

République Algérienne Démocratique et Populaire
Ministère de L'enseignement Supérieur et de la Recherche Scientifique
École Normale Supérieure des Professeurs de l'enseignement Technologique

- Skikda -

Département de Mathématique et informatique



Mémoire du projet de fin d'études

Spécialité : Professeur d'Enseignement Moyen en Informatique

THEME :

**DÉTECTION DES ERREURS DANS UNE RÉCITATION
CORANIQUE EN UTILISANT LE DEEP LEARNING**

Réalisé par :

Chenta Nour El yakine

Abed Sara

Supervisé par :

Mme. Bekhouche Safia

Devant le jury composé de :

Président : Boudaoued faicel

Examineurs : Sellami samir, Abdoun leila

Année universitaire : 2024/2025

Remerciements

Louange à Allah, Seigneur des mondes, et que la paix et les bénédictions soient sur Son noble Prophète Muhammad(ﷺ) ainsi que sur sa famille et ses compagnons.

Nous souhaitons exprimer nos plus sincères remerciements à notre encadrant(e) 'Bekhouche Safia' pour sa précieuse orientation, sa disponibilité constante et la qualité de son accompagnement tout au long de la réalisation de ce mémoire. Son expertise, sa rigueur scientifique et ses conseils avisés ont été d'une grande valeur pour l'aboutissement de ce travail.

Nos remerciements s'adressent également à l'ensemble du corps enseignant du département Mathématiques et Informatique, pour l'excellence de leur enseignement et leur engagement tout au long de notre formation.

Nous tenons à remercier nos camarades et collègues pour les échanges constructifs, l'entraide et le climat de coopération qui ont enrichi cette expérience.

Nous adressons enfin notre gratitude la plus profonde à nos familles respectives, qui n'ont cessé de nous soutenir par leurs prières, leur patience, leur affection et leurs encouragements inconditionnels, tout au long de ce parcours.

À toutes celles et ceux qui, de près ou de loin, ont contribué à la réussite de ce travail, nous témoignons ici notre reconnaissance la plus distinguée.

Dédicace

À ma chère mère,

Toi qui es la lumière de mon chemin, le refuge de mes incertitudes et la source de ma force,

Je te dédie ce travail avec une gratitude infinie.

Tes prières silencieuses, ton amour inconditionnel et tes sacrifices discrets Ont été le socle sur lequel j'ai bâti chaque étape de ce parcours.

À mon père bien-aimé,

Exemple de sagesse, de rigueur et de dignité, Je te remercie pour ton soutien constant, ta confiance en moi et ton regard toujours encourageant.

Que Dieu te récompense pour tout ce que tu as semé dans mon cœur.

À mes frères et sœurs,

Pour votre affection sincère, vos encouragements et votre présence bienveillante à chaque moment important,

Merci de m'avoir entouré(e) d'un amour fraternel qui donne sens à chaque réussite.

À mes honorables enseignants,

Je vous adresse mon profond respect et ma sincère reconnaissance. Par votre dévouement, votre exigence intellectuelle et votre passion pour la transmission du savoir, Vous avez façonné mon esprit et nourri mon ambition.

Ce mémoire est en partie le fruit de votre accompagnement éclairé.

À tous ceux qui, de près ou de loin, ont cru en moi, m'ont soutenu(e), inspiré(e) ou simplement tendu la main,

Je vous dédie cette modeste réalisation comme un hommage à la patience, à l'effort et à

l'espoir partagé.

Dédicace

C'est avec une immense joie et une profonde gratitude que je dédie l'aboutissement de ce travail, un jalon important de ma vie, aux piliers de mon existence.

À mon très cher père, Abdelkader,

Tu as toujours incarné pour moi l'exemple même de la droiture, du respect et de l'intégrité.

Grâce à toi, papa, j'ai appris la valeur inestimable du travail de la responsabilité

Ce modeste travail est le fruit de tous les sacrifices que tu as déployés pour mon éducation et ma formation. J'implore le tout-puissant pour qu'il t'accorde une bonne santé et une vie longue et heureuse

À ma très chère mère, Amel,

Aucun mot, aucune action, ne sauraient exprimer la plénitude de ma gratitude envers toi. Ton affection enveloppante, ta bienveillance ta présence constante à mes côtés ont toujours été la source intarissable de ma force pour surmonter les défis et franchir les obstacles. Tu es mon étoile guide, ma confidente.

À mes merveilleux frère et sœur, Zineb, Merci pour votre patience inconditionnelle durant mes moments de doute et de difficulté

À ma chère cousine Hana, Ta présence et ton soutien indéfectibles ont été une véritable bénédiction. Merci d'être toujours là, avec ta gentillesse

À mes chères amies, Hasna, Imen et mon binôme Nour, Vous êtes bien plus que des amies ; vous êtes des sœurs sur qui je sais pouvoir compter, des complices dans les rires comme dans les défis.

Enfin, je souhaite adresser mes sincères remerciements à toutes celles et ceux, sans exception, qui m'ont soutenue et accompagnée, de près ou de loin, dans cette étape de ma vie.

Résumé

Le Saint Coran constitue la source fondamentale de la législation islamique, mais sa mémorisation et sa récitation correcte selon les règles du Tajwīd représentent un défi pour de

nombreux apprenants, en particulier les non-arabophones. Ce mémoire présente une approche innovante fondée sur l'apprentissage profond, visant à assister les apprenants dans la récitation du Coran. Un modèle de réseau de neurones convolutifs (CNN) a été développé et entraîné à partir de la base de données "quran-reciters", permettant de reconnaître les versets récités et de détecter les erreurs de prononciation. Les résultats obtenus, avec une précision de 95 %, confirment l'efficacité du système. Ce travail propose une solution technologique respectueuse des exigences spirituelles et pédagogiques de l'apprentissage du Coran, en combinant rigueur scientifique, traitement vocal et intelligence artificielle.

Mots clés : Apprentissage profond, CNN, Coran, Tajwīd, Réseau de neurones, MFCC, Reconnaissance vocale, Verset.

Abstract

The Holy Quran is the foundational source of Islamic law, yet memorizing and accurately reciting it according to Tajwid rules remains a major challenge, especially for non-Arabic speakers. This research proposes an innovative deep learning-based approach to assist learners in Quranic recitation. A Convolutional Neural Network (CNN) model was developed and trained using the "quran-reciters" dataset, enabling the identification of recited verses and detection of pronunciation errors. The model achieved an accuracy of 95% on the test set, demonstrating its effectiveness. This work presents a technological solution that combines scientific rigor, voice signal processing, and artificial intelligence while respecting the sacred aspects of Quranic learning.

Keywords: Deep Learning, CNN, Quran, Tajwīd, Neural Network, MFCC, Voice Recognition, Verse.

ملخص

يمثل القرآن الكريم المصدر الأساسي للتشريع الإسلامي، إلا أن حفظه وتلاوته بشكل صحيح وفق قواعد التجويد يُعدّ تحديًا كبيرًا للعديد من المتعلمين، خاصةً من غير الناطقين باللغة العربية. يقدم هذا البحث مقاربة مبتكرة تعتمد على التعلم العميق، وتهدف إلى مساعدة المتعلمين في تلاوة القرآن الكريم. تم تطوير نموذج لشبكة عصبية التلافيفية (CNN) وتدريبه اعتمادًا على قاعدة البيانات "quran-reciters"، مما مكّنه من التعرف على الآيات المُرتّلة واكتشاف أخطاء النطق. وقد أكدت النتائج المحققة، بدقة بلغت 95%، فعالية هذا النظام. يقدم هذا العمل حلًا تكنولوجيًا يحترم المتطلبات الروحية والتربوية لتعلم القرآن الكريم، من خلال الدمج بين الصرامة العلمية، ومعالجة الصوت، والذكاء الاصطناعي.

Table Des Matières

Introduction générale.....	1
Chapitre 1 : Intelligence Artificielle & Deep Learning.....	4
1.1 Introduction.....	5
1.2 Définition de l'intelligence Artificielle.....	5
1.3 Histoire et Évolution de l'IA.....	6
1.3.1 Canard de Vaucanson 1738	6
1.3.2 Gestation de l'IA (1943-1955)	6
1.3.3 Le test de TURING (1950)	6
1.3.4 Naissance d'IA (1956).....	6
1.3.5 Systèmes Experts (1969-1979).....	7
1.3.6 Le retour des réseaux de neurones	7
1.4 Les Branches de l'intelligence artificielle.....	7
1.4.1 Apprentissage Automatique (Machine Learning)	7
1.4.2 Réseaux de Neurones Artificiels (Neural Networks).....	7
1.4.3 Vision par Ordinateur (Computer Vision.....	7
1.4.4 Traitement du Langage Naturel (Natural Language Processing – NLP)	8
1.4.5 Robotique.....	8
1.5 Système multi agent.....	8
1.6 Réseaux de Neurones Artificiels.....	7
1.7 Deep Learning.....	9
1.7.1 Domaines d'application de l'apprentissage profond.....	9
1.7.2 Principe et fondamentaux de l'apprentissage profond (DL)	10
1.7.3 Positionnement du Deep Learning dans l'IA	11
1.8 Les principales architectures de Deep Learning.....	12
1.8.1 Réseau de neurones convolutifs (CNN)	12
1.8.2 Les réseaux de neurones récurrents (RNN)	12
1.8.3Réseau antagoniste génératif (GAN)	12
1.8.4Transformer.....	13
1.9 Les techniques d'apprentissage.....	13

1.9.1 Apprentissage supervisé	13
1.9.2 Apprentissage non supervisé	13
1.9.3 L'apprentissage semi-supervisé	14
1.10 Les algorithmes d'apprentissage supervisé.....	14
1.10.1 Arbre de décision	14
1.10.2 Naive Bayes.....	14
1.10.3 Machine à vecteurs de support	14
1.11 Les algorithmes d'apprentissage non supervisé	15
1.11.1 K-Means.....	15
1.11.2 FC-Means.....	15
1.12 Conclusion	16
Chapitre 2 : La science de la récitation du Coran.....	17
2.1 Introduction.....	18
2.2 Définition du Coran.....	18
2.3 Définition et règles du Tajwid.....	18
2.4 Le développement de la science de Tajwid.....	19
2.5 Sections du Tajwid et Les jugement de chaque elle.....	19
2.6 Les niveaux de récitation.....	20
2.7 Les fautes de récitation (Al-Lahn).....	20
2.8 Le Waqf (L'arrêt dans la récitation).....	20
2.8.1 L'arrêt contraint (الوقف الاضطراري).....	21
2.8.2 L'arrêt évaluatif (الوقف التعليمي).....	21
2.8.3 L'arrêt d'attente (الوقف الانتظاري).....	21
2.8.4 L'arrêt optionnel (الوقف الاختياري).....	21
2.9 L'Initiation.....	22
2.9.1 Le début véritable.....	22
2.9.2 Le début additionnel.....	22
2.9.3 Le début absolu.....	22
2.9.4 Le début approprié.....	22

2.10 Dessin coranique	23
2.10.1 Définition du dessin coranique.....	23
2.10.2 Avantages du dessin coranique.....	23
2.11. Définir et Comparer les travaux existants	25
2.12 Conclusion.....	25
Chapitre3 : Application de Deep Learning à la reconnaissance vocale du Coran.....	27
3.1 Introduction.....	28
3.2 La reconnaissance automatique de la parole.....	28
3.3 La Reconnaissance Vocale.....	28
3.3.1 Acquisition du signal vocal (Signal de parole).....	29
3.3.2 Prétraitement du signal.....	29
3.3.3 La modélisation acoustique.....	29
3.3.4 Décodage phonétique.....	30
3.4 Les différentes approches de Deep Learning pour la reconnaissance vocale.....	30
3.4.1 Réseaux de Neurones Récurrents (RNN).....	30
3.4.2 Réseaux de Neurones Convolutifs (CNN).....	31
3.4.3 Modèles de langage neuronal (Les Transformers).....	32
3.5 La reconnaissance vocale de la récitation coranique.....	33
3.6 Choix de la solution.....	36
3.7 Conclusion.....	37
Chapitre 4 : Conception du système	38
4.1 Introduction.....	39
4.2 Diagramme de cas d'utilisation	39
4.3 Diagramme de Séquence (Sequence Diagram)	41
4.4 Diagramme de Classe (Class Diagram).....	43
4.5 Diagramme d'activité.....	44
4.6 Conclusion	45
Chapitre 5 : Expérimentation & Résultats.....	46
5.1 Introduction	47

5.2 Outils et bibliothèques utilisés pour le développement.....	47
5.2.1 Langages et Bibliothèques.....	47
5.2.2 Environnement de développement.....	48
5.2.3 Les outils utilisés pour le développement.....	48
5.3 Méthodologie d'entraînement	49
5.3.1Collecte les données..	49
5.3.2 Prétraitement des données audio	49
5.3.3 Architecture du Modèle.....	51
5.3.4 L'entraînement.....	47
5.4 Métriques d'évaluation.....	52
5.5 Résultats comparatifs.	53
5.6 Discussions.....	54
5.7 Architecture de l'Interface.....	55
5.7.1 Composants Clés.....	55
5.7.2 Fonctionnalités Démontrées.....	55
5.7.3 Intégration du Modèle.....	56
5.8 Conclusion.....	58
Conclusion générale	59

Liste Des Abréviations

CNN : Convolutionnels Neural Network " Réseau de Neurones Convolutifs"

DL : Deep Learning « Apprentissage profond »

CAN : Conversion Analogique Numérique.

