différents sous-domaines, dont l'apprentissage profond (deep learning).

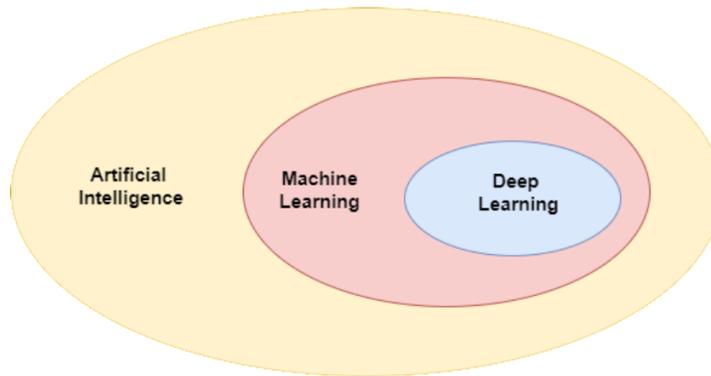


FIGURE 2.1 – Intelligence artificielle, machine learning et deep learning [8]

## 2.3 Apprentissage automatique (Le machine learning(ML))

L'apprentissage automatique est une discipline de l'intelligence artificielle axée sur le développement de méthodes permettant aux ordinateurs d'apprendre et de s'améliorer à partir de données, sans programmation explicite. Plutôt que de suivre des instructions précises, les algorithmes d'apprentissage automatique sont conçus pour identifier des modèles et des relations dans les données, leur permettant de prendre des décisions ou de prédire des résultats avec un certain degré de précision. Utilisant des techniques telles que la régression, la classification, le clustering et les réseaux de neurones, l'apprentissage automatique permet aux machines d'acquérir des connaissances et de s'adapter à de nouvelles situations. [9]

### 2.3.1 A quoi sert le Machine Learning

L'apprentissage automatique, également connu sous le nom de Machine Learning, est un domaine de l'intelligence artificielle qui ouvre de vastes perspectives et offre de multiples applications. Parmi celles-ci, voici quelques exemples significatifs :

- Classification et reconnaissance de modèles : l'apprentissage automatique peut classer les données dans des catégories prédéfinies et reconnaître des modèles dans des ensembles de données volumineux. Par exemple, il peut être utilisé pour la détection de fraude, la reconnaissance vocale, la reconnaissance d'images et la traduction automatique. [10]
- Prévisions et prédictions : l'apprentissage automatique permet de créer des modèles prédictifs à partir de données historiques. Ces modèles peuvent être utilisés pour prédire des événements futurs ou estimer des valeurs manquantes. Par exemple, il peut être utilisé pour la prévision de la demande, la prévision des

prix, la prévision météorologique et la détection des tendances. [11]

- Optimisation et prise de décision : l'apprentissage automatique résout les problèmes d'optimisation et aide à la prise de décision. Par exemple, il peut être utilisé pour optimiser les processus de production, la planification des ressources, recommander des produits ou du contenu et optimiser les stratégies marketing. [12]
- Apprentissage automatique et automatisation : L'apprentissage automatique permet de créer des systèmes autonomes capables d'apprendre et de s'adapter à de nouvelles situations. Il est utilisé dans les véhicules autonomes, les robots industriels, les assistants virtuels et autres systèmes d'intelligence artificielle. [13]
- Analyse des données et extraction des connaissances : l'apprentissage automatique peut extraire des informations précieuses à partir de grandes quantités de données. Il peut être utilisé pour analyser des données non structurées, découvrir des relations cachées, effectuer un regroupement et une analyse des sentiments. [14]

En bref, l'apprentissage automatique est utilisé pour résoudre des problèmes complexes, automatiser des tâches, prendre des décisions basées sur des données et extraire des connaissances à partir de grandes quantités d'informations. Son utilisation est répandue dans de nombreuses industries telles que la santé, la finance, la technologie, la science, l'industrie et bien d'autres.

### 2.3.2 Différents types de l'apprentissage automatique

Il existe plusieurs types d'apprentissage automatique, qui sont des approches différentes pour entraîner les modèles et effectuer des prédictions. Voici quelques-uns des types les plus courants : [15]

#### **Apprentissage supervisé :**

Dans cette approche, les modèles sont entraînés à l'aide d'un ensemble de données d'entraînement étiquetées, où chaque exemple est associé à une étiquette ou à une valeur cible. Le modèle apprend à prédire l'étiquetage des données non étiquetées sur la base d'exemples étiquetés. L'apprentissage supervisé se concentre sur la classification et la régression. Ces approches aident à résoudre diverses tâches en attribuant des étiquettes ou en prédisant des valeurs. Ils sont essentiels pour construire des modèles prédictifs à partir de données étiquetées.

- La Classification : Dans un problème de classification, l'objectif est de regrouper des objets ou des échantillons dans différentes classes prédéfinies. La tâche consiste à prédire la valeur d'une variable discrète qui représente la classe à laquelle chaque objet appartient. un Algorithme de classification est présenter

dans la figure 2.2

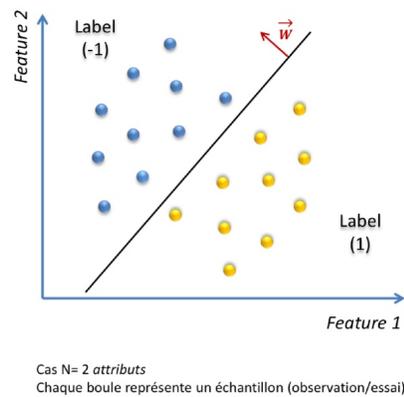


FIGURE 2.2 – Algorithme de Classification

[16]

- La Régression : Les problèmes de régression consistent à essayer de prédire la valeur d'une variable continue, c'est-à-dire une variable qui peut prendre un nombre infini de valeurs. un Algorithme de régression est présenter dans la figure 2.3

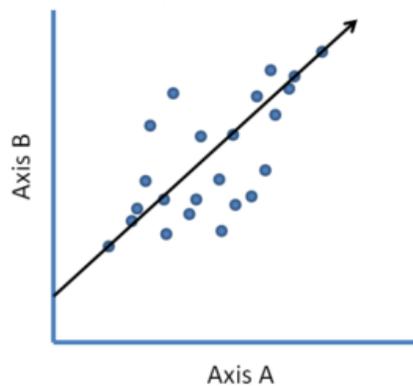


FIGURE 2.3 – Algorithme de Régression

[17]

### Apprentissage non supervisé :

Contrairement à l'apprentissage supervisé, l'apprentissage non supervisé n'a pas d'étiquettes dans les données d'apprentissage. Le modèle recherche des structures internes ou des modèles dans les données pour regrouper ou réduire leur dimensionnalité.

- Clustering : le sous-problème d'apprentissage non supervisé est le regroupement en classes homogènes qui consiste à représenter un nuage de points d'espace arbitraire dans un ensemble de groupes connus sous le nom cluster. Il s'agit d'un traitement sur un ensemble d'objets qui n'ont pas été marqués par le gestionnaire. L'algorithme doit découvrir la structure par lui-même à partir des données. un Algorithme de clustering est présenter dans la figure 2.4.

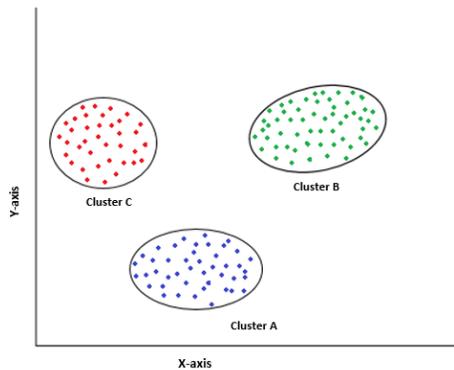


FIGURE 2.4 – Algorithme de Clustering

[18]

### Apprentissage semi-supervisé :

Cette approche combine des éléments d'apprentissage supervisé et non supervisé. Il utilise à la fois des données étiquetées et non étiquetées pour entraîner le modèle. Cela peut être utile lorsque l'étiquetage des données est coûteux ou difficile.

### Apprentissage par renforcement :

Dans cette approche, le modèle apprend à prendre des décisions séquentielles en interagissant avec l'environnement. Le modèle reçoit des récompenses ou des punitions en fonction de ses actions, ce qui lui permet d'apprendre quelles actions sont optimales dans différentes situations.

## 2.4 Apprentissage profond (deep learning(DL))

L'apprentissage en profondeur est une sous-catégorie de l'apprentissage automatique qui implique l'utilisation de réseaux de neurones artificiels profonds pour effectuer des tâches d'apprentissage à partir de données. Les réseaux de neurones profonds artificiels sont capables de modéliser et extraire des caractéristiques complexes à partir de données brutes, permettant la création de modèles d'apprentissage automatique hautes performances dans des domaines tels que la reconnaissance d'images, la reconnaissance

vocale, la traduction automatique, la génération de texte et la prédiction. L'apprentissage en profondeur est particulièrement efficace pour les tâches qui nécessitent de grandes quantités de données et une forte capacité de généralisation.. [19]

### 2.4.1 Réseaux de neurones artificielle :

Un réseau neuronal artificiel est un modèle mathématique qui imite le fonctionnement du cerveau humain. Il se compose de plusieurs unités de traitement appelées neurones. Chaque neurone reçoit des signaux d'entrée, les pèse à l'aide de poids, effectue des calculs et génère une sortie. Les poids représentent l'importance accordée à chaque signal d'entrée dans le calcul. En utilisant des méthodes d'apprentissage telles que la rétropropagation d'erreurs, le réseau de neurones apprend à ajuster les poids pour mieux répondre à des tâches spécifiques [20]. la figure 2.5 représente l'architecture des réseaux de neurones

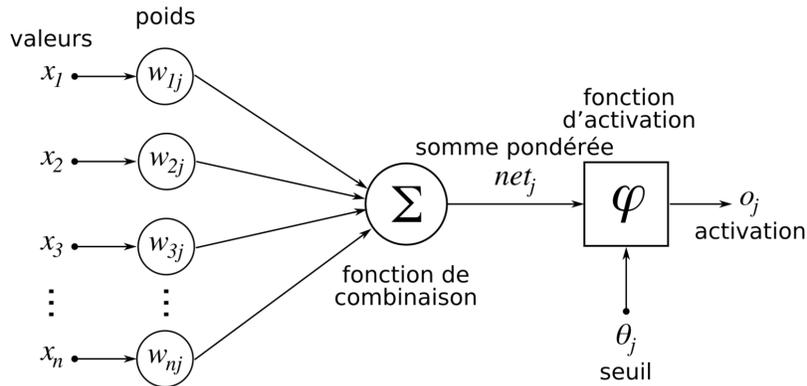


FIGURE 2.5 – Architecture des réseaux de neurones

[21]

### 2.4.2 L'architecture des réseaux de neurones profonds :

Dans l'apprentissage en profondeur, les réseaux de neurones profonds sont utilisés pour représenter et apprendre des hiérarchies de caractéristiques à partir de données. L'architecture des réseaux de neurones profonds se compose de plusieurs couches, dont les plus importantes sont les couches d'entrée, les couches cachées et la couche de sortie.