CONV: Convolutional Layer.

GAN: Generative Adversarial Network. "Réseaux Antagonistes Génératifs".

GRU : Gated Recurrent Unit"Unité Récurrente Fermée".

IA : Intelligence Artificielle.

KNN: K-Nearest Neighbors «K plus proches voisins"

ML : Machine Learning "Apprentissage automatique"

MFCC : Mel Frequency Cepstral Coefficients "Coefficients Cepstraux en Fréquence de Mel"

RNN : Recurrent Neural Network " Réseau de Neurones Récurrent"

RAP : Reconnaissance Automatique de la Parole

ReLU : Rectified Linear Unit " Unité Linéaire Rectifiée"

SVM : Support Vector Machine "Machine à Vecteurs de Support"

SMA: Système Multi-Agents.

RBM: Restricted Boltzmann Machine

LSTM: Long Short-Term Memory.

Liste Des Figures

Figure 1 : Fonctionnement du neurone artificiel -Deep Learning.

Figure 2 : Un travail typique d'apprentissage à distance pour résoudre des problèmes du monde réel.

Figure 3 : Relation entre IA, DL, ML.

Figure 4 : SVM.

Figure 5 : Le processus de la reconnaissance vocale.

Figure 6 : Processus de numération.

Figure 7 : Présentation d'un système ASR.

Figure 8 : Illustration de la carte des caractéristiques 2D-CNN pour la reconnaissance vocale .

Figure 9 : Comparaison de l'efficacité des techniques de reconnaissance vocale du coran.

Figure 10 : Résumé de la recherche sur la récitation du Saint Coran.

Figure 11 : La description modulaire du système

Figure 12 : Les différentes bibliothèques utilisées dans notre implémentation

Figure13 : MFCC Features

Figure 14 : La table de matrice de confusion

Figure15 : Résultat 1.

Figure 16 : Résultat 2.

Figure17 : Résultat 3.

Figure 18 : Interface d'accueil.

Figure 19 : Entrée de l'application.

Figure 20 : Interface de test.

Figure 21 : Interface de résultat.

Introduction Générale

Le Saint Coran, parole divine révélée au Prophète Mohammed (ص), constitue le fondement spirituel, juridique et culturel de la civilisation islamique. Allah a dit : *إِنَّا نَحْنُ نَزَّلْنَا الذِّكْرَ وَإِنَّا لَهُ* «لُحُوظُونَ» (Sourate Al-Hijr, 15 :9). Ce texte sacré, préservé dans sa forme originale, est récité selon des règles strictes de Tajwīd, garantissant une transmission fidèle à travers les siècles.

Les différentes lectures du Coran, telles que Hafs, Warsh, Qaloun, reflètent la richesse de sa transmission orale. Ces variations, authentifiées par le Prophète (ص), soulignent l'importance de la récitation précise. Comme le rapporte Uthmān ibn 'Affān (رضي الله عنه), le Messenger d'Allah (ص) a dit : « Le meilleur d'entre vous est celui qui apprend le Coran et l'enseigne » (Bukhari). Ainsi, la mémorisation (Hifz) et la récitation correcte (Tilāwa) sont des actes d'adoration majeurs, mais elles représentent un défi, en particulier pour les non-arabophones.

Dans ce contexte, les technologies modernes, notamment le Deep Learning, offrent des solutions innovantes pour faciliter l'apprentissage du Coran. Contrairement aux méthodes traditionnelles de traitement du signal ou de reconnaissance vocale basique, les réseaux de neurones profonds, comme les CNN (Convolutionnels Neural Networks), permettent une analyse fine des nuances phonétiques et mélodiques propres à la récitation coranique. Leur capacité à traiter de larges jeux de données et à s'adapter aux différentes Qira'at en fait un outil privilégié pour la validation automatique de la prononciation et des règles de Tajwīd.

Ce mémoire propose donc d'explorer l'application du Deep Learning dans l'assistance à la récitation coranique, combinant précision technologique et préservation du patrimoine religieux, tout en respectant les spécificités des différentes lectures.

Contexte de travail

L'apprentissage traditionnel du Coran repose sur l'enseignement direct par des professeurs qualifiés (*Shuyūkh*), ce qui peut présenter des limitations en termes d'accessibilité et de disponibilité. Parallèlement, on observe :

- Une demande croissante d'outils numériques d'apprentissage du Coran
- Des avancées majeures en reconnaissance vocale grâce au *Deep Learning*

Ce projet s'inscrit dans cette dynamique en proposant une solution technologique qui :

- ✓ Permet une évaluation objective de la récitation
- ✓ Offre un retour instantané aux apprenants
- ✓ Respecte les règles strictes du *Tajwīd*

Problématique

Bien que plusieurs applications existent aujourd'hui pour l'apprentissage du Coran, la validation automatique et précise de la récitation soulève plusieurs défis :

Premièrement, la complexité phonétique et linguistique de l'arabe coranique exige une modélisation fine des subtilités articulatoires. Deuxièmement, le respect strict des règles du *Tajwīd* impose un seuil de précision exceptionnel, où la moindre erreur de prononciation ou d'intonation compromet la validité religieuse de la récitation. Troisièmement, la variabilité des voix, des accents et des niveaux de maîtrise parmi les apprenants nécessite une robustesse algorithmique accrue. Ces contraintes soulèvent trois questions de recherche fondamentales :

Comment l'apprentissage profond peut-il contribuer à améliorer la mémorisation et la récitation précise du Coran malgré sa complexité linguistique ? Quelle est la manière de concevoir un système qui puisse vérifier de manière précise et fiable une récitation Coranique? Enfin, comment garantir que le système respecte intégralement les règles sacrées de récitation?.

Ce mémoire vise à apporter des réponses méthodologiques à ces interrogations tout en démontrant expérimentalement la faisabilité d'un système automatisé, précis et fiable.

Organisation du mémoire

Ce travail est structuré en cinq chapitres. Le premier présente les fondements théoriques de l'intelligence artificielle et de la reconnaissance vocale, en exposant les concepts clés tels que le Machine Learning, le Deep Learning et les principales architectures de réseaux de neurones.

Le deuxième chapitre est consacré à la science de la récitation coranique (*Tajwīd*), avec une attention particulière portée aux règles de prononciation et à l'étude du Rasm (orthographe coranique). Le troisième chapitre explore l'apport du Deep Learning à la reconnaissance vocale coranique, notamment à travers l'adaptation des réseaux convolutifs (CNN) à l'analyse de la récitation, et une revue critique des travaux existants dans ce domaine. Le quatrième chapitre décrit l'architecture du système développé, ainsi que les diagrammes UML consacrés à la conception de notre application.

Enfin, le cinquième chapitre présente les modalités d'implémentation, l'analyse des performances et la validation des résultats, avant de conclure ce mémoire par une synthèse générale et des perspectives d'amélioration future du système proposé.

Chapitre 1 :
Intelligence Artificielle &
Deep Learning

1.1 Introduction

Il n'est pas facile de cerner cette activité en constante évolution et même de définir exactement l'Intelligence artificielle (IA). L'objectif des premiers chercheurs était réellement de concevoir des machines intelligentes, capables d'imiter l'intelligence humaine dans ses tâches intellectuelles les plus complexes, l'IA englobe une variété de techniques et de technologies, dont le machine Learning (apprentissage automatique) et le Deep Learning (apprentissage profond), qui permettent aux machines d'apprendre à partir de données et de s'améliorer avec le temps.

L'IA est utilisée dans de nombreux domaines, tels que :

- **La reconnaissance vocale et faciale** : Les assistants virtuels comme Alexa utilisent l'IA pour comprendre et répondre aux commandes vocales.
- **Les véhicules autonomes** : Les voitures sans conducteur utilisent des algorithmes d'IA pour naviguer et prendre des décisions en temps réel.
- **Les finances** : Les systèmes d'IA sont utilisés pour détecter les fraudes, analyser les marchés financiers et gérer les portefeuilles d'investissement.

Avec les avancées récentes en IA, telles que l'apprentissage profond (DeepLearning), l'IA est devenue un domaine d'innovation qui continue de transformer de nombreuses industries et de susciter des discussions sur ses implications éthiques et sociales.

L'IA a le potentiel de transformer de nombreux aspects de notre vie quotidienne, mais elle soulève également des questions on éthiques et sociales importantes, telles que la confidentialité des données, l'impact sur l'emploi et la prise de décision automatisée.

1.2 Définition de l'intelligence Artificielle

Depuis ses débuts en tant que domaine d'étude au milieu du XXe siècle, l'intelligence artificielle (IA) représente une frontière constamment repoussée. [1]

L'Intelligence Artificielle est une branche de l'informatique dont le but est de réaliser des systèmes intégrant un grand nombre de connaissances et de traitements dits systèmes intelligents, l'IA cherche à créer des systèmes capables de réaliser des tâches qui nécessiteraient normalement l'intelligence humaine, Par conséquent, l'Intelligence Artificielle conduit à cet ensemble d'interprétation où :

L'automatisation d'activités qui nous associons à la pensée humaine, comme la prise de décision, la résolution de problème ou l'apprentissage (BELLMAN 1978). [2]

1.3 Histoire et Évolution de l'IA

L'intelligence artificielle (IA), née dans les années 1950, depuis les premières réflexions sur la pensée mécanique jusqu'aux réseaux neuronaux profonds et à l'IA générative, cette discipline a été marquée par des avancées révolutionnaires.

1. Canard de Vaucanson 1738

Le Canard Digérateur est un automate créé par Jacques de Vaucanson. Il était possible de programmer les mouvements de cet automate, grâce à des pignons placés sur un cylindre gravé, qui contrôlaient des baguettes traversant les pattes du canard. Le mécanisme placé dans l'imposant piédestal, visible par tous, dans le but de montrer la complexité du travail accompli. Toutefois, il fallait attendre les années 30 pour que certains mathématiciens fassent progresser la logique ; analyse des fondements mêmes des raisonnements mathématiques et les années 50 pour que les calculateurs rentrent dans le domaine civil. Ces deux outils rassemblés ont provoqué la naissance de l'IA. [3]

2. Gestation de l'IA (1943-1955)

Pendant cette période furent menés les premiers travaux qui peuvent être considérés comme les débuts de l'intelligence artificielle (même si le terme n'existait pas encore). On peut citer les travaux de McCulloch et Pitts qui ont introduit en 1943 un modèle de neurones artificiels. Quelques années après, Hebb proposa une règle pour modifier des connections entre neurones, et Minsky et Edmonds construisirent le premier réseau de neurones. Ce fut aussi durant cette période que Turing publia son fameux article dans lequel introduit le test de Turing.[4]

3. Le test de TURING 1950

Développé par le mathématicien Alan Turing, le test de Turing consiste en une procédure expérimentale visant à attester de l'intelligence des machines. La preuve supposée s'obtient à l'aide de questions et de réponses tendant à corroborer l'impossibilité d'opérer une distinction entre l'intelligence humaine et artificielle dès lors que les interrogateurs (humains) sont incapables de différencier un interlocuteur humain d'un interlocuteur artificiel.[5]

4. Naissance d'IA (1956)

- Apparition des méthodes statistiques (plus proches voisins).
- Recherches en arbres : structures de listes, Géométrie, jeux.
- apparition du langage LISP pour l'IA développé par McCarthy.[6]

5. Systèmes Experts (1969-1979)

Le premier système expert, appelé DENDRAL, créé en 1969 pour déterminer La structure moléculaire d'une molécule étant donnés sa formule et les résultats de sa spectrométrie de masse. DENDRAL, est basé sur un grand nombre de règles heuristiques élaborées par des experts humains. Après le succès du DENDRAL, d'autres systèmes d'experts furent créés, tel que MYCIN, qui réalisait un diagnostic des infections sanguines.[7]

6. Le retour des réseaux de neurones :

Au milieu des années 80, quatre groupes de chercheurs ont découvrirent indépendamment la règle d'apprentissage "back-propagation" qui permit le développement de réseaux de neurones capables d'apprendre des fonctions très complexes. Depuis, l'apprentissage automatique est devenu l'un des domaines les plus actifs de L'IA, et a été appliqué avec succès a de nombreux problèmes pratiques (comme par exemple la fouille de données).[7]

1.4 Les Branches d'intelligence artificielle

1.4.1 Apprentissage Automatique (Machine Learning)

Le machine Learning consiste à développer des algorithmes et des modèles qui permettent aux ordinateurs d'apprendre à partir de données et de s'améliorer avec l'expérience. Il comprend des sous-domaines tels que l'apprentissage supervisé, l'apprentissage non supervisé, et l'apprentissage profond.[8]

1.4.2 Réseaux de Neurones Artificiels (Neural Networks)

Les réseaux de neurones artificiels sont des modèles informatiques inspirés de la structure du cerveau humain. Ils sont utilisés dans des domaines tels que la vision par ordinateur, la reconnaissance vocale, et la compréhension du langage naturel.[8]

1.4.3 Vision par Ordinateur (Computer Vision)

La vision par ordinateur se concentre sur la capacité des machines à interpréter et à comprendre des images et des vidéos. Elle est utilisée dans la reconnaissance d'objets, la détection de visages, la réalité augmentée, et bien d'autres applications.[8]

1.4.4 Traitement du Langage Naturel (Natural Language Processing – NLP)

Le NLP se penche sur la compréhension et la génération du langage humain par des machines. Il est utilisé dans les chatbots, les traducteurs automatiques, la recherche sémantique, et la classification de texte.[8]

1.4.5 Robotique

L'IA est utilisée pour la conception de robots et de systèmes autonomes capables de prendre des décisions, d'interagir avec leur environnement, et d'accomplir des tâches variées.[8]

1.5 Les systèmes multi-agents

Un agent intelligent représente un programme informatique capable de fonctionner de manière autonome. Il interagit avec son environnement en percevant des événements et en agissant en fonction d'objectifs spécifiques définis par sa programmation.[9]

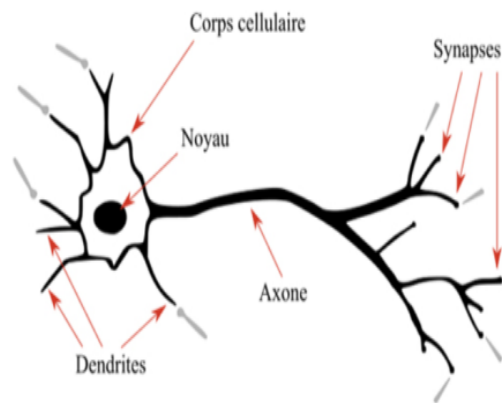
L'agent intelligent peut percevoir les conditions actuelles de la circulation, telles que la densité du trafic, les accidents, les travaux routiers.[10]

1.6 Les réseaux de neurones artificiels

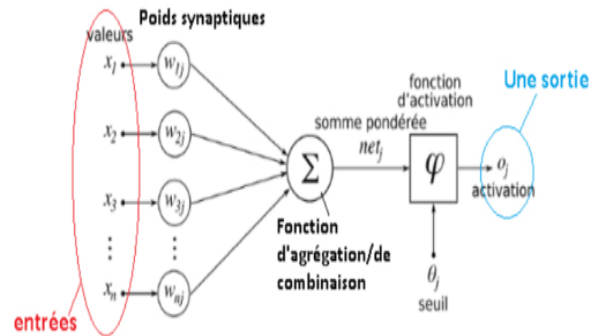
C'est les briques de base du Deep Learning, les réseaux de neurones artificiels tentent d'imiter le cerveau humain via une combinaison d'entrées de données, de poids et de biais.

Les réseaux neuronaux profonds sont constitués de plusieurs couches de nœuds interconnectés, chacune s'appuyant sur la précédente pour affiner et optimiser la prédiction ou la catégorisation. Cette progression des calculs à travers le réseau est ce que l'on appelle la propagation avant. Les couches d'entrée et de sortie d'un réseau neuronal profond sont appelées couches visibles. La couche d'entrée est celle où le modèle de Deep Learning ingère les données à traiter, et la couche de sortie est celle où la prédiction ou la classification finale est effectuée. Un autre processus appelé rétropropagation utilise des algorithmes, l'algorithme de descente de gradient par exemple, pour calculer les erreurs dans les prédictions, puis ajuste les poids et les biais de la fonction en remontant dans les couches pour entraîner le modèle. La propagation avant et la rétropropagation permettent au réseau neuronal de faire des prédictions et de corriger les erreurs éventuelles. Au fil du temps, l'algorithme devient plus précis.

Le mot « profond » dans l'apprentissage profond est attribué à ces couches cachées profondes. Chaque neurone de la couche cachée est connecté à de nombreux autres. Chaque flèche est associée à un poids, qui contrôle dans quelle mesure l'activation de ce neurone affecte les autres qui lui sont attachés.[11]



NEURONE BIOLOGIQUE



NEURONE ARTIFICIEL

Figure 1 : Fonctionnement du neurone artificiel

1.7 Deep Learning

Le Deep Learning (DL) est une méthode d'apprentissage automatique qui utilise des structures multicouches pour traiter l'information de manière complexe. Il se concentre sur la création de représentations hiérarchiques des données, en identifiant des caractéristiques de plus haut niveau à partir de celles de plus bas niveau. Ces caractéristiques sont apprises à partir des données elles-mêmes, ce qui permet au système de reconnaître des motifs et de classer les informations de manière automatique. Le point principal est que le Deep Learning utilise des réseaux neuronaux artificiels pour apprendre des niveaux de représentation des données, facilitant ainsi la compréhension et le traitement des informations [12]

1.7.1 Domaines d'application de l'apprentissage profond

Ces techniques se développent dans le domaine de l'informatique appliquée aux NTIC (reconnaissance visuelle par exemple d'un panneau de signalisation par un robot ou une voiture autonome — et vocale notamment) à la robotique, à la bio-informatique, la reconnaissance ou comparaison de formes, la sécurité, la santé, etc..., la pédagogie assistée par l'informatique, et plus généralement à l'intelligence artificielle. L'apprentissage profond peut par exemple permettre à un ordinateur de mieux reconnaître des objets hautement déformables et/ou analyser par exemple les émotions révélées par un visage photographié ou filmé, ou analyser les mouvements et position des doigts d'une main, ce qui peut être utile pour traduire le langage des signes, améliorer le positionnement automatique d'une caméra, etc... Elles sont utilisées

pour certaines formes d'aide au diagnostic médical (ex. : reconnaissance automatique d'un cancer en imagerie médicale), ou de prospective ou de prédiction (ex. : prédiction des propriétés).[13]

1.7.2 Principe et fondamentaux de l'apprentissage profond (DL)

Un modèle DL suit généralement les mêmes étapes de traitement que la modélisation par apprentissage automatique. Dans la figure 2, nous avons montré un travail d'apprentissage profond pour résoudre des problèmes du monde réel, qui se compose de trois étapes de traitement, telles que la compréhension et le prétraitement des données, la construction de modèles DL, l'entraînement, la validation et l'interprétation.

Cependant, contrairement à la modélisation ML, l'extraction des caractéristiques dans le modèle DL est automatisée plutôt que manuelle. K plus proche voisin, l'arbre de décision, la forêt aléatoire, le clustering k-means, sont quelques exemples de techniques d'apprentissage automatique couramment utilisées dans divers domaines d'application. D'autre part, le modèle DL comprend les réseaux neuronaux artificiels. [14]

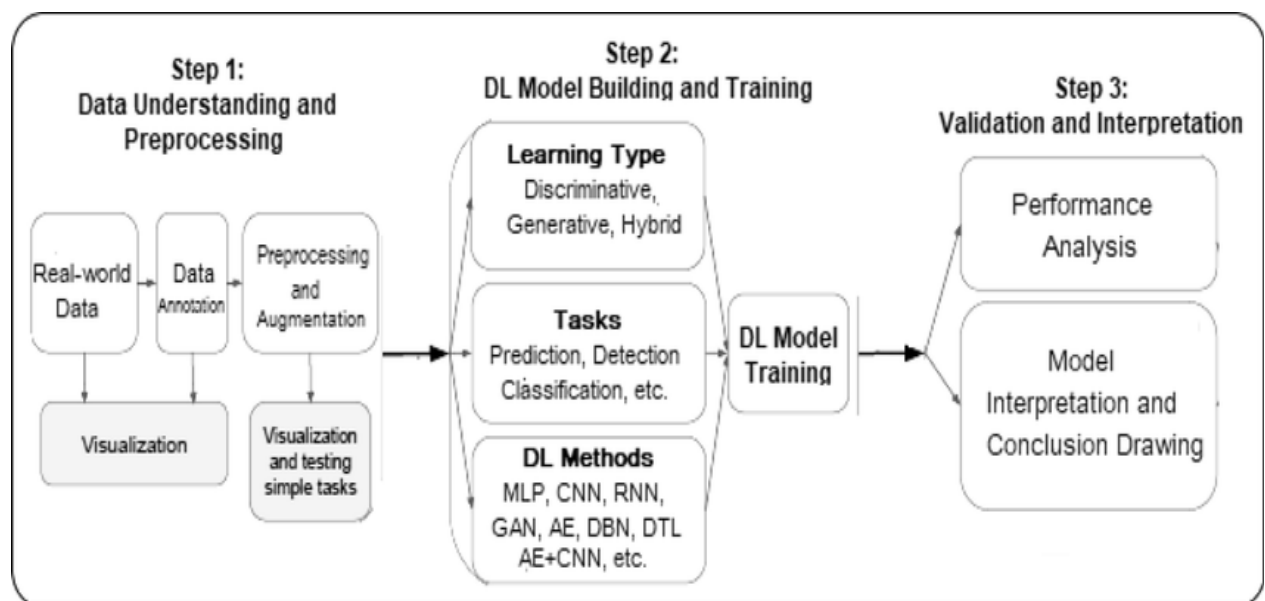


Figure 2 : Illustration d'un processus d'apprentissage pour la résolution de problèmes réelles

1.7.3 Positionnement du Deep Learning dans l'IA

Les termes : intelligence artificielle, machine Learning (apprentissage automatique ou apprentissage statistique) et Deep Learning (apprentissage profond) se sont popularisés. Ils sont parfois utilisés de manière interchangeable [15]

L'IA est l'imitation d'une réponse d'un expert dans un domaine, Le ML est un sous domaine de l'IA, adapte son algorithme mais a besoin d'un expert pour les données, le Deep Learning est du ML utilisé avec des réseaux de neurones, il extrait l'information pertinente des données et adapte son algorithme. [16]

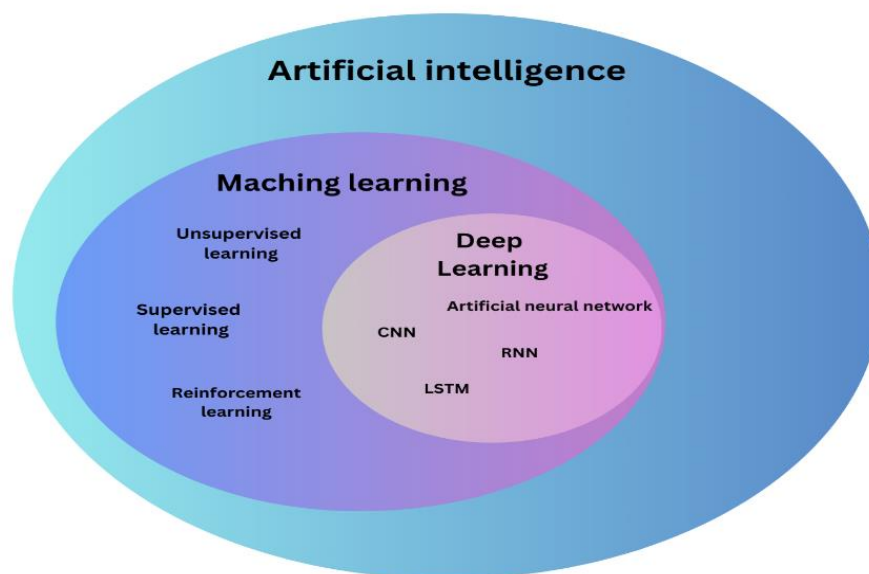


Figure 3: Relation entre IA, ML et DL

La principale différence entre le Deep Learning et le machine Learning réside dans la structure de l'architecture du réseau neuronal sous-jacent. Les modèles de machine Learning traditionnels (non profonds), utilisent des réseaux neuronaux simples, avec une ou deux couches de calcul. Les modèles de Deep Learning utilisent trois couches ou plus, mais généralement des centaines ou des milliers de couches, pour l'entraînement.