- **Les couches d'entrée** : La première couche reçoit des données d'entrée brutes telles que des images, du texte ou des signaux audio. Chaque nœud de la couche d'entrée représente une caractéristique spécifique des données.
- **Les couches cachées** : situées entre les couches d'entrée et de sortie, elles effectuent des calculs pour extraire des représentations de plus en plus abstraites des

données. Chaque couche cachée se compose de plusieurs neurones interconnectés qui calculent les poids et les biais pour prendre en compte les signaux d'entrée.

- **Les couche de sortie** : génère les prédictions ou les résultats finaux du modèle. La configuration de la couche de sortie dépend du type de problème à résoudre, comme la classification multiclasse ou la prédiction numérique. [22]

### 2.4.3 Les différents algorithmes de l'Apprentissage profond

Il existe plusieurs algorithmes de deep learning, chacun étant adapté à des tâches et des architectures de réseau de neurones spécifiques. Voici une liste de quelques-uns des algorithmes de deep learning les plus couramment utilisés avec leur brève définition :

- Réseau de neurones profond (**DNN**) : Il s'agit d'une architecture de réseau de neurones qui comprend plusieurs couches cachées entre la couche d'entrée et la couche de sortie. Les DNN se caractérisent par leur capacité à apprendre des représentations hiérarchiques et abstraites des données. Chaque couche du réseau traite les informations de la couche précédente en effectuant des calculs de poids et de biais associés aux connexions entre neurones [23].
- Réseaux de neurones convolutionnels (**CNN**) : un type particulier de réseau neuronal utilisé principalement pour le traitement d'images et la vision par ordinateur. Ils sont capables d'apprendre automatiquement à détecter et extraire les éléments visuels importants des images, tels que les bords, les textures et les formes. Les CNN utilisent des couches convolutionnelles qui appliquent des filtres pour analyser différentes parties de l'image. Cette capacité à apprendre et à hiérarchiser les fonctionnalités visuelles en fait une architecture puissante pour des tâches telles que la reconnaissance d'objets, la détection d'objets et la segmentation d'images [23]. la figure 2.6 représente un réseau neuronal convolutif.

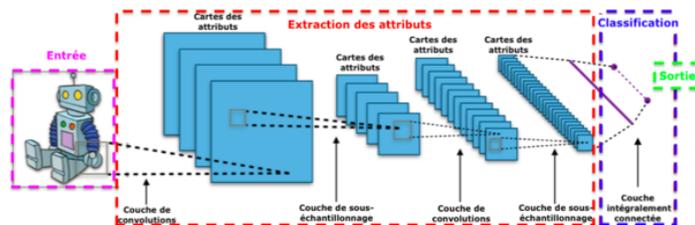


FIGURE 2.6 – Réseau neuronal convolutif

[24]

- Réseaux de neurones récurrents (**RNN**) : sont un type de réseau de neurones utilisé pour traiter des données séquentielles telles que du texte, de la parole ou des séries chronologiques. Contrairement aux réseaux de neurones classiques, les

RNN sont capables de se souvenir des informations du passé et de les utiliser pour prendre des décisions dans le futur.

Les **LSTM** (Long Short-Term Memory) sont une variante avancée des RNN qui ont été développées pour surmonter le problème de disparition du gradient qui se produit lorsque la formation des RNN sur de longues séquences devient difficile. Les LSTM utilisent des cellules de mémoire spéciales qui permettent de conserver des informations importantes pendant une longue période. Cela leur permet de capturer et de mémoriser les dépendances à long terme dans les séquences, améliorant ainsi les performances des RNN [23]. la figure 2.7 représente un réseaux de neurones récurrents et Les LSTM .

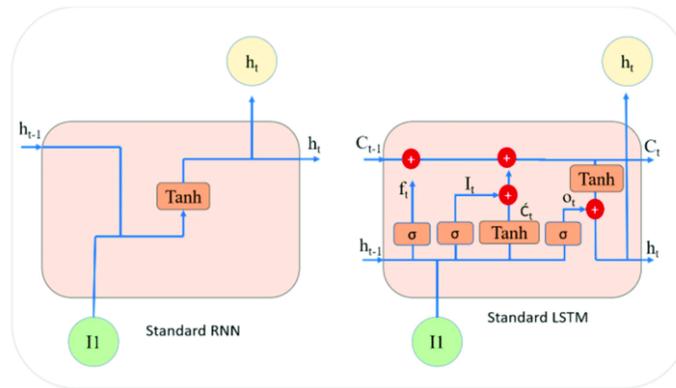


FIGURE 2.7 – Réseaux de neurones récurrents [25]

- Réseaux de neurones auto-encodeurs (**AE**) : une architecture de réseau de neurones utilisée pour apprendre une représentation compacte des données en réduisant la dimensionnalité de l'entrée et en essayant de la reconstruire à partir de cette représentation. Les AE sont souvent utilisés pour la compression de données et la détection d'anomalies [23]

## 2.5 Entraînement des réseaux de neurones

Le processus d'entraînement d'un réseau de neurones consiste à ajuster les poids et les biais du réseau afin qu'il puisse apprendre à effectuer une tâche spécifique. Voici les étapes clés du processus d'entraînement : [26]

**Initialisation des poids** : Les poids et les biais du réseau sont généralement initialisés de manière aléatoire. Cette étape permet de donner au réseau une capacité d'apprentissage initiale.

**Propagation avant (forward propagation)** : Lors de la propagation avant, les

données d'entrée sont propagées à travers le réseau couche par couche. Chaque couche effectue une combinaison linéaire des entrées pondérées par les poids, puis applique une fonction d'activation non linéaire pour produire une sortie. Cette propagation avant se poursuit jusqu'à ce que la sortie finale soit obtenue.

**Rétropropagation (backpropagation)** : Une fois que la sortie est obtenue, l'étape de rétropropagation est utilisée pour calculer le gradient de la fonction de perte par rapport aux poids du réseau. Cela se fait en utilisant l'algorithme de rétropropagation, qui propage l'erreur à travers le réseau de manière rétrograde, de la couche de sortie jusqu'à la couche d'entrée. Le gradient indique la direction et l'amplitude de la modification des poids nécessaires pour réduire la fonction de perte.

### 2.5.1 L'objectif de l'entraînement d'un réseau de neurones

L'objectif de l'entraînement d'un réseau de neurones est de minimiser la fonction de perte (loss function). La fonction de perte mesure à quel point les prédictions du réseau sont éloignées des valeurs réelles attendues pour un ensemble de données d'entraînement donné. En minimisant cette fonction de perte, le réseau de neurones cherche à ajuster ses poids de manière à obtenir des prédictions plus précises. L'entraînement du réseau se déroule en itérations appelées "epochs". Au cours de chaque epoch, les données d'entraînement sont propagées à travers le réseau (forward propagation) pour obtenir des prédictions, puis le gradient de la fonction de perte est calculé à l'aide de la rétropropagation. Les poids sont ensuite mis à jour en fonction de ce gradient, ce qui permet au réseau d'ajuster progressivement ses paramètres pour réduire la perte. Le processus d'entraînement se poursuit pendant plusieurs epochs jusqu'à ce que la fonction de perte converge vers un minimum, indiquant que le réseau a atteint un niveau d'apprentissage suffisant. Une fois l'entraînement terminé, le réseau est prêt à être utilisé pour faire des prédictions sur de nouvelles données. [27]

### 2.5.2 Les fonctions d'activation

Les fonctions d'activation sont utilisées dans les réseaux de neurones artificiels pour introduire la non-linéarité dans le modèle. Elles permettent de modéliser des relations complexes entre les entrées et les sorties des neurones, ce qui rend le réseau capable d'apprendre et de représenter des fonctions non linéaires. [26]

Voici quelques-unes des fonctions d'activation couramment utilisées dans les réseaux de neurones, ainsi que leur formule mathématique [26] :

- **Fonction Sigmoid (Sigmoid)** : La fonction sigmoïde est une fonction couramment utilisée dans les couches cachées des réseaux de neurones. Elle produit une sortie comprise entre 0 et 1, ce qui la rend adaptée pour modéliser des

probabilités ou des valeurs de décision binaires.

$$f(x) = \frac{1}{1 + \exp(-x)}$$

- **Fonction Tangente Hyperbolique (Tanh)** : La fonction tangente hyperbolique est similaire à la fonction sigmoïde, mais elle produit une sortie dans la plage  $[-1, 1]$ . Elle est souvent utilisée dans les couches cachées des réseaux de neurones.

$$f(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$$

- **Rectified Linear Unit (ReLU)** : La fonction ReLU est une fonction d'activation très populaire qui retourne la valeur d'entrée si elle est positive, sinon elle renvoie 0. Elle introduit une non-linéarité mais conserve les activations positives, ce qui la rend efficace pour résoudre le problème de disparition du gradient.

$$f(x) = \max(0, x)$$

- **Fonction Leaky ReLU** La fonction Leaky ReLU est une variante de ReLU qui introduit une pente légèrement négative pour les valeurs négatives. Cela permet de remédier à certains problèmes liés à la saturation des neurones dans la fonction ReLU.

$$f(x) = \max(0.01x, x)$$

- **Fonction Softmax** : La fonction Softmax est utilisée principalement dans la couche de sortie des réseaux de neurones pour la classification multiclass. Elle transforme les valeurs en scores normalisés qui représentent les probabilités relatives des différentes classes.

$$f(x_i) = \frac{\exp(x_i)}{\sum_{j=1}^N \exp(x_j)}$$

## 2.6 Systèmes de recommandation

Un système de recommandation est un type de système d'information qui utilise des techniques d'intelligence artificielle pour prédire les préférences et les intérêts d'un utilisateur en fonction des interactions passées avec le système ou de données similaires d'autres utilisateurs. Les systèmes de recommandation sont utilisés pour fournir aux

utilisateurs des suggestions personnalisées et pertinentes, qu'il s'agisse de produits, de services, de contenus ou d'autres éléments. Les algorithmes de recommandation peuvent utiliser diverses techniques d'analyse de données pour identifier les modèles et les similitudes des données utilisateur et produit pour générer des recommandations précises et utiles. [27]

### 2.6.1 Comment fonctionnent les systèmes de recommandation

Les systèmes de recommandation fonctionnent en analysant les données des utilisateurs pour identifier les tendances et les modèles de leur comportement afin de suggérer des éléments pertinents et adaptés à leurs préférences. Ces systèmes peuvent utiliser plusieurs techniques pour analyser les données, notamment [27] :

#### Recommandation basée sur la connaissance

un type de système de recommandation qui utilise des modèles de connaissances pour générer des recommandations. Ces systèmes s'appuient sur une base de connaissances qui décrit les caractéristiques des éléments à recommander et les préférences des utilisateurs. la figure 2.8