Alors que les modèles d'apprentissage supervisé nécessitent des données d'entrée structurées et étiquetées pour obtenir des résultats précis, les modèles de Deep Learning peuvent s'appuyer sur un apprentissage non supervisé. Avec l'apprentissage non supervisé, les modèles de Deep Learning peuvent extraire les caractéristiques, les éléments et les relations dont ils ont besoin pour produire des résultats précis à partir de données brutes non structurées. [17]

1.8 Les principales architectures de DeepLearning

1.8.1 Réseau de neurones convolutifs (CNN)

Le réseau neuronal convolutif (CNN ou ConvNet) est une architecture d'apprentissage profond discriminative populaire qui apprend directement à partir de l'entrée sans avoir besoin d'extraire des caractéristiques humaines. En conséquence, Chaque couche dans CNN prend en compte des paramètres optimaux pour une sortie significative et réduit la complexité du modèle. Les CNN sont spécifiquement destinés à traiter une variété de formes 2D et sont donc largement utilisés dans la reconnaissance visuelle, l'analyse d'images médicales, la segmentation d'images, le traitement du langage naturel. La capacité de découvrir automatiquement les caractéristiques essentielles à partir de l'entrée sans avoir besoin d'intervention humaine le rend plus puissant qu'un réseau traditionnel. Qui peuvent être utilisés dans divers domaines d'application en fonction de leurs capacités d'apprentissage.[18]

1.8.2 Les réseaux de neurones récurrents (RNN)

Un réseau neuronal récurrent (RNN) est un autre réseau neuronal populaire, qui utilise des données séquentielles ou de séries chronologiques et alimente la sortie de l'étape précédente en tant qu'entrée à l'étape actuelle. Comme le feedforward et le CNN, les réseaux récurrents apprennent des entrées d'entraînement, mais se distinguent par leur « mémoire », qui leur permet d'influencer l'entrée et la sortie actuelles en utilisant les informations des entrées précédentes. Contrairement aux DNN typiques, qui supposent que les entrées et les sorties sont indépendantes les unes des autres, la sortie des RNN dépend d'éléments antérieurs dans la séquence. Cependant, les réseaux récurrents standard ont le problème de la disparition des gradients, ce qui rend difficile l'apprentissage de longues séquences de données. Dans ce qui suit, nous abordons plusieurs variantes populaires du réseau récurrent qui minimisent les problèmes et fonctionnent bien dans de nombreux domaines d'application du monde réel.[18]

1.8.3 Réseau antagoniste génératif (GAN)

Un réseau antagoniste génératif (GAN), développé par Ian Goodfellow, est une architecture de réseau neuronal utilisée pour la modélisation générative, permettant de créer de nouveaux échantillons réalistes à partir de données d'entrée. Il se compose de deux réseaux neuronaux en compétition : un générateur qui produit des données similaires aux données d'origine, et un discriminateur qui évalue si les échantillons proviennent de données réelles ou sont générés. Le générateur cherche à tromper le discriminateur en créant des données de plus

en plus réalistes, tandis que le discriminateur s'efforce de distinguer les données authentiques des fausses. Ce processus d'apprentissage mutuel permet d'améliorer la qualité des données générées au fil du temps.[18]

1.8.4 Transformer

Est un modèle d'apprentissage profond de premier plan qui a été largement adopté dans divers domaines, tels que le traitement du langage naturel (NLP), la vision par ordinateur (CV) et le traitement de la parole. Transformer a été initialement proposé comme un modèle séquence à séquence pour la traduction automatique . Des travaux ultérieurs montrent que les modèles pré-entraînés basés sur Transformer (PTM) peuvent atteindre *des performances de pointe* sur diverses tâches. En conséquence, Transformer est devenu l'architecture de référence en NLP, en particulier pour les PTM. Outre les applications liées au langage, Transformer a également été adopté dans le CV;le traitement audio et même d'autres disciplines, telles que la chimie et les sciences de la vie.[19]

1.9 Les techniques d'apprentissage

1.9.1 Apprentissage supervisé

L'apprentissage supervisé est la tâche d'apprentissage automatique Consistant à apprendre une fonction qui mappe une entrée à une sortie en fonction d'exemples de paires entrée-sortie. Il déduit une fonction à partir de données d'entraînement étiquetées constituées d'un ensemble d'exemples d'entraînement. Les algorithmes d'apprentissage automatique supervisé sont ceux qui nécessitent une assistance externe. L'ensemble de données d'entrée est divisé en ensemble de données d'entraînement et de test. L'ensemble de données d'entraînement a une variable de sortie qui doit être prédite ou classée. Tous les algorithmes.

1.9. 2 Apprentissage non supervisé

L'apprentissage non supervisé est une méthode où le système apprend sans la présence d'étiquettes de données. Dans ce processus, le réseau apprend à représenter les données internes ou à identifier des caractéristiques importantes pour découvrir des relations ou une structure inconnue dans les données d'entrée. Les approches courantes incluent le regroupement, la réduction de la dimensionnalité et les techniques génératives. Les techniques populaires de Deep Learning pour cela comprennent les encodeurs automatiques (AE), les machines de Boltzmann restreintes (RBM) et les réseaux adversaires génératifs (GAN). De plus, les réseaux de neurones récurrents (RNN) tels que le LSTM sont également utilisés dans de nombreux

domaines pour l'apprentissage non supervisé. [20]

1.9.3 L'apprentissage semi-supervisé

Est une méthode qui repose sur des ensembles de données partiellement étiquetés, parfois également appelés apprentissage par renforcement. Cette approche utilise des techniques telles que les réseaux adversaires génératifs (GAN), les réseaux de neurones récurrents (RNN) comme LSTM et GRU, ainsi que l'apprentissage par renforcement profond (DRL). Elle permet d'exploiter à la fois les données étiquetées et non étiquetées pour entraîner un modèle.[21]

1.10 Les algorithmes d'apprentissage supervisé

1.10.1 Arbre de décision

L'arbre de décision est un graphique qui représente les choix et leurs résultats sous forme d'arbre. Les nœuds du graphique représentent un événement ou un choix et les arêtes du graphique représentent les règles ou conditions de décision. Chaque arbre est constitué de nœuds et de branches. Chaque nœud représente des attributs dans un groupe qui doit être classé et chaque branche représente une valeur que le nœud peut prendre.

1.10.2 Naïve Bayes

Il s'agit d'une technique de classification basée sur le théorème de Bayes avec une hypothèse d'indépendance entre les prédicteurs. En termes simples, un classificateur naïf de Bayes suppose que la présence d'une caractéristique particulière dans une classe n'est pas liée à la présence d'une autre caractéristique. Naïve Bayes cible principalement le secteur de la classification de texte. Il est principalement utilisé à des fins de clustering et de classification, qui dépend de la probabilité conditionnelle de ce qui se produit.

1.10.3 Machine à vecteurs de support

Une autre technique de pointe en matière d'apprentissage automatique est la machine à vecteurs de support (SVM). Dans l'apprentissage automatique, les machines à vecteurs de support sont des modèles d'apprentissage supervisés avec des algorithmes d'apprentissage associés qui analysent les données utilisées pour la classification et l'analyse de régression. En plus d'effectuer une classification linéaire, les SVM peuvent effectuer efficacement une classification non linéaire en utilisant ce que l'on appelle l'astuce du noyau, en mappant implicitement leurs entrées dans des espaces de caractéristiques de grande dimension. En gros,

cela permet de tracer des marges entre les classes. Les marges sont tracées de telle manière que la distance entre la marge et les classes soit maximale et donc de minimiser l'erreur de classification. .[22]

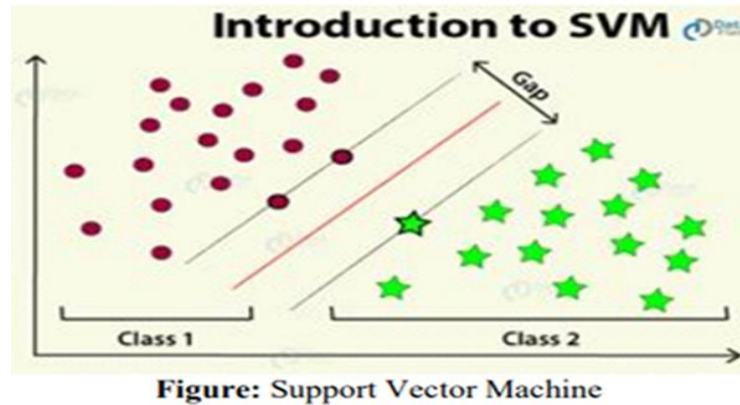


Figure 4 : SVM

1.11 Les algorithmes d'apprentissage non supervisé

1.11.1 K-Means

Le K-means est l'un des algorithmes d'apprentissage non supervisé les plus simples qui résolvent le problème bien connu du clustering. La procédure suit une méthode simple et facile pour classer un ensemble de données donné à travers un certain nombre de clusters. L'idée principale est de définir k centres, un pour chaque cluster. Ces centres doivent être placés de manière astucieuse car des emplacements différents entraînent des résultats différents. Ainsi, le meilleur choix est de les placer le plus loin possible les uns des autres.[22]

1.11.2 FCMeans " Le Fuzzy C-Means"

Est un algorithme de clustering flou utilisé en apprentissage non supervisé pour regrouper des données en plusieurs clusters, en permettant à chaque point d'appartenir à plusieurs groupes avec un degré d'appartenance compris entre 0 et 1. Contrairement à l'algorithme K-Means, où chaque donnée appartient à un seul cluster, FC Means attribue des poids d'appartenance à chaque point, qui sont ajustés à chaque itération en fonction de leur distance aux centres des clusters, lesquels sont également mis à jour. Il vise à minimiser une fonction de coût basée sur ces distances pondérées. Ce modèle est particulièrement utile dans

les cas où les frontières entre les classes ne sont pas clairement définies, comme en segmentation d'images, en reconnaissance vocale ou en bio-informatique, où un élément peut appartenir partiellement à plusieurs catégories.

1.12. Conclusion

Dans ce premier chapitre, nous avons présenté les fondements de l'Intelligence Artificielle (IA) ainsi que les principes clés du Deep Learning, une de ses branches les plus prometteuses. Nous avons expliqué l'évolution de l'IA, ses domaines d'application, et ses principales approches. Une attention particulière a été accordée aux réseaux neuronaux profonds, en mettant en évidence leur capacité à apprendre automatiquement à partir de grandes quantités de données. Ces concepts constituent la base théorique indispensable à la compréhension et à la mise en œuvre de notre projet, qui s'appuie sur les techniques du Deep Learning pour répondre à un problème réel : la détection des erreurs dans la récitation coranique. Le chapitre suivant sera consacré à une étude approfondie du domaine d'application, à savoir les spécificités de la récitation coranique.

Chapitre 2 :
La science de la récitation du
Coran

2.1 Introduction

Dans ce chapitre, nous allons explorer un pilier central de la pratique religieuse islamique : la récitation du Saint Coran, ainsi que les règles qui l'encadrent, connues sous le nom de Tajwid. La récitation correcte du Coran revêt une importance capitale, tant sur le plan spirituel que culturel, dans la vie des musulmans. Ce chapitre mettra en lumière la signification profonde de la récitation coranique, en insistant sur son rôle dans la préservation du message divin, et présentera les principes fondamentaux du Tajwid. L'objectif est d'offrir une compréhension claire et approfondie de cette pratique essentielle qui unit les musulmans à travers le monde dans leur rapport au texte sacré.

2.2 Définition du Coran

Le Coran est la parole divine révélée par Allah à Son dernier messenger, le prophète Muhammad (ص), par l'intermédiaire de l'ange Jibril (Gabriel). Il constitue la source principale de la foi et de la législation islamique, et se distingue par son éloquence, sa structure inimitable, et la profondeur de ses enseignements. Malgré les répétitions apparentes de certains passages, le Coran continue de fasciner et d'inspirer ceux qui le lisent ou l'écoutent, dévoilant sans cesse de nouvelles merveilles. Plusieurs versets témoignent de sa grandeur et encouragent les croyants à s'y attacher par la lecture, la méditation et la récitation conforme aux règles établies. La parole de Dieu Tout-Puissant : { إِنَّ هَذَا الْقُرْآنَ يَهْدِي لِلَّتِي هِيَ أَقْوَمُ وَيُبَشِّرُ الْمُؤْمِنِينَ الَّذِينَ يَعْمَلُونَ الصَّالِحَاتِ أَنَّ لَهُمْ أَجْرًا كَبِيرًا } ... (الإسراء : 9)

2.3 Définition et règles du Tajwid

Le Tajwid est l'art de réciter le Coran de manière correcte et harmonieuse, en articulant chaque lettre avec précision selon ses caractéristiques phonétiques et articulatoires. Fondé sur des règles codifiées et transmises fidèlement depuis le Prophète Muhammad (ص), il vise à garantir une récitation conforme au modèle prophétique et à préserver le sens authentique des versets en évitant les erreurs de prononciation. La connaissance du Tajwid constitue un devoir collectif, tandis que sa mise en pratique est un devoir individuel pour toute personne souhaitant réciter ne serait-ce qu'une partie du Saint Coran, afin d'en respecter le caractère sacré et d'en assurer la préservation. [23]

2.4 Le développement de la science de Tajwid

L'élaboration systématique de la science du Tajwid a véritablement commencé au début du quatrième siècle de l'Hégire. Toutefois, des références à ses principes apparaissent bien avant cette période dans des œuvres non exclusivement consacrées à cette discipline. On peut citer, par exemple, une lettre attribuée à Abū 'Amr ibn al-'Alā' al-Baṣrī († 145 H), ainsi qu'une versification de la récitation coranique attribuée à Qālūn al-Madanī († 220 H). Le premier ouvrage autonome entièrement dédié au Tajwid remonte à Abu Mazāhim al-Khāqānī († 325 H), qui composa un poème fondateur sur le sujet. [23]

Par la suite, plusieurs savants ont enrichi cette science par leurs contributions. Parmi eux figure Al-Sa'dī' Alī ibn Ja'far († 410 H), auteur de Al-Tanbīh'alā al-Laḥn al-Jalīwa al-Laḥn al-Khafī, une œuvre traitant des fautes de récitation manifestes et subtiles, ainsi que des divergences entre récitateurs concernant les lettres lām et nūn. D'autres ouvrages importants ont suivi, notamment Al-Ri'āyah li-Tajwīd al-Qirā'ahwaTaḥqīqLafz al-Tilāwah de Makkī ibn AbīTālib al-Qaysī († 437 H), Al-Taḥdīdfī al-Itqānwa al-Tajwīd d'Al-Dānī († 444 H), et Al-Mawḍiḥfī'Ilm al-Tajwīd de 'Abd al-Wahhāb al-Qurṭubī († 462 H). [23]

2.5 Sections du Tajwid et les jugements de chaque elle

La science du Tajwid se structure autour de deux dimensions fondamentales et complémentaires:

La dimension théorique : Elle consiste en l'apprentissage des règles, des fondements et des principes propres à la science du Tajwid. Cette connaissance comprend la mémorisation, la compréhension et l'analyse des notions qui régissent la bonne récitation du Coran. D'un point de vue juridique, l'acquisition de ces connaissances est obligatoire (fard 'ayn) dans la mesure où elle permet d'éviter les erreurs altérant le sens du texte sacré.

La dimension pratique : Elle repose sur la mise en application des règles théoriques lors de la récitation coranique. Cette pratique peut s'accomplir selon deux approches :

L'approche par transmission orale (riwāya) : Elle implique l'apprentissage de la récitation sous la supervision d'un enseignant qualifié (cheikh), jusqu'à atteindre une maîtrise précise et fidèle à la tradition.

L'approche par l'étude intellectuelle (dirāya) : Elle se base sur la compréhension autonome des règles théoriques suivie de leur application correcte lors de la lecture des versets.

Il est impératif pour toute personne récitant le Coran de posséder une compétence pratique en Tajwid, afin de garantir une prononciation juste et conforme à la tradition prophétique. [23]

2.6 Les niveaux de récitation

La Méditation (Tahqiq) : C'est similaire au Tartil, mais avec un niveau de réflexion plus profond, et c'est une étape utilisée dans l'enseignement.

La vigilance (Hadur) : implique une récitation rapide tout en respectant strictement les règles Du Tajwid.

La modulation (Tadwir) : correspond à une récitation se situant entre la méditation et la vigilance, tout en respectant rigoureusement les règles du Tajwid. [23]

2.7 Les fautes de récitation (Al-Lahn)

Justesse de la prononciation ou des règles de récitation. Ces fautes sont classées en deux catégories principales :

Le Lahn Jalî (اللحن الجلي) : Il s'agit d'une faute manifeste et audible qui altère la structure du mot ou en modifie le sens. Cela peut inclure la substitution d'une lettre par une autre, ou encore le changement d'une voyelle par une autre, entraînant ainsi une déformation du sens.

Exemples : remplacer la lettre « د » par « ذ », ou « ح » par « ع », comme dans : Le verset 7 de la Sourate An-Nazi'at : changer une lettre peut altérer profondément le sens.

Ou encore, changer une voyelle comme passer d'une 'Fateha' à une 'Kasra', par exemple dans : (صِرَاطَ الَّذِينَ أَنْعَمْتَ عَلَيْهِمْ), en lisant incorrectement " أَنْعَمْتَ " au lieu de " أَنْعَمْتَ ".

Le Lahn Khafî (اللحن الخفي) : Il s'agit d'une faute subtile et moins perceptible, qui consiste à enfreindre certaines règles de Tajwid, comme l'omission d'une ghunna (nasalisation) ou d'une madd (prolongation).

Statut juridique : Cette erreur est réprouvée (makrûh), voire interdite selon certains avis. Elle est qualifiée de "khafî" (cachée) car seule une personne maîtrisant le Tajwid peut la détecter avec précision. [23]

2.8 Le Waqf (L'arrêt dans la récitation)

Le waqf désigne l'action de marquer une pause dans la récitation du Coran, que ce soit à la fin ou au milieu d'un verset, permettant au lecteur de reprendre son souffle sans altérer le sens global du texte. La pause s'effectue en s'arrêtant à la fin d'un mot, en suspendant la voyelle finale le cas échéant, et en reprenant la récitation sans qu'il soit nécessaire de répéter la basmala ou la formule de protection (isti'ādha), même si la pause est prolongée. Il est essentiel de ne pas couper un mot en deux ou d'interrompre une structure grammaticale liée, afin de préserver la fluidité et la cohérence du sens.

On distingue plusieurs types d'arrêts, classés selon les circonstances de la pause [23]

2.8.1 L'arrêt contraint (الوقف الاضطراري)

Cet arrêt est imposé par des facteurs extérieurs à la volonté du lecteur, comme une toux, un étouffement, un essoufflement, ou un oubli. Il n'est donc pas considéré comme une erreur. Le lecteur peut alors s'arrêter sur n'importe quel mot, même si le sens reste incomplet, comme après une conjonction sans complément, ou une condition sans conséquence. Toutefois, la reprise de la lecture doit se faire à partir du mot approprié pour restaurer le sens grammatical et sémantique.

Exemple : Arrêter sur « إن » (Si) sans lire la suite rompt le sens, mais peut être accepté dans un arrêt contraint. [23]

2.8.2 L'arrêt évaluatif (الوقف التعليمي)

Utilisé principalement dans un contexte pédagogique, cet arrêt est demandé par un enseignant à un élève pour évaluer sa compréhension ou corriger sa récitation. Il permet de mettre en évidence les mots isolés, les liaisons grammaticales, ou les omissions. Ce type de pause ne fait pas partie de la récitation normale, mais sert d'outil d'apprentissage.

Exemple : L'enseignant peut demander à l'élève de s'arrêter sur un mot pour discuter de sa signification ou de son lien syntaxique avec le suivant.

2.8.3 L'arrêt d'attente (الوقف الانتظاري)

Ce type d'arrêt est utilisé pour comparer différentes lectures (qirā'āt). Le lecteur marque une pause après chaque mot ou groupe de mots afin d'examiner les variantes de prononciation ou de vocalisation, avant de poursuivre. Il est principalement appliqué dans les sciences de la récitation avancée.

Exemple : Comparer la lecture de « مالك يوم الدين » (Mālikyawmi-d-dīn) avec « ملك يوم الدين » selon les différentes qirā'āt. [23]

2.8.4 L'arrêt optionnel (الوقف الاختياري)

Cet arrêt est choisi librement par le lecteur, sans contrainte externe. Il se divise en plusieurs sous-catégories, selon la relation entre la partie arrêtée et ce qui suit :

L'arrêt complet (تام) : La pause intervient après un énoncé totalement indépendant, sans lien sémantique ou grammatical avec la suite.

Exemple : « والله سميع عليم » – la pause ici est complète car ce qui suit n'est pas directement lié.

L'arrêt adéquat (كافٍ) : L'arrêt se fait sur un sens complet, mais le discours suivant présente un lien logique, bien que non grammatical.

Exemple : « إن الله غفور رحيم. والله يحب المحسنين » – sens complet, mais continuité logique.

L'arrêt convenable (حسن) : Il y a un lien à la fois sémantique et grammatical entre la phrase arrêtée et ce qui suit. La pause est tolérée si elle n'altère pas le sens.

Exemple : « لا إله إلا الله. محمد رسول الله »

L'arrêt inapproprié (قبيح) : C'est une pause inacceptable qui coupe le sens, entraînant une déformation du message ou une incompréhension.

Exemple : S'arrêter sur « لا تعبدون » dans le verset « ولا أنا عابد ما عبدتم » peut altérer gravement le sens. [23]

2.9 L'Initiation

Commencer la lecture après une pause ou un arrêt les défèrent types :

2.9.1 Le début véritable

Il correspond à une reprise de la récitation après une longue interruption ou un changement d'activité. C'est un commencement authentique, indépendant du passage précédent.

Exemple : Commencer une nouvelle séance de récitation par le début d'une sourate.

2.9.2 Le début additionnel

Il survient après une courte pause, généralement pour respirer, sans quitter l'intention de continuer la récitation. Ce type de début est temporaire et dépend du contexte immédiat.

Exemple : S'arrêter un instant après un mot comme « الحمد » dans « الحمد لله رب العالمين », puis reprendre sans changer de sujet. [23]

2.9.3 Le début absolu

Ce début se fait à partir d'un énoncé complètement autonome, sans lien ni sémantique ni grammatical avec ce qui précède. C'est typiquement le cas au début d'une sourate.

Exemple : Commencer par « تبارك الذي بيده الملك » ou « قل هو الله أحد » [23]

2.9.4 Le début approprié

Il s'agit d'un commencement lié au discours précédent à la fois sur le plan du sens et de la grammaire. Il est permis uniquement si le verset est repris à partir d'un point qui maintient le sens intact. Il n'est pas valable comme début indépendant.

Exemple : Commencer par « فبأي آلاء ربكما تكذبان » après une pause dans la sourate Ar-Rahmān

nécessite un retour réfléchi pour garder le lien avec ce qui précède. [23]

2.10 Dessin coranique

2.10.1 Définition du dessin coranique

Dessiner le Coran, ou comme certains érudits l'appellent le dessin d'Uthmani, c'est écrire quelques mots du Saint Coran, dans un dessin autre que le dessin orthographique connu. Elle a été approuvée par le Prophète - que la paix et la bénédiction de Dieu soient sur Lui et a été acceptée par les nobles compagnons. Elle a été adoptée par Uthman ibn Affan - qu'Allah soit satisfait de lui - dans le mus'haf (coran) qu'il a fait écrire par le comité de Zaid ibn Thabit. Ce mus'haf a été nommé le mus'haf d'Uthman, et il a été transmis par la nation avec un consensus continu jusqu'à nos jours.