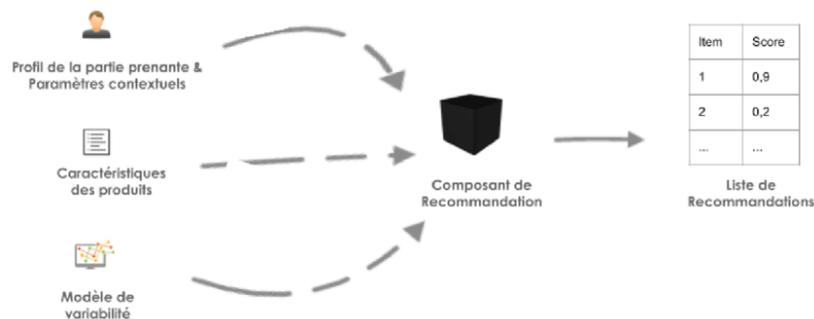


FIGURE 2.8 – Recommandation basées sur la connaissance

[23]

#### Recommandation basée sur le contenu

un type de système de recommandation qui utilise des informations sur les caractéristiques des éléments pour suggérer des éléments similaires à ceux qui ont été appréciés ou choisis par un utilisateur. Le système analyse les attributs de l'élément, tels que les mots clés, les catégories, les auteurs, les acteurs ou les réalisateurs, et utilise ces informations pour trouver des éléments qui ont des caractéristiques similaires. Figure 2.9

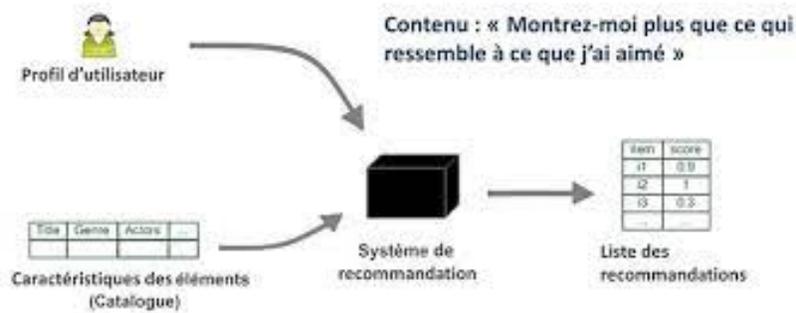


FIGURE 2.9 – Recommandation basées sur le contenu

[28]

### Recommandation basée sur le filtrage collaboratif

sont une méthode de recommandation qui utilise les données d'interaction de plusieurs utilisateurs pour suggérer des éléments ou des produits à un utilisateur en particulier. Ces systèmes se basent sur le principe que les préférences d'un utilisateur peuvent être déduites à partir des préférences similaires. la figure 2.10

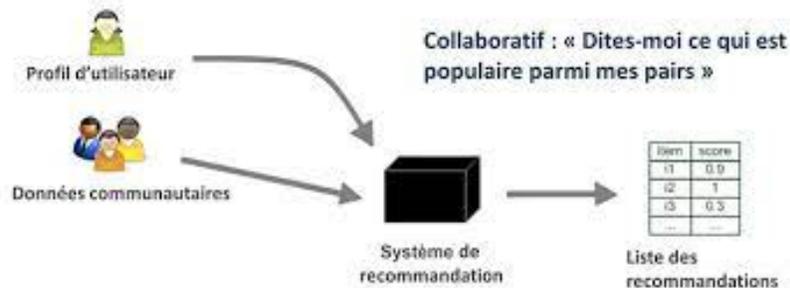


FIGURE 2.10 – Recommandation basées sur le filtrage collaboratif

[28]

### Recommandation Hybrid

un type de système de recommandation qui combine plusieurs approches pour améliorer la précision et la pertinence des recommandations. Ces systèmes combinent généralement des méthodes basées sur le contenu et des méthodes collaboratives pour tirer le meilleur des deux approches. la figure 2.11

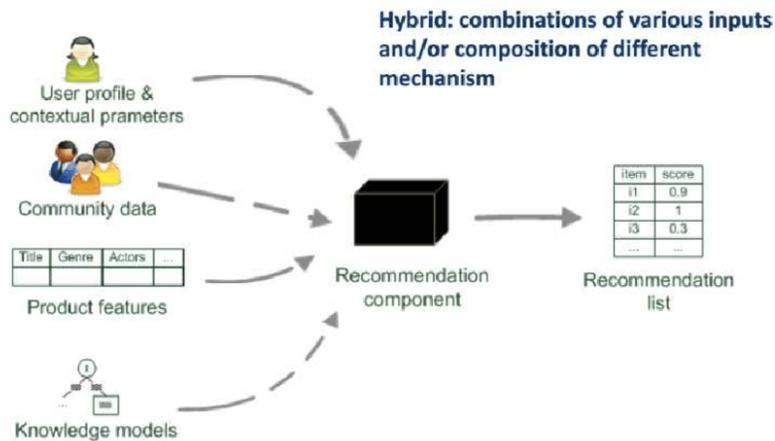


FIGURE 2.11 – Recommendation hybrid

28

## 2.7 Conclusion

nous avons défini l'intelligence artificielle (IA), l'apprentissage automatique (ML) et l'apprentissage profond (DL), L'intelligence artificielle vise à créer des systèmes qui effectuent des tâches de type humain. Le ML se concentre sur l'apprentissage à partir des données, tandis que le DL utilise des réseaux de neurones artificiels profonds pour comprendre des données complexes. Nous avons également couvert les outils de recommandation qui utilisent l'IA et le ML pour faciliter la prise de décision. Ces outils analysent les préférences et le comportement des utilisateurs et suggèrent des fonctionnalités pertinentes. Il existe deux principaux types de systèmes de recommandation : les systèmes basés sur le contenu, qui recommandent des éléments similaires à ceux déjà appréciés, et les systèmes collaboratifs, qui sont basés sur les préférences d'autres utilisateurs similaires.

# Chapitre 3

## Etat de l'art

### 3.1 Introduction

L'utilisation de la recommandation pour aider à la prise de décision est devenue de plus en plus courante dans notre monde numérique. Dans un monde où nous sommes constamment exposés à une quantité écrasante de données et d'options, Les systèmes de recommandation ont émergé comme une solution efficace pour aider les utilisateurs à découvrir des produits, des services, des informations ou des contenus pertinents, en fonction de leurs besoins et de leurs préférences. Diverses approches ont été développées pour créer des systèmes de recommandation, mais ces dernières années, les méthodes basées sur le deep learning ont connu un succès remarquable. Dans le cadre de ce chapitre, nous nous intéresserons aux approches d'apprentissage en profondeur (deep learning) appliquées à la recommandation. Notre objectif est de fournir une analyse comparative de ces différentes méthodes, permettant de comprendre leurs avantages et leurs limites respectives.

### 3.2 Etat de l'art

L'utilisation de l'apprentissage profond pour augmenter les performances des systèmes de recommandation et d'aide à la décision a suscité un vif intérêt dans les domaines de l'informatique et de la technologie. Dans Cette section nous présentons plusieurs travaux qui abordent le thème de la recommandation basée sur le Deep Learning

Dans l'article [29] Juan Carlos CepedaPacheco et Mari Carmen Domingo ont étudié les avantages de l'utilisation de l'apprentissage en profondeur pour développer un Système de recommandation IoT pour les attractions touristiques Smart City, le système de recommandation présenté utilise un réseau de neurones profonds (DNN) composé de plusieurs couches cachées avec des fonctions d'activation ReLU et binaire a été utilisée. Les résultats ont été évalués sur des données de voyageurs et ont révélé une

précision et un rappel de plus de 99% , ce qui démontre l'efficacité de cette approche pour recommander des attractions touristiques. Ces résultats sont encourageants car ils soulignent le potentiel de l'apprentissage en profondeur pour améliorer les systèmes de recommandation IoT et aider les utilisateurs à prendre des décisions éclairées. En outre, cette approche peut être appliquée à d'autres domaines, tels que la recommandation de produits ou de services, la recommandation de contenu en ligne, la recommandation de restaurants ou de bars, pour n'en nommer que quelquesuns. En résumé, cette étude illustre la puissance de l'apprentissage en profondeur pour améliorer les performances des systèmes de recommandation, et souligne le potentiel de cette approche pour développer des systèmes de recommandation IoT efficaces dans le domaine du tourisme et au-delà

L'étude présentée dans [30] par l'auteur Qusai Shambour expose une approche innovante pour améliorer les performances algorithmes de recommandation en concevant un algorithme de recommandation multicritères basé sur un auto-encodeur profond (AEMC). L'objectif de cet algorithme est de prédire les préférences de l'utilisateur pour divers éléments à l'aide d'une matrice brute d'évaluation des éléments utilisateur multicritères en entrée. Le fonctionnement de l'algorithme repose sur l'utilisation de couches cachées avec des fonctions d'activation sigmoïde, ainsi que sur l'optimiseur de réseau RMSprop. Des expériences ont été réalisées sur deux jeux de données multicritères de Yahoo! Films et TripAdvisor et les résultats ont été évalués pour leur précision et leur rappel. Les résultats expérimentaux ont montré que l'algorithme de recommandation proposé dans cette étude est extrêmement efficace pour faire des prédictions plus précises par rapport aux algorithmes de recommandation de pointe. Cette découverte est extrêmement prometteuse car elle montre que l'utilisation d'auto-encodeurs profonds peut améliorer considérablement les performances des systèmes de recommandation multicritères pouvant être utilisés dans différents domaines . En résumé, l'étude présenté dans [30] contribue de manière significative au domaine des systèmes de recommandation multicritères en proposant un algorithme innovant basé sur l'auto-encodeur. Les résultats prometteurs de cette étude montrent que cette approche est un moyen d'explorer l'amélioration des performances des systèmes de recommandation, ce qui pourrait avoir un impact significatif dans les industries où la recommandation est essentielle.

Dans cet article [31], les auteurs abordent la problématique des recommandations de films basées sur le filtrage et l'apprentissage en profondeur. Ils proposent une approche novatrice qui combine le filtrage collaboratif avec des techniques d'apprentissage en profondeur afin d'améliorer la précision des recommandations. Pour cela, les auteurs utilisent un modèle de réseau de neurones appelé AutoRec, spécialement conçu pour le filtrage collaboratif. Ils entraînent ce modèle en utilisant l'ensemble de données MovieLens, qui comprend des évaluations de films fournies par les utilisateurs ainsi que des informations sur les films eux-mêmes. Les résultats obtenus montrent que l'approche basée sur l'apprentissage en profondeur permet d'améliorer significativement la préci-

sion des recommandations par rapport aux méthodes traditionnelles telles que la factorisation matricielle. Les performances sont évaluées à l'aide de mesures couramment utilisées dans le domaine des systèmes de recommandation, telles que la précision et le rappel. En résumé, cet article présente une approche novatrice basée sur l'apprentissage en profondeur pour les recommandations de films collaboratives. Les auteurs utilisent le modèle AutoRec et l'ensemble de données MovieLens pour modéliser les relations complexes entre les utilisateurs et les films. Les résultats expérimentaux démontrent une amélioration significative de la précision des recommandations par rapport aux approches traditionnelles.