2.10.2 Avantages du dessin coranique

Indiquer l'origine du mouvement

La lettre "ya" a été ajoutée dans la parole de Dieu Tout-Puissant {وَلَقَدْ جَاءَكَ مِنْ نَبِيِّ الْمُرْسَلِينَ} [Al-An'am : 34] pour indiquer le kasrah, et la lettre "wa" a été ajoutée dans la parole du Tout-Puissant {سَأْرِيكُمْ دَارَ الْفَاسِقِينَ} [Al-A'raf : 1 45] pour indiquer le dhamma, car le Coran a été rédigé sans hamza, ponctuation et voyelles.

Indication de l'origine de la lettre

Écrire "صلاة" avec un "و" au lieu d'un "أ" dans : "صلوة" [Al-Baqarah : 3] indique que l'origine de la lettre est le "و" dans le mot "صلوة" et écrire "يا" et "ياء" indique qu'ils sont parmi Les lettres "ياء" qui, de par leur nature, inclinent vers leur inclinaison, comme dans : « Ad-Duha : والضحى » [1]. [23]

Prise en compte des différentes lectures du dessin

Le mot "عباد" dans la parole de Dieu Tout-Puissant : "وجعلوا الملائكة الذين هم عباد الرحمن اناثا" [AzZukhruf : 19] a été lu comme "عباد" et "عبد" et le dessin est possible pour les deux lectures. [23]

Se référant à une autre lecture

Le mot [بينة]. Apparaît dans dix-neuf endroits. Dans tous ces cas, il est écrit avec un "ت" Attaché, sauf dans un seul endroit où il est écrit avec un "ت" ouvert, dans le verset : « أم آتيناهم) » [Fatir : 40]. Cela indique une autre lecture. Certains récitants lisent ce verset

Au pluriel et d'autres au singulier, donc il est écrit avec un "د" ouvert pour permettre les deux lectures.

Indiquer un sens caché précis

L'ajout de la lettre "ا" après le verbe dans le mot "يعفو" dans la parole divine : " وَيَعْفُوا عَنْكثير " [Fatir: 34] indique l'abondance et la continuité du pardon de Dieu

- L'augmentation de la Construction renforce le sens.

L'écriture de «ايد» avec deux "ياء" dans la parole divine ، Dhariyat: وَالسَّمَاءَ بَنَيْنَاهَا بِأَيْدٍ وَإِنَّا لَمُوسِعُونَ { [47] est pour souligner la grandeur de Dieu.

L'ajout d'un "همزة" sur la "واو" et l'ajout de la lettre "ا" après cela dans le mot "تفتوا" dans la Parole divine : « { تَاللَّهِ تَفْتَأُ تَذُكُرُ يُوسُفُ } » [Yusuf : 85] indique la fréquence de la mention de Joseph par Jacob, que la paix soit sur eux. - L'ajout de la lettre "ا" dans le mot "وجاء" dans la parole divine : { وَجِيءَ يَوْمَئِذٍ بِجَهَنَّمَ } [Al-Fajr : 23] Est pour l'accentuation, l'exagération et la menace. - La suppression de la lettre "ا" après la "واو" du pluriel dans le mot "وجاؤا" dans la parole Divine : " [Al-A'raf : 116] (116) وَجَاءُوا بِسِحْرِ عَظِيمٍ indique que leur arrivée est dominée par le Mensonge et la tromperie. [23]

Fournir des significations différentes d'une manière qui n'est pas cachée

La séparation de "م" dans la parole divine : " أَمْ مَنْ يَكُونُ عَلَيْهِمْ وَكَيْلًا " [An-Nisa : 109] Indique que "م" est détaché, ce qui signifie "plutôt", tandis que "من" est lié. La liaison de "م" dans la parole divine : " أَفَمِنْ يَمْشِي مَكْبَأًا عَلَىٰ وَجْهِهِ أَهْدَىٰ أَمْ مَنْ يَمْشِي سَوِيًّا عَلَىٰ صِرَاطٍ مُسْتَقِيمٍ " [Al-Mulk : 22] indique que "م" est relié à "من". [23]

Encouragez les gens à recevoir le Saint Coran de sources fiables

Le dessin uthmanique différait du dessin standard par certains mots. Par exemple, au début de 29 sourates du Saint Coran, il y a des lettres disjointes composées de 14 lettres regroupées En : " ن ، ق ، ص " ; Trois d'entre elles commençaient par une lettre : " طرق سمعك النصيحة " ; neuf Commençaient par deux lettres : " طه ، يس ، حم ، طس " , treize commençaient par trois lettres telles Que طسم Que المر ، المص ، المر ، et deux commençaient par quatre lettres : " ألم ، الر ، طسم Que كهيصص ، حم عسق Ces lettres sont lues par leurs noms, et non par leur prononciation. Cinq de ces lettres Regroupées dans : " حي طهر " se prononcent comme deux lettres comme ceci ها ، را ، يا ، طا ، ها ، Et les huit lettres restantes regroupées dans : " نقص عسلكم " se prononcent comme trois lettres Comme ceci : " نون ، قاف ، صاد ، عين ، سين ، لام ، كاف ، ميم " Si la question avait été laissée au format Standard, les gens ne se seraient pas précipités pour recevoir le Saint

Coran de sources fiables, Et "ألف لام ميم. [Al-Baqarah : 1] aurait été récité "الم

2.11 Définir et Comparer les travaux existants

Plusieurs projets technologiques récents ont été développés pour faciliter l'interaction avec le Coran à l'aide de l'intelligence artificielle. Parmi eux, Tarteel AI est une application basée sur le Deep Learning, spécialisée dans la reconnaissance vocale de la récitation coranique. Elle permet aux utilisateurs de réciter à haute voix et détecte automatiquement les erreurs de prononciation, les oublis ou les ajouts, tout en offrant un retour en temps réel. Ce projet met l'accent sur l'amélioration individuelle de la récitation en respectant les règles de Tajwid.

En parallèle, iReceptor est une plateforme collaborative destinée à l'analyse et au partage de données immunogénétiques, qui bien qu'elle ne soit pas directement liée au Coran, sert ici d'exemple d'un système distribué de traitement de données massives à des fins scientifiques. Il peut être mentionné pour illustrer les approches techniques et méthodologiques utilisées dans des projets d'IA avancés.

Son comparaison, Tarteel AI se distingue par son application directe au domaine religieux, en particulier la récitation coranique, avec une interface conviviale et une spécialisation dans l'analyse audio linguistique. iReceptor, quant à lui, représente un modèle d'architecture collaborative pour la gestion de grandes bases de données et le traitement complexe, ce qui peut inspirer des projets similaires dans le domaine de la récitation ou de l'enseignement du Coran. Ainsi, ces deux projets, bien que différents dans leurs objectifs, offrent des perspectives complémentaires sur l'usage de l'intelligence artificielle dans des contextes variés.

2.11 Conclusion

Ce chapitre a mis en lumière le rôle fondamental de la récitation correcte du Coran, régie par les règles du Tajwid, dans la préservation de l'authenticité et de la beauté du texte sacré. En définissant les concepts du Coran et du Tajwid, et en exposant les différents niveaux de récitation ainsi que les types d'erreurs pouvant survenir, nous avons souligné l'importance cruciale de la maîtrise de ces éléments pour une lecture conforme et respectueuse.

Par ailleurs, le rasm al-mushaf (le dessin orthographique du Coran) occupe une place essentielle dans la facilitation d'une récitation fidèle aux enseignements prophétiques. Il constitue un

repère visuel et phonétique permettant d'éviter les erreurs et d'assurer une transmission précise du texte.

En somme, l'application rigoureuse des règles de Tajwid et la compréhension du rasm contribuent à renforcer la transmission orale du Coran, pilier central de la tradition islamique, tout en nourrissant la dimension spirituelle de la récitation.

Chapitre 3 :
Application de Deep Learning
à la reconnaissance vocale du
Coran

3.1 Introduction

Le domaine de la reconnaissance vocale de la récitation coranique est connu des avancées remarquables ces dernières années, principalement en raison des progrès réalisés dans le domaine du Deep Learning. Les techniques modernes basées sur les réseaux de neurones profonds ont permis de dépasser les méthodes traditionnelles, offrant ainsi une capacité d'analyse plus fine des nuances de la récitation. Cela inclut l'identification des différents styles de récitation et la reconnaissance des voix individuelles, éléments essentiels dans le contexte de la lecture du Coran où la précision est primordiale. Dans ce cadre, deux méthodes principales de validation de la récitation sont utilisées : la comparaison vocale, qui impacte la performance phonétique, et la comparaison textuelle, qui assure l'exactitude du contenu récité. Ces approches contribuent à garantir une expérience d'apprentissage enrichissante et conforme aux règles du tajwid.

3.2 La reconnaissance automatique de la parole

La reconnaissance automatique de parole, ou transcription automatique de parole, est une technologie qui convertit la voix humaine en texte à l'aide de microphones et d'analyses sonores. Elle est utilisée dans des applications telles que la dictée vocale, les commandes vocales et l'interaction avec des assistants virtuels. Ce domaine de recherche en intelligence artificielle cherche à imiter la capacité humaine de transcription de la parole. [24]

3.3 La Reconnaissance Vocale

La reconnaissance vocale consiste à convertir la parole humaine en texte ou en commandes numériques. Elle utilise des algorithmes et des modèles pour interpréter des signaux audios.

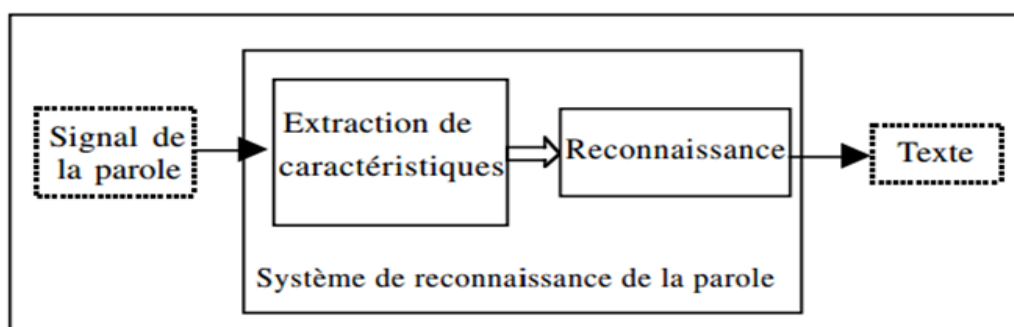


Figure 5 : Le processus de la reconnaissance vocale

3.3.1 Acquisition du signal vocal (Signal de parole)

Le processus débute par la capture du signal vocal à l'aide d'un microphone ou d'un autre dispositif d'enregistrement audio. Ce signal constitue une représentation analogique de la parole humaine, traduisant les nuances et les inflexions de la voix.

3.3.2 Prétraitement du signal

Avant d'être analysé, le signal vocal brut subit un prétraitement visant à supprimer le bruit de fond, à ajuster le volume pour une uniformité et éventuellement à réduire sa complexité dimensionnelle.[25]

L'objectif fondamental de cette étape est de transformer le signal analogique recueilli précédemment en un signal numérique, un processus connu sous le nom de conversion analogique-numérique (CAN), afin qu'il puisse être traité et compris par la machine.[26]

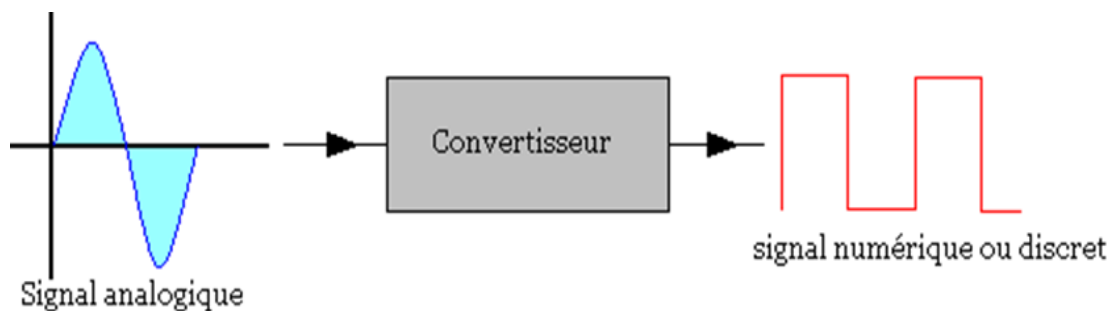


figure 1: Un processus de numérisation

Figure 6 : Processus de numérisation

3.3.3 La modélisation acoustique

Dans le domaine de la reconnaissance vocale débute par la capture et le traitement du signal audio, suivis de l'extraction de caractéristiques à l'aide de techniques telles que les Coefficients Cepstraux de Fréquence Mel (MFCC) ou les spectrogrammes. Ces caractéristiques sont essentielles pour l'entraînement de modèles acoustiques, qui reposent souvent sur des architectures de réseaux de neurones récurrents (RNN) ou convolutifs (CNN), permettant d'aligner les séquences audios avec les phonèmes correspondants. Par la suite, un modèle linguistique est intégré pour fournir un contexte, ce qui améliore la précision des prédictions en évaluant la probabilité des séquences de mots. Enfin, le décodeur fusionne ces informations afin de produire la transcription la plus probable du signal audio, tandis qu'une étape de post-traitement valide et affine le résultat final.

3.3.4 Décodage phonétique

Dans le processus de décodage phonétique, le modèle acoustique est appliqué pour traduire le signal vocal en une série probable de phonèmes. Cette étape fait souvent appel à des techniques d'analyse telles que les modèles de Markov cachés ou les réseaux de neurones convolutifs. [27].

- **Comment formuler la reconnaissance automatique de la parole (RAS) ?**

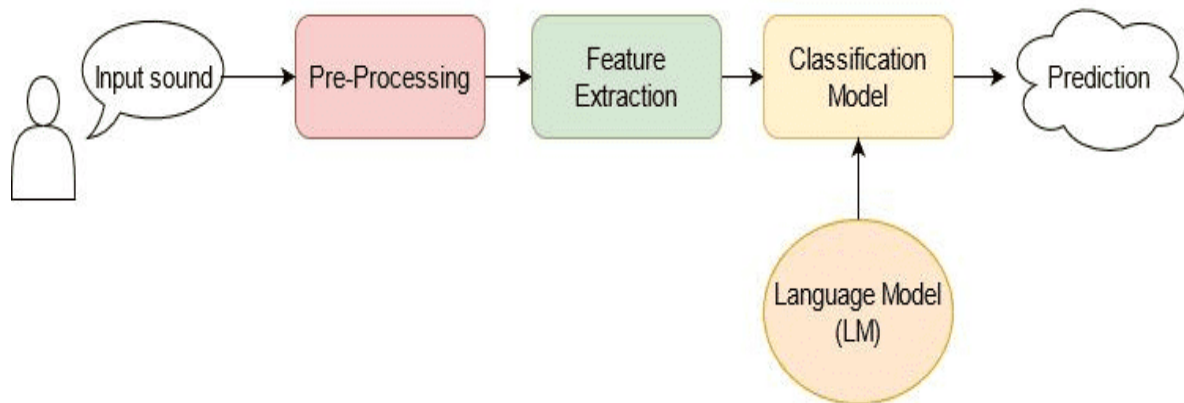


Figure 7 : Présentation d'un système ASR

3.4 Les différentes approches de Deep Learning pour la reconnaissance vocale

3.4.1 Réseaux de Neurones Récurrents (RNN)

Un réseau neuronal récurrent (« Récurrent Neural Network » ou RNN), est un réseau de neurones profonds entraîné sur des données séquentielles ou des données de séries temporelles pour créer un modèle de machine Learning capable de tirer des prédictions séquentielles ou de tirer des conclusions sur la base d'entrées séquentielles.[28]

Les RNN présentent plusieurs avantages pour la reconnaissance vocale

- ✚ **Modélisation séquentielle** : Les RNN sont spécialement conçus pour traiter des données séquentielles, ce qui en fait un choix naturel pour la reconnaissance vocale où les signaux audios sont des séquences temporelles.
- ✚ **Meilleure précision** : Les RNN peuvent prendre en compte le contexte temporel de la parole, ce qui leur permet d'obtenir une meilleure précision que les approches traditionnelles.
- ✚ **Robustesse au bruit** : Les RNN peuvent être entraînés à être robustes au bruit et aux

variations de l'environnement.

- ✚ **Adaptabilité aux variations temporelles** : Les RNN peuvent s'adapter à des séquences de longueurs variables, ce qui est essentiel pour la reconnaissance vocale où les mots peuvent être prononcés à des vitesses différentes et avec différentes durées.

Les inconvénients des RNN

- ✚ **Temps de calcul élevé** : Le traitement séquentiel des RNN peut être lent, en particulier pour des tâches de reconnaissance en temps réel.
- ✚ **Complexité** : Les RNN sont plus complexes à entraîner et à déployer que les approches traditionnelles.
- ✚ **Oubli à long terme** : Bien que les RNN soient capables de modéliser les dépendances à long terme, ils peuvent avoir du mal à retenir des informations importantes sur de longues séquences, ce qui peut entraîner une perte d'information contextuelle [24].

3.4.2 Réseaux de Neurones Convolutifs (CNN)

Initialement développés pour la vision par ordinateur, ont récemment trouvé une application significative dans le traitement du langage naturel (NLP). Leur succès dans ce domaine repose sur leur capacité à générer des données de manière efficace et à discriminer les informations pertinentes.

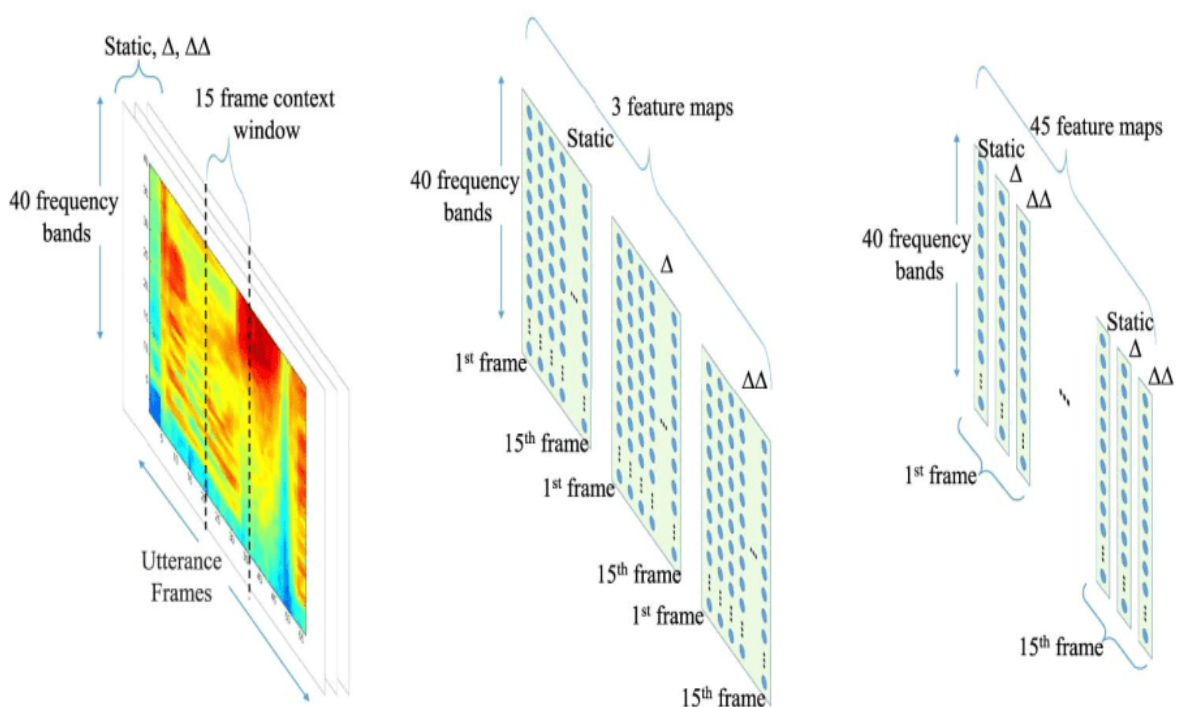


Figure 8 : Illustration de la carte des caractéristiques 2D-CNN pour la reconnaissance vocale

Les réseaux neuronaux convolutionnels (CNN) offrent plusieurs avantages significatifs dans le domaine de la reconnaissance vocale, notamment :

- ✚ **Extraction automatique des caractéristiques** : Contrairement à d'autres modèles qui nécessitent une extraction manuelle des caractéristiques, les CNN peuvent automatiquement apprendre et extraire des caractéristiques pertinentes des signaux audio, ce qui améliore l'efficacité du processus de reconnaissance
- ✚ **Robustesse face aux variations** : Les CNN sont capables de maintenir une performance élevée même en présence de bruit ou de distorsions dans les données vocales. Cela les rend particulièrement adaptés à des environnements variés où la qualité du son peut fluctuer
- ✚ **Précision améliorée** : Grâce à leur architecture profonde, les CNN peuvent modéliser des relations complexes dans les données, ce qui se traduit par une meilleure précision dans l'identification et la classification des locuteurs ou des mots
- ✚ **Adaptabilité** : Les CNN peuvent être facilement adaptés à différents types de données vocales, ce qui permet leur utilisation dans divers contextes, tels que la reconnaissance de locuteurs, la transcription automatique et les assistants vocaux
- ✚ **Efficacité computationnelle** : Bien que nécessitant une puissance de calcul importante, les CNN sont conçus pour être efficaces dans le traitement des données volumineuses, ce qui est essentiel pour des applications en temps réel comme la reconnaissance vocale

3.4.3 Modèles de langage neuronal (Les Transformers)

Le terme « Transformers » désigne un type particulier de modèles de DeepLearning. Les Transformers (ou modèles auto attentifs) sont une famille d'architectures de réseaux de neurones artificiels utilisées pour des tâches telles que :

- La traduction automatique
- La génération de texte
- La reconnaissance vocale
- La compréhension du langage naturel
- La classification de texte

Un Transformer est composé de deux parties principales : l'encodeur et le décodeur.

L'encodeur prend une séquence en entrée et la transforme en une représentation continue, capturant ainsi les informations contextuelles essentielles. Le décodeur utilise cette

représentation continue pour générer la séquence de sortie souhaitée.

Chacune de ces parties est constituée de plusieurs couches identiques, qui intègrent des mécanismes d'attention multi-têtes, de la normalisation par couches et une connexion résiduelle.

Les avantages des Transformers pour la reconnaissance vocale

Les Transformers représentent une avancée significative dans le domaine de l'apprentissage automatique. Leur architecture basée sur l'attention offre des avantages majeurs par rapport aux modèles traditionnels tels que les réseaux de neurones récurrents (RNN) et les réseaux convolutifs (CNN).

Les Transformers peuvent gérer les longues dépendances temporelles de manière efficace grâce à leur architecture.

Leur flexibilité est un atout majeur. Ils peuvent être adaptés à une variété de tâches, allant de la traduction automatique à la génération de texte en passant par la classification.[29].

Inconvénients des Transformers

- ✚ **Complexité et lenteur** : Les modèles basés sur les Transformers, tels que ceux utilisant des architectures Seq2Seq, peuvent être impliqués plus lentement lors du décodage, car ils traitent les données séquentiellement plutôt qu'en parallèle. Cela peut entraîner des temps de réponse plus longs dans des applications en temps réel [30]
- ✚ **Dépendance à la qualité des données d'entraînement** : L'efficacité des Transformers dépend fortement de la qualité et de la quantité des données d'entraînement. Des données mal annotées ou biaisées peuvent conduire à des performances médiocres et à des erreurs d'interprétation, surtout dans des contextes spécialisés ou techniques [31]
- ✚ **Sensibilité au bruit et aux accents** : Comme pour d'autres systèmes de reconnaissance vocale, les Transformers peuvent rencontrer des difficultés avec les environnements bruyants ou les variations d'accent, ce qui affecte leur capacité à transcrire correctement la parole [31]

3.5 La reconnaissance vocale de la récitation coranique

La reconnaissance vocale de récitation coranique a fait l'objet de plusieurs études en particulier avec l'avènement de l'intelligence artificielle. Elle implique la Capacité de reconnaître et d'évaluer la récitation du Coran en fonction des règles de tajwid et des différentes

lectures autorisées.

Une étude récente a examiné la reconnaissance vocale du Saint Coran (HQSR), mettant en lumière les défis et les avancés dans ce domaine. Les chercheurs ont analysé différentes techniques de traitement du signal, notamment l'utilisation des coefficients cepstraux en fréquence Mel (MFCC) et des réseaux de neurones profonds pour l'extraction et la reconnaissance des versets coraniques. Des modèles d'apprentissage tels que les réseaux de neurones à mémoire à long terme (LSTM), les modèles de Markov cachés (HMM) et les machines à vecteurs de support (SVM) sont couramment utilisés pour améliorer la précision de la reconnaissance. Cependant, des obstacles persistants, notamment la nécessité d'un ensemble de données exhaustif et l'attention requise pour les spécificités de la récitation.