Cet article [32] aborde l'importance de la recommandation séquentielle pour améliorer les systèmes de recommandation, mais il souligne également les défis auxquels elle est confrontée, tels que le problème de démarrage à froid des éléments en raison du manque de commentaires des utilisateurs. Pour résoudre ce problème, une nouvelle approche est proposée, basée sur un réseau neuronal récurrent multi-vues (MV-RNN), qui utilise des informations visuelles et textuelles pour mieux comprendre l'intérêt de l'utilisateur. Cette approche a été testée sur deux ensembles de données réels, et les résultats ont montré que ce modèle a obtenu des performances de pointe en résolvant le problème des modalités manquantes grâce à la fusion de fonctionnalités visuelles et textuelles. Cependant, il convient de noter que cette approche peut être limitée dans certaines situations où ces informations ne sont pas disponibles. De plus, le modèle MV-RNN suppose que les utilisateurs ont des préférences stables et cohérentes au fil du temps, ce qui peut ne pas être le cas dans certains scénarios réels. Par conséquent, des recherches futures doivent être menées pour améliorer ce modèle en prenant en compte ces facteurs et pour mieux comprendre les habitudes de consommation des utilisateurs dans différents contextes.

Dans cet article [33], les auteurs présentent une étude approfondie sur un système de recommandation innovant basé sur un réseau de neurones récurrent. L'objectif est d'aider les étudiants à sélectionner les cours les mieux adaptés en fonction de leur niveau de connaissance et du domaine de développement qui les intéresse le plus. Les résultats de l'étude indiquent que le modèle de recommandation présente des performances modérées, évaluées à l'aide d'indicateurs tels que la précision et le score F, pour deux tâches de classification binaire. Les auteurs ont identifié plusieurs limites du modèle proposé, notamment la complexité de la tâche de prédiction, le manque de données d'entraînement et les problèmes d'interprétation des résultats. Malgré ces limites, cette approche représente une avancée importante dans le domaine des systèmes de recommandation pédagogique et peut apporter des avantages significatifs aux étudiants et aux enseignants en aidant à optimiser l'apprentissage et les résultats scolaires. Les recherches futures pourraient se concentrer sur l'amélioration des performances du modèle en utilisant des techniques d'apprentissage plus avancées et en intégrant des données supplémentaires pour améliorer la précision des prévisions.

Cet article [34] présente une méthode pour prédire les futurs achats des joueurs dans un jeu vidéo afin de leur recommander d'autres objets. Les auteurs utilisent un modèle prédictif basé sur des arbres extrêmement aléatoires (ERTs) et des réseaux de neurones profonds (DNNs). L'objectif est de recommander aux joueurs les objets qui leur conviennent le mieux en se basant sur leur comportement de jeu. Les auteurs utilisent un jeu de données provenant du jeu de cartes japonais Age Of Ishtaria. Ce jeu de données contient des informations sur les achats quotidiens de chaque joueur, sur une période allant de 24-09-2014 à 08-05-2017. Les données comprennent le nombre d'achats par objet et les ventes totales par objet pour chaque joueur. Pour représenter le comportement de jeu des joueurs sous forme de vecteurs statiques, les auteurs calculent des statistiques générales sur les données temporelles. Ces statistiques comprennent la moyenne, la variance, l'asymétrie, le kurtosis et le maximum. Ces caractéristiques sont utilisées pour représenter le comportement de jeu de chaque joueur. Étant donné que les joueurs effectuent généralement plusieurs achats, il peut y avoir plusieurs achats à prédire pour chaque joueur. Les auteurs résolvent ce problème en prenant un sous-échantillon de chaque série temporelle jusqu'à un certain moment, puis ils prédisent le prochain achat après ce moment. Ils utilisent ces échantillons pour former leur modèle. Le modèle proposé est une combinaison d'ERTs et de DNNs. Les ERTs sont utilisés pour leur efficacité computationnelle et leur capacité à réduire le surajustement, tandis que les DNNs sont utilisés pour leur capacité à apprendre des relations non linéaires. L'évaluation du modèle consiste à mesurer l'exactitude des prédictions sur une période temporelle. Les mesures d'évaluation incluent la vérification si l'objet prédit a été réellement acheté par le joueur lors de son prochain achat, s'il a été acheté à un moment donné dans la période temporelle, etc. Les résultats montrent que le modèle proposé est capable de prédire avec précision les futurs achats des joueurs. Les prédictions des modèles ERT et DNN sont similaires, avec de légères variations. Chaque joueur a des probabilités d'achat différentes pour chaque objet. Les avantages de cette approche sont sa faisabilité opérationnelle, sa facilité de parallélisation, sa scalabilité pour des millions d'utilisateurs et son potentiel d'augmentation des revenus grâce à une meilleure rétention des joueurs et à une augmentation des achats en jeu. Cependant, cette approche présente également des inconvénients, tels que la nécessité de connaissances spécifiques au jeu pour trouver les bonnes caractéristiques et la gestion de grandes quantités de données lors de l'entraînement du modèle.

Cet article [35] présente une application mobile appelée "Travigate" qui recommande des attractions touristiques, des restaurants et des hôtels en fonction des intérêts de l'utilisateur. Le système utilise l'algorithme de clustering K-modes pour regrouper les données d'entraînement provenant de l'ensemble de données de TripAdvisor. De plus, il utilise des réseaux de neurones convolutionnels (CNN) pour effectuer une recherche d'image inverse et reconnaître les lieux à partir d'images. La méthodologie du système de recommandation de lieux comprend le nettoyage et la préparation des données, l'utilisation du clustering K-modes pour recommander des lieux en fonction des intérêts similaires des utilisateurs, et l'utilisation des réseaux de neurones convolutionnels

(CNN) pour la reconnaissance d'image. Le jeu de données utilisé comprend des profils d'utilisateurs et des avis sur les lieux provenant de TripAdvisor, ainsi qu'un ensemble de données d'images obtenu en effectuant un web scraping d'images de Google. L'architecture du système de recommandation comprend une base de données MySQL, un serveur web et une application frontale.

Cet article [36] présente une approche appelée "RecDNNing" (Recommendation with Deep Neural Network embedding) pour la recommandation basée sur des systèmes de notation utilisateur-élément. L'approche vise à prédire les scores de notation en utilisant l'historique des utilisateurs et des éléments. Elle se compose de deux phases principales. Dans la première phase, les variables catégorielles sont transformées en vecteurs continus denses à l'aide de la technique Skip-gram, créant ainsi des intégrations d'utilisateurs et d'éléments. Ces intégrations sont ensuite moyennées et concaténées pour former des vecteurs d'entrée. Dans la deuxième phase, un réseau neuronal profond est utilisé pour prédire les scores de notation. Le réseau est composé d'une couche d'entrée, de couches cachées et d'une couche de sortie. La fonction d'activation ReLU est utilisée dans les couches cachées. Les vecteurs d'entrée sont propagés à travers le réseau, et la fonction SOFTMAX est appliquée à la couche de sortie pour obtenir les prédictions de notation. La différence entre les valeurs supervisées et prédites est évaluée à l'aide de l'entropie croisée. En résumé, RecDNNing utilise des intégrations d'utilisateurs et d'éléments moyennées et concaténées comme entrée pour un réseau neuronal profond, afin de prédire les scores de notation pour les utilisateurs actifs. L'approche se base sur des techniques d'intégration et d'apprentissage profond pour améliorer la précision des recommandations.

Dans cet article [37], les auteurs présentent un nouveau système de recommandation hybride basé sur l'apprentissage profond. L'objectif est de combler les lacunes des systèmes de filtrage collaboratif et d'atteindre une précision prédictive de pointe en utilisant des techniques d'apprentissage profond. Les systèmes de filtrage collaboratif sont largement utilisés et ont obtenu d'excellents résultats dans le domaine des systèmes de recommandation, mais ils rencontrent des problèmes de démarrage à froid lorsqu'il n'y a pas suffisamment d'historique sur les utilisateurs et les articles. De plus, les facteurs latents appris par ces méthodes sont linéaires par nature, ce qui limite leur capacité à capturer des relations non linéaires. Pour remédier à ces lacunes, les auteurs proposent un système de recommandation hybride qui utilise des embeddings pour représenter les utilisateurs et les articles, permettant ainsi d'apprendre des facteurs latents non linéaires. En intégrant des informations supplémentaires sur les utilisateurs et les articles dans un réseau neuronal profond, le système peut atténuer le problème du démarrage à froid. Les auteurs utilisent également une stratégie de taux d'apprentissage décroissant pour améliorer la précision des prédictions. Le système est évalué en termes de précision prédictive et de temps d'exécution sur plusieurs ensembles de données populaires. Les résultats montrent que la méthode proposée surpasse les approches existantes tant dans les cas de démarrage à froid que dans les cas sans démarrage à froid. En conclusion, ce

cadre de solution générique et performant peut être appliqué à d'autres ensembles de données de prédiction de notes dans le domaine des systèmes de recommandation.

nous clôturons cette section avec un tableau comparatif des articles étudiés, mettant en évidence les différences et les similitudes entre les approches proposées. Ce tableau nous permettra de mieux comprendre les performances de chaque méthode et de déterminer les meilleures pratiques pour la mise en œuvre de systèmes de recommandation basés sur le Deep learning.

Après avoir étudié tous ces articles, nous avons décidé d'utiliser les DNN (Deep Neural Networks) pour développer notre système de recommandation. Nous avons choisi cette approche car, par rapport aux autres méthodes traditionnelles de recommandation, les DNN offrent plusieurs avantages significatifs.

Tout d'abord, les DNN ont une capacité unique à apprendre des représentations de données hautement abstraites à partir des caractéristiques brutes. Cela signifie qu'ils sont capables d'extraire automatiquement des informations pertinentes et complexes à partir des données d'entrée, ce qui est essentiel pour comprendre les préférences des utilisateurs et générer des recommandations précises. Les DNN sont capables d'apprendre automatiquement ces caractéristiques, simplifiant ainsi considérablement le processus de développement du système de recommandation.

De plus, les DNN sont connus pour leur capacité à capturer les interactions non linéaires entre les différentes variables d'entrée. Dans le contexte des systèmes de recommandation, cela est particulièrement important car les préférences des utilisateurs peuvent être influencées par de multiples facteurs interdépendants et complexes. Les DNN peuvent modéliser ces relations complexes de manière plus efficace que les méthodes linéaires ou basées sur des modèles probabilistes.

En résumé, nous avons choisi de travailler avec les DNN pour notre approche de recommandation en raison de leur capacité à apprendre des représentations abstraites des données, leur aptitude à modéliser des interactions non linéaires et leur flexibilité pour s'adapter à différents types de recommandations.

### 3.3 Conclusion

L'utilisation de la recommandation basée sur l'apprentissage en profondeur (deep learning) est devenue courante dans notre monde numérique. Les systèmes de recommandation sont des outils puissants pour aider les utilisateurs à découvrir des produits, des services, des informations ou des contenus pertinents en fonction de leurs besoins et de leurs préférences. Différentes approches basées sur le deep learning ont été développées pour créer ces systèmes de recommandation, et elles ont montré des résultats remarquables dans divers domaines tels que le tourisme, le commerce électronique, la musique, l'éducation, etc. Les articles étudiés représentent différentes approches d'apprentissage en profondeur pour la recommandation, évaluant leurs performances et discutant de leurs avantages et limites. Ces approches comprennent l'utilisation de réseaux de neurones profonds, d'auto-encodeurs profonds, de réseaux de neurones récurrents et de combinaisons de modèles basés sur des arbres extrêmement aléatoires et des réseaux de neurones profonds. Les études présentées mettent en évidence l'efficacité de ces approches pour améliorer les performances des systèmes de recommandation dans différents domaines d'application et soulignent leur potentiel pour aider les utilisateurs à prendre des décisions éclairées.

Titre	Auteurs	Dataset	L'approche proposée	Résultats	Les limites de l'approche
Deep learning and Internet of Things for tourist attraction recommendation in smart cities (2022)	Juan Carlos CepedaPacheco , Mari Carmen Domingo	-La base de données utilisée est une base de données IoT qui regroupe des informations sur les attractions touristiques de Barcelone	-DNN : un réseau de neurones artificiels profonds, constitués de plusieurs couches de neurones interconnectés. Chaque couche traite les données d'entrée et produit des sorties qui sont utilisées comme entrées pour la couche suivante.	Les résultats montrent que plus le réseau de neurones est profond, meilleures sont les performances de l'algorithme. Comparé aux modèles traditionnels, le DNN obtient une précision et un rappel de plus de 99% .	-Les DNN nécessitent généralement de grandes quantités de données pour s'entraîner efficacement.
A Deep Learning Based Algorithm for Multi Criteria Recommender Systems (2020)	Qusai Shambour	Le jeu de données utilisé dans cette étude est composé du Yahoo! Movies MC dataset et du TripAdvisor MC dataset.	Les auteurs proposent l'algorithme AEMC, qui utilise des autoencodeurs profonds pour apprendre des représentations latentes des évaluations multicritères. L'algorithme est implémenté pour chaque critère, permettant ainsi de capturer les relations complexes entre les critères et les préférences des utilisateurs.	-Les résultats expérimentaux montrent que l'algorithme AEMC obtient une précision de prédiction plus élevée, mesurée par des métriques telles que l'erreur absolue moyenne (MAE) et l'erreur quadratique moyenne (RMSE)	la complexité de l'algorithme, qui peut nécessiter des ressources de calcul plus importantes que les méthodes de recommandation traditionnelles.
Movie Recommendations Using the Deep Learning Approach (2018).	Jeffrey Lund , Yiu-Kai Ng	Le jeu de données complet contient des évaluations pour 45 115 films provenant de 270 896 utilisateurs différents	l'article propose une approche novatrice qui combine le filtrage collaboratif avec des techniques d'apprentissage en profondeur en utilisant un modèle de réseau de neurones appelé AutoRec.	Les résultats obtenus montrent que l'approche basée sur l'apprentissage en profondeur améliore significativement la précision des recommandations par rapport aux méthodes traditionnelles telles que la factorisation matricielle	

Titre	Auteurs	Dataset	L'approche proposée	Résultats	Les limites de l'approche
MV-RNN : A Multi-View Recurrent Neural Network for Sequential Recommendation (2020)	Qiang Cui, Shu Wu, Qiang Liu, Wen Zhong, and Liang Wang	Dataset de Taobao, Dataset d'Amazon	Utiliser un réseau de neurones récurrents à vues multiples pour intégrer des informations multimodales dans la recommandation séquentielle et résoudre le problème du démarrage à froid des articles.	Le modèle MV-RNN surpasse les méthodes existantes en termes de précision de recommandation.	Difficulté à modéliser les changements dans les intérêts des utilisateurs au fil du temps.
Goal-based Course Recommendation (2019)	Weijie Jiang, Zachary A.Pardos, Qiang Wei	Les auteurs utilisent des données historiques sur les inscriptions, les notes et les majeures des étudiants provenant de l'Université de Californie, Berkeley.	Les auteurs proposent une approche basée sur les réseaux de neurones récurrents (RNN) pour recommander des cours en fonction des objectifs spécifiques des étudiants.	Les résultats montrent que leur modèle est capable de fournir des recommandations pertinentes et personnalisées.	Les résultats de l'approche dépendent de la qualité et de la pertinence des données historiques utilisées.
A Machine-Learning Item Recommendation System for Video Games (2018)	Paul Bertens, Anna Guitart, Pei Pei Chen and Africa Periañez	Le modèle utilise un jeu de données provenant du jeu de cartes Age Of Ishtaria	Les auteurs proposent une méthode basée sur des arbres extrêmement aléatoires (ERTs) et des réseaux de neurones profonds (DNNs) pour prédire les prochains achats des joueurs dans un jeu vidéo et leur recommander d'autres objets	Les résultats montrent que le modèle proposé est capable de prédire avec précision les prochains achats des joueurs.	cette approche présente des inconvénients tels que la nécessité de connaissances spécifiques au jeu pour trouver les bonnes caractéristiques et la gestion de grandes quantités de données lors de l'entraînement du modèle.

Titre	Auteurs	Dataset	L'approche proposée	Résultats	Les limite de l'approche
A Tourist Place Recommendation and Recognition System (2018)	Viken Parikh , Madhura Keskar , Dhwanil Dharia , Pradnya Gotmare	les données de TripAdvisor, comprenant environ 7000 utilisateurs et 32000 avis sur plus de 26000 lieux à travers le monde.	le modèle est entraîné à l'aide de l'algorithme de clustering K-modes et utilisant des CNN pour la reconnaissance d'image	le modèle de clustering K-modes est capable de classifier les utilisateurs en clusters avec une précision de 89,5%. Le modèle CNN entraîné atteint une précision moyenne de 46,4% pour la reconnaissance d'images des lieux	il repose sur des données provenant de TripAdvisor, ce qui signifie qu'il est limité aux lieux répertoriés sur cette plateforme.
RecDNNing : a recommender system using deep neural network with user and item embeddings (2019)	Hafed Zarzour , Ziad A. Al-Sharif , and Yaser Jararweh	le jeu de données MovieLens-100k, qui contient les évaluations de 943 utilisateurs sur 1682 vidéos, avec près de 100 000 scores de notation.	une approche appelée "RecDNNing" (Recommendation with Deep Neural Network embedding). RecDNNing vise à utiliser l'historique des utilisateurs et des éléments à partir d'une matrice de notation utilisateur-élément pour prédire les scores de notation.	Les résultats expérimentaux montrent que l'approche RecDNNing surpasse à la fois les méthodes SVD et KNN en termes de RMSE sur le jeu de données MovieLens-100k	Les réseaux de neurones nécessitent généralement une grande quantité de données.
DNNRec : A Novel Deep Learning based Hybrid Recommender System (2019)	Kiran R, Pradeep Kumar, Bharat Bhasker	MovieLens 100K, FilmTrust, Book-Crossing et MovieLens 1M	Le modèle utilisé dans cette proposition est un système de recommandation hybride basé sur l'apprentissage en profondeur, qui combine des techniques de filtrage collaboratif avec des réseaux de neurones profonds	Les résultats montrent que la technique proposée surpasse les méthodes existantes tant dans les cas de démarrage à froid que dans les cas non-cold start	Complexité computationnelle : Les réseaux de neurones profonds sont connus pour leur complexité computationnelle élevée, ce qui peut rendre l'entraînement et l'inférence du modèle relativement lents

TABLE 3.1 – Étude comparative des travaux connexes 1

# Chapitre 4

## Approche proposée SR-DNN

### 4.1 Introduction

Dans ce chapitre, nous présenterons en détail notre approche proposée qui vise à aider les voyageurs à prendre des décisions de voyage en leur fournissant de bonnes recommandations de séjours. Notre approche utilise un réseau neuronal profond (DNN) qui exploite les avantages du deep learning pour fournir des recommandations plus précises et personnalisées.

### 4.2 Approche proposée SR-DNN

Notre approche appelée **SR-DNN** (système de recommandation basée sur Deep Neural Network), vise à exploiter la puissance des réseaux de neurones profonds pour la recommandation de contenu personnalisé. SR-DNN vise à capturer des modèles complexes et latents à partir des données des utilisateurs et des articles afin de fournir des recommandations précises et pertinentes. Le principal avantage de SR-DNN est sa capacité à apprendre des représentations de haut niveau à partir données d'entrée. À l'aide d'une architecture de réseau neuronal profond, les caractéristiques complexes et les relations entre les utilisateurs et les éléments peuvent être automatiquement extraites

#### 4.2.1 Les étapes de SR-DNN

Dans cette section, nous allons vous présenter les différentes étapes de notre approche SR-DNN (Système de Recommandation basée sur Deep Neural Network). Les étapes clés de cette approche sont illustrées dans la figure 4.1 .

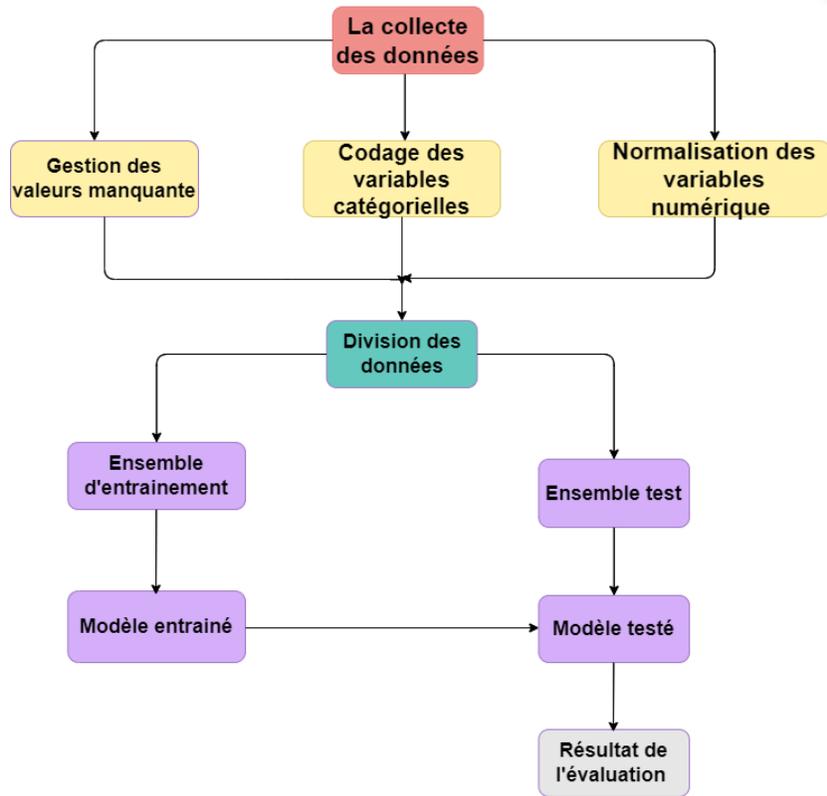


FIGURE 4.1 – Les étapes de SR-DNN

## La Collecte des données

Nous avons généré un dataset qui représentent les préférences de voyage et les caractéristiques d’hébergement des utilisateurs. L’ensemble de données est créé de manière aléatoire à l’aide d’une méthode de génération pseudo-aléatoire.

Nous avons décidé de créer un ensemble de données car nous n’avons pas accès aux ensembles de données existants qui répondaient à nos besoins . En générant des données, nous sommes en mesure de collecter les caractéristiques et les attributs des utilisateurs tels que le nom, l’âge, le budget, la durée du voyage, les intérêts de voyage, le type d’hébergement, la destination et l’étoile.

Il convient de noter que la génération aléatoire de l’ensemble de données nous permet de créer un ensemble diversifié de profils et de préférences d’utilisateurs , les ensembles de données nous permettent de créer un environnement contrôlé pour former et évaluer notre approche SR-DNN. Pour collecter ces données, nous avons chargé des données à partir d’un fichier Excel en utilisant la fonction `read_excel` de la bibliothèque `pandas`. Nous avons spécifié le chemin du fichier Excel dans la fonction `read_excel` pour charger les données à partir du fichier spécifié. Les donnée utiliser pour SR-DNN sont présenter dans la figure 4.2

	nom	age	Budget	durée du voyage	Intérêts Voyage	Type Hébergement	destination	etoile
0	Emma Smith	57	1376	25	plage	The Waldorf Astoria	Costa Rica	1
1	Olivia Johnson	38	644	14	ville	Pullman Hotels and Resorts	Inde	1
2	Ava Williams	39	664	28	ville	The Breakers	Kenya	1
3	Isabella Jones	27	1322	8	montagne	St. Regis Hotels & Resorts	Mexique	1
4	Sophia Brown	25	1355	17	montagne	Kimpton Hotels & Restaurants	Émirats arabes unis	2
...	...	...	...	...	...	...	...	...
10192	Isabella Adams	42	440	17	randonnée	The Ritz-Carlton, Chicago	Italie	5
10193	Lucas Foster	49	2839	21	montagne	Protea Hotels by Marriott	Mexique	2
10194	Stella Diaz	64	829	28	ville	Grand Hyatt	Espagne	1
10195	Emily Russell	38	780	18	randonnée	The Burj Khalifa	Grèce	5
10196	Gabriel Reed	40	992	9	montagne	COMO Hotels and Resorts	Autriche	4

10197 rows x 8 columns

FIGURE 4.2 – Le dataset

## Prétraitement des données

Dans cette partie, nous allons présenter les différentes étapes de prétraitement des données utilisées dans notre approche SR-DNN

**Gestion des valeurs manquantes :** Nous avons utilisé la méthode `dropna()` pour supprimer les lignes contenant des valeurs manquantes de notre dataset , En éliminant ces lignes , nous nous assurons que les données utilisées pour notre approche sont complètes et fiables.

**Codage des variables catégorielles :** Nous avons identifié des colonnes de catégorie dans notre ensemble de données, telles que le Nom, l'Intérêt de voyage, le Type d'hébergement et la Destination. Afin de traiter ces variables catégorielles, nous avons converti les valeurs catégorielles en nombres entiers à l'aide d'un `LabelEncoder` de la bibliothèque `scikit-learn`.

**Normalisation des variables numériques :** Nous avons identifié des colonnes numériques dans notre ensemble de données, telles que l'âge, le budget et la durée du voyage. Afin de préparer ces variables numériques pour l'analyse, nous avons utilisé le `StandardScaler` de la bibliothèque `scikit-learn` pour normaliser les valeurs de chaque colonne numérique. Dans la figure ci-dessous, nous présentons le dataset après avoir été normalisé pour les valeurs numériques et encodé pour les variables catégorielles. la figure 4.3 représente Le dataset après la normalisation et l'encodage

	nom	age	Budget	durée du voyage	Intérêts Voyage	Type Hébergement	destination	etoile
0	1038	0.951298	-0.346732	1.014228	1	83	12	1
1	2799	-0.345390	-1.281045	-0.450694	3	38	20	1
2	466	-0.277143	-1.255517	1.413752	3	54	26	1
3	1656	-1.096103	-0.415657	-1.249742	0	47	29	1
4	3529	-1.232597	-0.373536	-0.051170	0	25	49	2
...	...	...	...	...	...	...	...	...
10192	1649	-0.072403	-1.541427	-0.051170	2	73	24	5
10193	2159	0.405324	1.520617	0.481529	0	37	29	2
10194	3564	1.429024	-1.044914	1.413752	3	17	15	1
10195	1020	-0.345390	-1.107457	0.082005	2	56	18	5
10196	1255	-0.208896	-0.836863	-1.116567	0	10	4	4

10197 rows × 8 columns

FIGURE 4.3 – Le dataset après la normalisation et l’encodage

**Division des données :** Nous avons divisé notre dataset en deux parties distinctes : les entités (x) et les cibles (y). Les entités (X) comprennent toutes les colonnes dataset, à l’exception de la colonne "etoile". La colonne "etoile" est considérée comme la variable cible (y) que nous cherchons à prédire.

**Conversion les données en un tableau numpy :** Nous avons converti les données d’entité (x) et de cible (y) en tableaux numpy en utilisant la méthode `to_numpy()`. Cette opération nous permet d’obtenir une représentation numérique des données qui est compatible avec notre approche SR-DNN.

## L’architecture de SR-DNN

l’architecture proposé représente un modèle d’algorithme DNN (Deep Neural Network), également connu sous le nom de réseau de neurones profond. Un DNN est un type de modèle d’apprentissage profond qui est composé de plusieurs couches de neurones artificiels, organisées de manière séquentielle. Dans notre approche SR-DNN, nous avons utilisé trois couches denses, également connues sous le nom de couches entièrement connectées. Dans ce qui suit, nous vous présenterons en détail l’architecture de notre approche SR-DNN

**Couche dense (64 neurones, activation 'relu') :** Cette couche est la première couche de SR-DNN. Elle est définie avec 64 neurones, ce qui signifie qu’elle aura 64 sorties différentes. La fonction d’activation utilisée est **ReLU** (Rectified Linear Unit), qui est couramment utilisée dans les réseaux de neurones pour introduire une non-linéarité. Le paramètre `input_dim` est spécifié pour indiquer la dimension d’entrée de

cette couche. Dans notre code, `input_dim` est défini comme `num_features`, qui correspond au nombre total de caractéristiques du dataset.

**Couche dense (32 neurones, activation 'relu')** : Cette couche est la deuxième couche du SR-DNN elle est définie avec 32 neurones, ce qui signifie qu'elle aura 32 sorties différentes. La fonction d'activation utilisée est également **ReLU**.

**Couche dense (1 neurone, activation 'linear')** : Cette couche est la dernière couche du SR-DNN elle est définie avec 1 neurone, ce qui signifie qu'elle produira une seule sortie. La fonction d'activation utilisée est 'linear'.

En résumé, l'architecture du modèle comporte deux couches intermédiaires de 64 et 32 neurones respectivement, avec des fonctions d'activation 'relu' pour introduire de la non-linéarité. La dernière couche a 1 neurone avec une fonction d'activation 'linear', ce qui produit la sortie continue pour la prédiction de la variable cible 'etoile'.

La figure 4.4 représente l'architecture de notre approche SR-DNN

Model: "mon\_modele\_de\_recommandation"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 64)	576
dense_1 (Dense)	(None, 32)	2080
dense_2 (Dense)	(None, 1)	33

=====  
Total params: 2,689  
Trainable params: 2,689  
Non-trainable params: 0

FIGURE 4.4 – L'architecture de RS-DNN

## Entraînement et Evaluation de SR-DNN

**Compilation du RS-DNN** : Le RS-DNN est compilé en spécifiant la fonction de perte (loss function) et l'optimiseur à utiliser. Dans notre cas, la fonction de perte est définie comme 'mean\_squared\_error' (erreur quadratique moyenne) et l'optimiseur choisi est 'adam', qui est un optimiseur populaire et efficace pour l'apprentissage en profondeur.

### Entraînement du RS-DNN :

L'entraînement du RS-DNN est effectué à l'aide de la méthode fit du modèle. Nous spécifions les données d'entraînement `X_train` et les étiquettes d'entraînement `y_train` comme entrées. Le nombre d'épochs est défini à 10, ce qui signifie que notre SR-DNN passera 10 fois à travers l'ensemble des données d'entraînement lors du processus d'entraînement.

Le paramètre `batch_size` est défini à 64, ce qui indique que 64 échantillons d'entraînement sont utilisés à chaque itération pour mettre à jour les poids de notre approche.

Pour évaluer les performances de SR-DNN pendant l'entraînement, nous utilisons les données de validation `X_test` et les étiquettes de validation `y_test`. Cela nous permet de surveiller la perte (loss) et la précision (accuracy). Cela nous donne une indication de la capacité de notre approche à généraliser et à faire des prédictions précises sur de nouvelles données.

La figure 4.5 ci-dessous présente des informations qui donnent un aperçu des performances du modèle à la dernière époque d'entraînement. Ces informations incluent la valeur de la perte (loss) et la précision (accuracy).

```
Epoch 10/10  
128/128 [=====] - 1s 4ms/step - loss: 0.0166 - accuracy: 0.2039
```

FIGURE 4.5 – L'entraînement du RS-DNN

**Prédiction des données de test :** Une fois l'entraînement terminé, nous utilisons SR-DNN pour effectuer des prédictions sur les données de test `X_test`. Ces prédictions sont obtenues en appelant la méthode `predict` du modèle et les résultats sont stockés dans la variable `y_pred`.

**Évaluation de SR-DNN :** Les performances de SR-DNN sont évaluées en comparant les prédictions `y_pred` aux véritables étiquettes de test `y_test`. Différentes mesures d'erreur sont calculées pour évaluer la qualité des prédictions et évaluer les performances de notre approche. Dans cette partie, nous présentons les métriques que nous utilisons pour évaluer les performances de notre approche :

La Mean Squared Error (**MSE**), ou erreur quadratique moyenne, est une mesure couramment utilisée pour évaluer la performance d'un modèle. Elle calcule la moyenne des carrés des écarts entre les prédictions du modèle et les valeurs réelles correspondantes. La formule mathématique de la MSE est la suivante [38] :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

où  $n$  est le nombre total d'échantillons,  $y_i$  est la valeur réelle de l'échantillon  $i$ , et  $\hat{y}_i$  est la valeur prédite de l'échantillon  $i$ .

La Mean Absolute Error(), ou erreur absolue moyenne, est une autre mesure couramment utilisée pour évaluer la performance d'un modèle . Contrairement à la MSE qui calcule les écarts au carré, la **MAE** calcule la moyenne des valeurs absolues des écarts entre les prédictions du modèle et les valeurs réelles correspondantes[38].

La formule mathématique de la MAE est la suivante :

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Root Mean Squared Error (**RMSE**) : il s'agit de la racine carrée de l'erreur quadratique moyenne, qui mesure l'erreur moyenne entre les prédictions et les vraies valeurs en prenant en compte la direction de l'erreur.[38]

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

où n est le nombre d'échantillons ou d'observations, y\_i est la valeur réelle de la variable cible pour le i-ème échantillon, y\_ est la valeur prédite de la variable cible pour le i-ème échantillon, et sum représente la somme des valeurs pour tous les échantillons.

R-squared (R<sup>2</sup>) : il s'agit d'une mesure de l'ajustement du modèle aux données , qui indique la proportion de la variance totale expliquée par le modèle. La valeur du R<sup>2</sup> varie de 0 à 1. Voici sa formule mathématique :[38]

$$R^2 = 1 - \frac{SSE}{SST}$$

où SSE (Sum of Squared Errors) représente la somme des carrés des écarts entre les valeurs prédites et les valeurs réelles, et SST (Total Sum of Squares) représente la somme des carrés des écarts entre les valeurs réelles et la moyenne des valeurs réelles.

En résumé, le modèle est entraîné pendant 10 epochs avec une taille de batch de 64. Les prédictions sont ensuite calculées sur les données de test et les performances du modèle sont évaluées à l'aide de différentes mesures d'erreur. Dans notre exemple, le modèle semble obtenir de bons résultats avec un faible MSE, MAE, RMSE et un R<sup>2</sup> élevé, ce qui suggère une bonne adéquation aux données de test.

## Validation de SR-DNN

Dans cette section, nous présentons les résultats obtenus par notre approche SR-DNN . Ces résultats sont essentiels pour évaluer la performance du modèle de recommandation. Lors de l'évaluation du modèle, les mesures suivantes ont été obtenues : une

loss (erreur) de 0.0240 et une accuracy (précision) de 0.2096. Étant donné que l'accuracy est relativement faible, différentes mesures d'erreur ont été utilisées pour évaluer la qualité des prédictions de notre approche.

En ce qui concerne les métriques d'erreur, la Mean Squared Error (MSE) est de 0.02, ce qui indique que les prédictions du modèle ont une erreur quadratique moyenne de 2%. Une MSE plus faible est généralement préférée, car cela signifie que le modèle est capable de prédire les valeurs cibles avec moins d'erreur.

De plus, la Mean Absolute Error (MAE) est de 0.09, ce qui signifie que les prédictions du modèle ont une erreur moyenne absolue de 9%. Une MAE plus faible est généralement préférée, car cela indique que le modèle parvient à prédire les valeurs cibles avec une plus grande précision.

Le Root Mean Squared Error (RMSE) est de 0.13, ce qui correspond à une erreur quadratique moyenne des prédictions du modèle. Une RMSE plus faible est préférée. Les résultats obtenus sont présentés dans le graphe dans la figure suivante 4.6

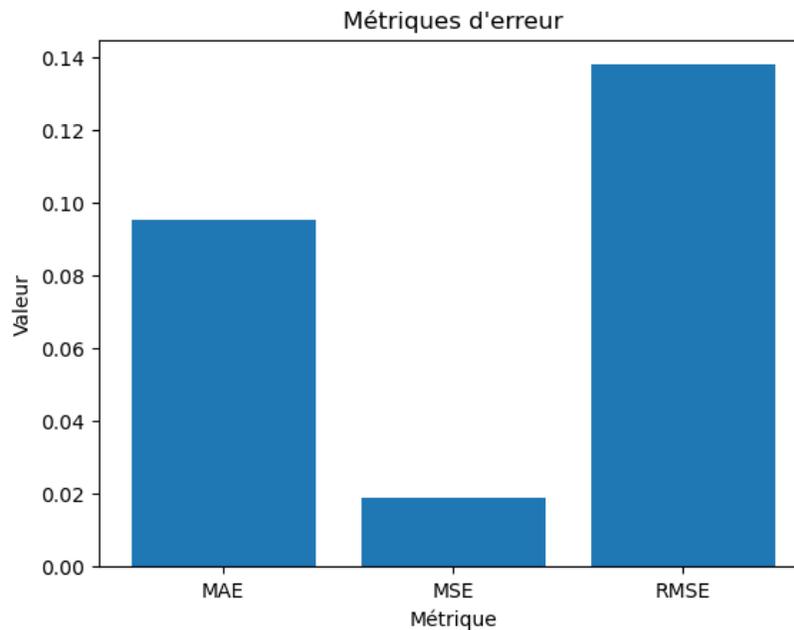


FIGURE 4.6 – Évaluation des métriques d'erreur

Une autre métrique utilisée est le coefficient de détermination  $R^2$  qui mesure la proportion de la variance de la variable cible. Un  $R^2$  de 0.99 indique que SR-DNN explique 99% de la variance de la variable cible. Un  $R^2$  élevé est généralement considéré comme un bon indicateur de performance, car il indique que notre approche parvient à capturer une grande partie de la variation de la variable cible.

La figure 4.7 représente un graphe de comparaison entre les valeurs réelles et les valeurs prédites, avec le coefficient de détermination. L'axe X représente les valeurs réelles de l'étoile, allant de 1 à 5, et chaque groupe de points dans le graphique représente un échantillon individuel. L'axe Y représente les valeurs prédites par notre modèle.

Une ligne diagonale en rouge est tracée pour illustrer l'idéal où les valeurs prédites seraient parfaitement égales aux valeurs réelles. Les points bleus sur le graphique indiquent dans quelle mesure les valeurs prédites s'éloignent ou se rapprochent des valeurs réelles de l'étoile. L'objectif de ce graphique est d'avoir les points bleus situés près de la ligne rouge, ce qui indiquerait que les valeurs prédites sont proches des valeurs réelles de l'étoile.

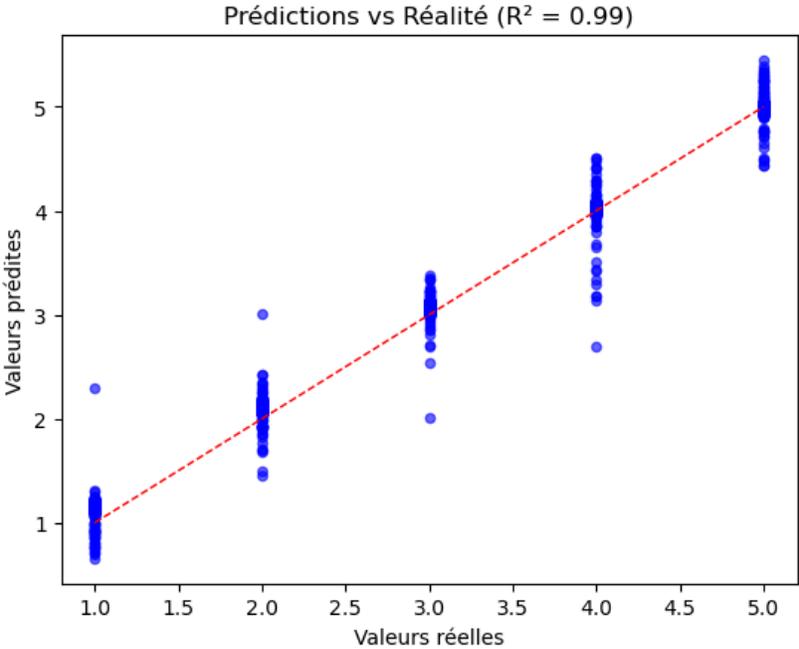


FIGURE 4.7 – Comparaison des valeurs réelles 'étoile' et prédites avec coefficient de détermination

# Chapitre 5

## Implémentation

### 5.1 Introduction

Dans ce chapitre, nous avons validé notre approche en l'appliquant au domaine de la recommandation d'un séjour touristique basé sur les informations fournies par l'utilisateur. Pour mieux comprendre le fonctionnement de notre approche SR-DNN, nous avons décrit en détail les différentes étapes à suivre.

### 5.2 Etude de cas

Pour valider notre approche, nous l'avons appliquée au domaine de la recommandation d'un séjour touristique en se basant sur les informations fournies par l'utilisateur. Afin de mieux comprendre le fonctionnement du programme, voici une explication détaillée des différentes étapes à suivre :

Tout d'abord, les noms de destinations et les types d'hébergement sont récupérés à partir du DataFrame appelé "train\_data". Ces informations sont ensuite stockées dans des dictionnaires afin de créer une correspondance entre les codes d'encodage et les noms réels des destinations et des types d'hébergement.

Ensuite, l'utilisateur est invité à fournir certaines informations essentielles, telles que son âge, son budget, son intérêt de voyage et la durée prévue du voyage. la figure représente les informations fournies par l'utilisateur 4.8 tel que l'âge le budget la durée de voyage et l'intérêt de voyage

---

```
Quel est votre âge ? 23
Quel est votre budget ? 2500
Quel est votre intérêt de voyage ? plage
Combien de jours durera votre voyage ? 22
```

FIGURE 5.1 – Les informations fournies par l'utilisateur

Un nouveau DataFrame, nommé "new\_data", est créé pour contenir les informations fournies par l'utilisateur, ainsi que des colonnes supplémentaires pour les variables catégorielles telles que "Type Hébergement" et "destination". le nouveau DataFrame est représenté dans la figure 4.9

	nom	age	Budget	durée du voyage	Intérêts Voyage	Type Hébergement	destination	etoile
0	utilisateur	23	2500	22	plage			0

FIGURE 5.2 – Un nouveau DataFrame

Les variables catégorielles sont encodées à l'aide de l'encodeur de labels (LabelEncoder) de la bibliothèque scikit-learn. Cela permet de convertir les valeurs catégorielles en valeurs numériques afin de les utiliser dans le modèle de prédiction. Les variables numériques sont normalisées à l'aide d'un objet StandardScaler de la bibliothèque scikit-learn. La normalisation est effectuée pour mettre toutes les variables numériques à la même échelle.

Une fois les données préparées, le modèle est utilisé pour faire une prédiction en utilisant les informations fournies par l'utilisateur.

Les noms des destinations recommandées et les types d'hébergement correspondants sont extraits des dictionnaires créés précédemment, afin d'obtenir les noms réels associés aux identifiants. La Figure 4.10 présente les résultats recommandés, qui incluent à la fois une destination et un type d'hébergement, tous deux encodés.

	destination	Type Hébergement
0	17	66

FIGURE 5.3 – Une destinations et un types d'hébergement encoder

Enfin, les recommandations sont affichées en utilisant les noms de destinations et les types d'hébergement correspondant aux identifiants, pour donner à l'utilisateur une recommandation de voyage. les résultats d'exécution de notre approche

```
# Afficher les recommandations avec les noms de destinations et de types d'hébergement correspondant aux identifiants
for index, row in recommandations.head(5).iterrows():
    destination = row['destination']
    type_hebergement = row['Type Hébergement']
    print("Pour votre séjour, je vous recommande de visiter {}, et de séjourner à {}".format(destination, type_hebergement))
```

Pour votre séjour, je vous recommande de visiter France, et de séjourner à Mövenpick Hotels & Resorts

FIGURE 5.4 – Résultats d'exécution de notre approche

En résumé, notre approche SR-DNN combine des informations fournies par l'utilisateur avec un modèle de prédiction pour recommander des destinations de voyage et des types d'hébergement adaptés à ses préférences.

### 5.3 Conclusion

Bien que le score de précision de SR-DNN sur les données de validation (`val_accuracy`) soit faible, ce qui suggère que le modèle pourrait avoir des difficultés à généraliser ses prédictions sur de nouvelles données, les résultats obtenus indiquent que notre approche SR-DNN a réussi à minimiser la perte lors de l'entraînement et présente des valeurs d'erreur (MSE, MAE, RMSE) relativement faibles. De plus, le coefficient de détermination  $R^2$  est élevé, ce qui indique que le modèle explique une grande partie de la variance de la variable cible. Cependant, pour garantir la performance globale de SR-DNN, des améliorations supplémentaires pourraient être nécessaires, notamment en augmentant la précision sur les nouvelles données.

# Conclusion

Dans notre étude, nous avons souligné l'importance de l'intelligence artificielle (IA) et de l'apprentissage en profondeur pour améliorer la prise de décision grâce aux systèmes de recommandation. Les systèmes de recommandation basés sur le deep learning, tels que notre approche SR-DNN, offrent des avantages significatifs en termes de précision et de pertinence des recommandations. Ces approches permettent d'optimiser les résultats en proposant des recommandations personnalisées qui prennent en compte les préférences de chaque utilisateur. Cela peut conduire à une meilleure satisfaction de l'utilisateur et à une amélioration de l'expérience. Malgré un score de précision (accuracy) relativement faible de 20% sur les données de validation, les résultats obtenus dans notre approche proposée, nommée SR-DNN, ont démontré une performance satisfaisante avec des erreurs quadratique moyenne (MSE) et moyenne absolue (MAE) de seulement 2% et 9%, respectivement, ainsi qu'un root mean squared error (RMSE) de 13%. Ces faibles valeurs d'erreur indiquent que le modèle est capable de minimiser l'erreur de prédiction. De plus, un  $R^2$  de 0.99 indique que SR-DNN explique 99% de la variance de la variable cible. Il est important de noter que, bien que notre approche ait montré une performance satisfaisante, il est essentiel d'améliorer l'accuracy pour garantir une meilleure précision et une meilleure généralisation des prédictions sur de nouvelles données. Des améliorations supplémentaires pourraient inclure l'ajout de nouvelles fonctionnalités, l'exploration de techniques d'optimisation avancées, ou encore l'augmentation de la taille et de la diversité des données d'entraînement. Ces améliorations pourraient contribuer à renforcer l'efficacité globale de notre approche en améliorant la précision et la fiabilité des prédictions.

# Références

- [1] Université de Sherbrooke. Brochure de prise de décision. (2021). Consulté le 23 mai 2023. Disponible sur : [https://www.usherbrooke.ca/etudiants/fileadmin/sites/etudiants/documents/Psychologie/Brochure\\_prise\\_de\\_decision\\_2021.pdf](https://www.usherbrooke.ca/etudiants/fileadmin/sites/etudiants/documents/Psychologie/Brochure_prise_de_decision_2021.pdf)
- [2] Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY : Farrar, Straus and Giroux.
- [3] Dobelli, R. (2015). *The Art of Thinking Clearly - Edisi Bahasa Melayu*. books.google.com.
- [4] Anterior prefrontal function and the limits of human decision-making E Koechlin, A Hyafil - *Science*, 2007 - science.org
- [5] Hammond, J. S., Keeney, R. L., Raiffa, H. (1998). *Smart Choices : A Practical Guide to Making Better Decisions*.
- [6] Briganti, G. (2023). Intelligence artificielle : une introduction pour les cliniciens. *Revue des Maladies Respiratoires*, Elsevier.
- [7] Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, 10(2), 1-37. DOI : 10.2478/jagi-2019-0002 <https://sciendo.com/downloadpdf/journals/jagi/10/2/article-p1.pdf>
- [8] <https://dataanalyticspost.com/quelques-pistes-developper-intelligence-artificielle/>
- [9] Janiesch, C., Zschech, P., Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets, Volume(31)*, 685–695. DOI : <https://doi.org/10.1007/s12525-021-00475-2>
- [10] Osisanwo, F.Y., Akinsola, J.E.T., Awodele, O., Hinmikaiye, J.O., Olakanmi, O., Akinjobi, J. (2017). Supervised Machine Learning Algorithms : Classification and Comparison. *International Journal of Computer Trends and Technology (IJCTT)*, Volume(48).
- [11] Mehrbakhsh Nilashi, Othman bin Ibrahim, Hossein Ahmadi, Leila Shahmoradi (2017). An Analytical Method for Diseases Prediction Using Machine Learning Techniques *Computers and Chemical Engineering*, (17), 30257-0. <http://dx.doi.org/doi:10.1016/j.compchemeng.2017.06.011>
- [12] H. Laude. *Prise de décision dans des environnements complexes. Management & Datascience*, 2017. Disponible sur : [management-datascience.org](http://management-datascience.org).

- [13] Stéphane Chrétien et Julien Velcin. *Interprétabilité en machine learning et interaction homme/machine*. Laboratoire ERIC - Université Lyon 2.
- [14] Gregor Leban, Blaž Zupan, Gaj Vidmar, Ivan Bratko. *VizRank : Data Visualization Guided by Machine Learning*. *Data Mining and Knowledge Discovery*, volume 13, pages 119-136, 2006.
- [15] Batta Mahesh. *Machine Learning Algorithms - A Review*. *International Journal of Science and Research (IJSR)*, Volume 9, Issue 1, January 2020. ISSN : 2319-7064. ResearchGate Impact Factor (2018) : 0.28 | SJIF (2018) : 7.426. Licensed Under Creative Commons Attribution CC BY. URL : <http://www.ijsr.net>.
- [16] <https://www.chegg.com/flashcards/cours-10-1-intelligence-artificielle-de-fb63b450-0e-deck>
- [17] <https://www.idbc.fr/tutoriel-modele-de-regression-lineaire-comment-evaluer-les-hy>
- [18] <https://towardsdatascience.com/efficient-k-means-clustering-algorithm-with-optimu>
- [19] Yann LeCun, Yoshua Bengio, Geoffrey Hinton. *Deep learning*. *Nature*, doi :10.1038/nature14539.
- [20] Miiikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Hodjat, B. *Evolving Deep Neural Networks*. *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, pages 293–312. doi :10.1016/b978-0-12-815480-9.00015-3.
- [21] Batta Mahesh. *Machine Learning Algorithms - A Review*. *International Journal of Science and Research (IJSR)*, volume 9, issue 1, January 2020. ISSN : 2319-7064. ResearchGate Impact Factor (2018) : 0.28 | SJIF (2018) : 7.426.
- [22] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. *A survey of deep neural network architectures and their applications*. *Neurocomputing*, 234, 11–26. doi :10.1016/j.neucom.2016.12.038.
- [23] Shrestha, A., & Mahmood, A. *Review of Deep Learning Algorithms and Architectures*. *IEEE Access*, 7, 53040–53065. doi :10.1109/access.2019.2912200.
- [24] <https://towardsdatascience.com/convolutional-neural-networks-e5a6745b2810>
- [25] [https://www.researchgate.net/figure/Standard-architecture-of-RNN-and-LSTM\\_fig4\\_339700136](https://www.researchgate.net/figure/Standard-architecture-of-RNN-and-LSTM_fig4_339700136)
- [26] **Université A. Mira de Béjaia**  
 Faculté de Sciences Exactes  
 Département d'Informatique  
 Apprentissage Avancé **Enseignant** : BOUCHEBBAH Fatah
- [27] Lü, L., Medo, M., Yeung, C. H., Zhang, Y.-C., Zhang, Z.-K., & Zhou, T. (2012). *Recommender systems*. *Physics Reports*, 519(1), 1–49. doi :10.1016/j.physrep.2012.02.006.
- [28] DR. Houda El Bouhissi. *Recommendation systems : Applications, Challenges and Issues*. LIMED Laboratory, University of Bejaia.

- [29] Cepeda-Pacheco, J.C., Domingo, M.C. (2022). Deep learning and Internet of Things for tourist attraction recommendations in smart cities. *Neural Computing and Applications*, 34(2022), 7691–7709. doi : 10.1007/s00521-021-06872-0.
- [30] Shambour, Q. (2020). A deep learning based algorithm for multi-criteria recommender systems. *Knowledge-Based Systems, to appear*. doi : 0.1016/j.knosys.2020.106545.
- [31] Lund, J., & Ng, Y.-K. (2018). Movie Recommendations Using the Deep Learning Approach. 2018 IEEE International Conference on Information Reuse and Integration (IRI). doi : 10.1109/iri.2018.00015.
- [32] Cui, Q., Wu, S., Liu, Q., Zhong, W., Wang, L. (2020). MV-RNN : A Multi-View Recurrent Neural Network for Sequential Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 32(2), 317.
- [33] Jiang, W., Pardos, Z. A., Wei, Q. (2019). Goal-based Course Recommendation. In *Proceedings of the 9th International Conference on Learning Analytics Knowledge - LAK19*. doi :10.1145/3303772.3303814
- [34] Bertens, P., Guitart, A., Chen, P. P., & Perianez, A. (2018). A Machine-Learning Item Recommendation System for Video Games. 2018 IEEE Conference on Computational Intelligence and Games (CIG). doi : 10.1109/cig.2018.8490456.
- [35] Parikh, V., Keskar, M., Dharia, D., & Gotmare, P. (2018). A Tourist Place Recommendation and Recognition System. 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). doi : 10.1109/icicct.2018.8473077.
- [36] Zarzour, H., Al-Sharif, Z. A., & Jararweh, Y. (2019). RecDNNing : a recommender system using deep neural network with user and item embeddings. 2019 10th International Conference on Information and Communication Systems (ICICS). doi : 10.1109/iacs.2019.8809156.
- [37] R, K., Kumar, P., & Bhasker, B. (2019). DNNRec : A Novel Deep Learning based Hybrid Recommender System. *Expert Systems with Applications*, 113054. doi : 10.1016/j.eswa.2019.113054.
- [38] <https://openclassrooms.com/fr/courses/4297211-evaluez-les-performances-dun-modele-de-machine-learning/4308276-evaluez-un-algorithme-de-regression>

## *Résumé*

---

Ce projet de fin d'études se concentre sur l'utilisation d'outils d'intelligence artificielle pour aider les utilisateurs à prendre des décisions . Notre approche repose sur l'utilisation de la recommandation comme moyen d'assistance. Pour cela, nous avons développé une nouvelle approche SR-DNN(système de recommandation basé sur l'apprentissage profond.)

L'objectif principal de notre approche était d'améliorer les recommandations fournies aux utilisateurs afin de les guider vers des décisions pertinentes. Pour illustrer son fonctionnement, nous avons choisi de l'appliquer spécifiquement dans le domaine du tourisme. Ainsi, notre système permet d'effectuer des choix concernant un séjour touristique en se basant sur les données fournies par l'utilisateur.

Grâce à l'utilisation des informations saisies par l'utilisateur, notre approche est en mesure de suggérer une destination et un type d'hébergement correspondant à ses préférences. En combinant les techniques de l'apprentissage profond et de la recommandation, nous avons pu proposer des recommandations personnalisées et pertinentes.

## *Abstract*

---

This final year project focuses on the utilization of artificial intelligence tools to assist users in making informed decisions. Our approach relies on recommendation as a means of assistance. To achieve this, we developed a novel approach called SR-DNN (Deep Learning-based Recommendation System).

The primary objective of our approach was to enhance the recommendations provided to users and guide them towards relevant decisions. To demonstrate its functionality, we specifically applied our system in the tourism domain. Consequently, our system facilitates decision-making regarding a tourist trip based on user-provided data.

By leveraging the user-inputted information, our approach is capable of suggesting a destination and accommodation type that align with their preferences. Through the combination of deep learning techniques and recommendation systems, we were able to provide personalized and meaningful recommendations.