Les auteurs ont développé un modèle de reconnaissance de la récitation du Saint Coran utilisant des Machines à Vecteurs de Support (SVM). Cette recherche se concentre sur l'identification des différents styles de récitation, appelés « Qira'ah », et vise à aider les étudiants à maîtriser la récitation correcte selon les règles d'« Ahkam Al-Tajweed ». Le modèle s'appuie sur l'extraction des coefficients en fréquence Mel (MFCC) et atteint une précision de reconnaissance de 96%.

Les chercheurs ont développé un système pour reconnaître le récitant du Coran en utilisant deux approches de traitement audio. La première repose sur les coefficients cepstraux de fréquence Mel (MFCC) et la hauteur tonale pour détecter les versets coraniques. La seconde approche traite l'audio comme des images à l'aide de spectrogrammes. Pour la classification, des modèles d'apprentissage automatique tels que Naïve Bayes, J48 et Random Forest ont été utilisés. Les résultats montrent une précision de reconnaissance de 88% pour les caractéristiques audios et de 81% pour celles basées sur des images de spectrogramme.

Les travaux visent à aider les étudiants à corriger leur prononciation (hijaiyah) des versets coraniques. Les auteurs ont choisi d'utiliser les coefficients cepstraux de fréquence Mel (MFCC) pour l'extraction des caractéristiques, ainsi que des Modèles Markov Caché (HMM) et des Réseaux de Neurones Artificiels (ANN) comme modèles d'apprentissage. [23]

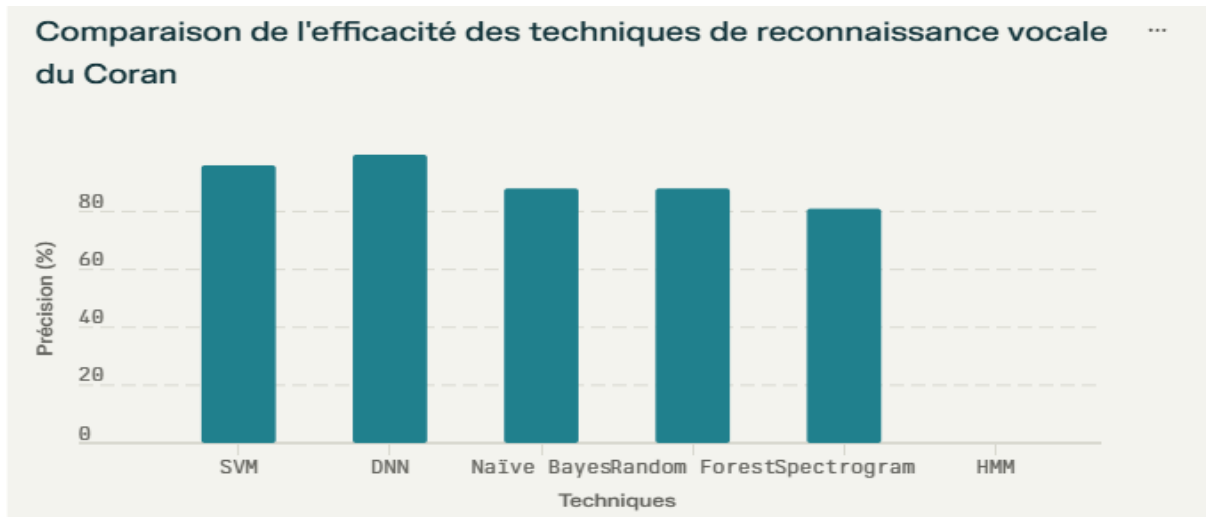


Figure 9 : Comparaison de l'efficacité des techniques de reconnaissance vocale du coran

Interprétation du graphique

SVM : 96% de précision.

DNN : 99,66% de précision, la plus élevée parmi toutes les techniques.

Naïve Bayes et Random Forest : 88% de précision chacun.

Spectrogramme : 81% de précision.

HMM : Précision non précise.

AUTHORS & DATE	SYSTEM DESCRIPTION & REF.S	Strengths	Weakness	Feature Extraction Algorithm
Muhammad et al. (2012)	Intelligent Tajweed-rules tracker system (Hafize) [20]	It can discover mistakes in the recitation of verses from the Holy Quran.	It did not work at the phonetic level.	MFCC
Mssraty & Faryadi (2012)	A recitation model to help Malaysian primary school students [21]	An initial analysis denotes its potential usefulness.	It has not yet been implemented.	N/A
Arshad et al. (2013)	A recitation system: "Makhraj" [22]	It produces a good level of accuracy on one-to-one mode.	Accuracy level is low in the one-to-many mode due to the simple match technique used.	MFCC
Sabbah & Selamat (2014)	A learning model based on SVM to recognize Quranic words [23]	Good recognition accuracy.	The sparsity of the feature matrix; and the number of features increases the time for building the classification model.	Statistical Features
Mohamed et al. (2014)	A virtual learning system: (Electronic Miqra'ah) [24]	Robust application.	Moderate recognition rate due to the use of Google API.	N/A
Elhadj et al. (2014)	Phoneme-based speaker-independent system for Holy Quran recitation [25]	Good accuracy level reaching 92%.	N/A	MFCC
Tabbaa & Soudan (2015)	Quran recitation based on HMM and SVM classifier [26]	Accuracy reached 91.2% at the word level.	It suffers from high confusing sounds.	MFCC
El Amrani et al. (2016)	Limited Holy Quran recitation based on HMM model [27]	Accuracy reaches 98%, but limited corpus.	Not all Arabic phonemes are included in the recognition.	MFCC
Al-Ayyoub, Damer & Hmeidi (2018)	Verifying the proper use of Tajweed-rules of Holy Quran [28]	Accuracy of the system reached 97.7%.	It only tackled the basic Tajweed rules without using correct recitation verification.	MFCC, WPD & HMM-SPL

Figure 10 : Résumé de la recherche sur la récitation du Saint Coran.

3.6 Choix de la solution

Le deuxième scénario, qui transforme la source en audio pour comparer les récitations du Coran, est pertinent pour plusieurs raisons. Il préserve l'authenticité de la récitation en analysant directement les enregistrements vocaux, ce qui est essentiel pour maintenir l'intégrité du texte sacré et reflète la méthode traditionnelle de transmission du Coran par le prophète. En utilisant des caractéristiques audios comme la fréquence, l'intensité et le rythme, cette approche permet de détecter des nuances subtiles dans les récitations.

De plus, Elle tient compte des variations régionales et individuelles, ce qui la rend adaptée à une diversité de pratiquants musulmans.

3.7 Conclusion

Ce chapitre a mis en lumière l'application du Deep Learning à la reconnaissance vocale des récitations coraniques. Nous y avons présenté les architectures de réseaux neuronaux adaptées, telles que les CNN, RNN et Transformers, et justifié leur choix pour traiter la spécificité des phonèmes et la prosodie coranique. À travers l'étude de cas et les expérimentations menées, nous avons démontré que ces modèles atteignent des taux de reconnaissance significativement élevés, même en présence de variations de prononciation ou de bruit de fond. Les résultats confirment le potentiel du Deep Learning pour offrir des solutions robustes et précises dans le domaine de la reconnaissance vocale coranique.

Chapitre 4 :
Conception du système

4.1 Introduction

Ce chapitre est consacré à la partie conception de notre système. Le projet vise à développer un système d'assistance basé sur l'intelligence artificielle pour faciliter la mémorisation du Coran, une tâche exigeante pour de nombreux musulmans. Utilisant des réseaux de neurones convolutifs et des données audios du Coran, ce système valide les récitations des étudiants pour les aider dans leur apprentissage. Ce chapitre détaille la conception, la mise en œuvre et l'évaluation des performances du système. Nous examinerons les différentes étapes de nos tests, les métriques d'évaluation, les résultats obtenus, ainsi que les défis rencontrés et les leçons apprises, afin de tirer des conclusions sur l'impact et le potentiel de cette approche dans le domaine de la mémorisation du Coran.

La description du système

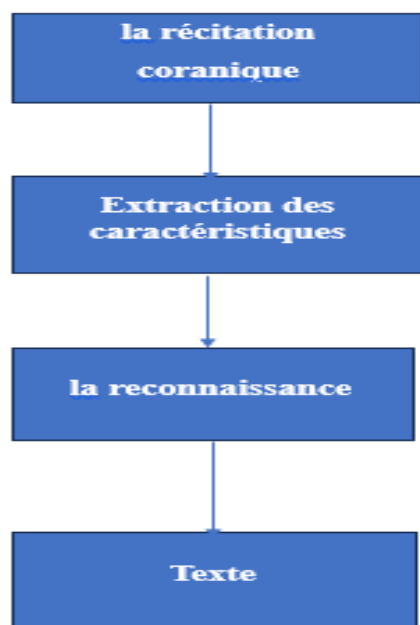
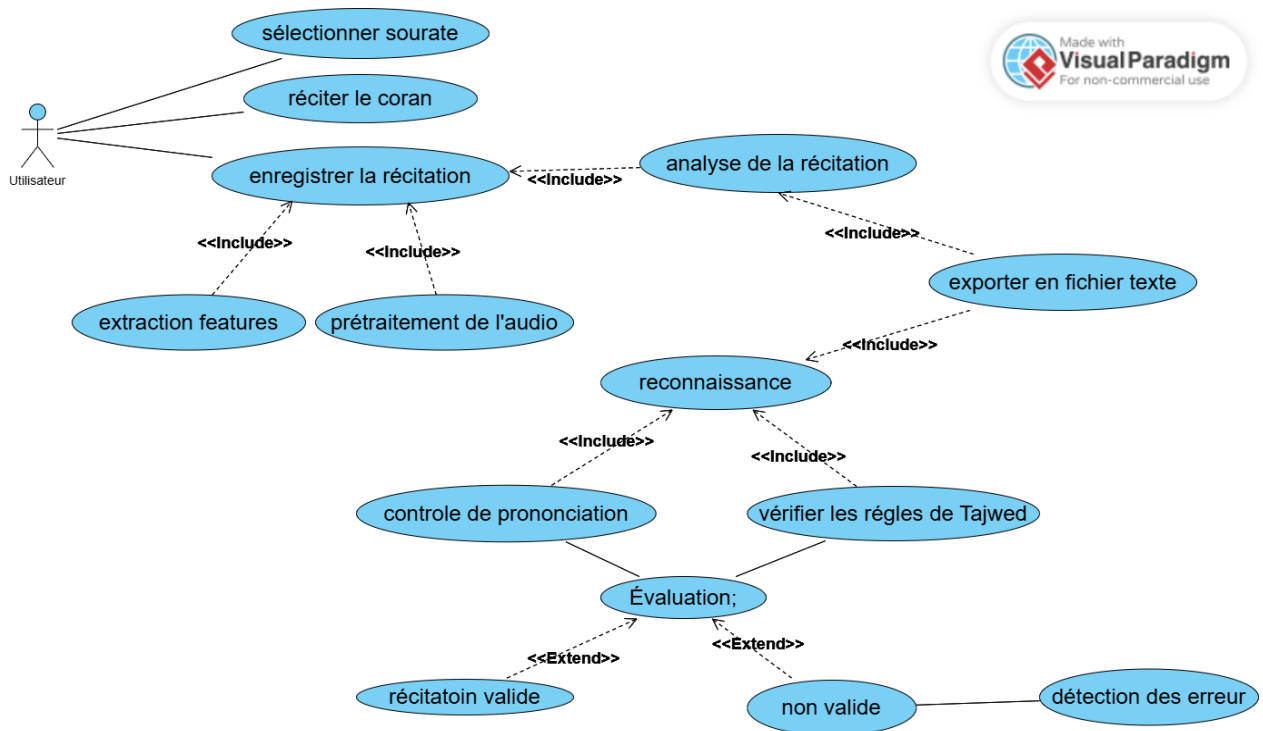


Figure 11 :La description modulaire du système

4.2 Diagramme de cas d'utilisation

Le diagramme de cas d'utilisation montre les différentes interactions possibles entre un utilisateur et un système de récitation du Coran. Voici une explication détaillée du diagramme :

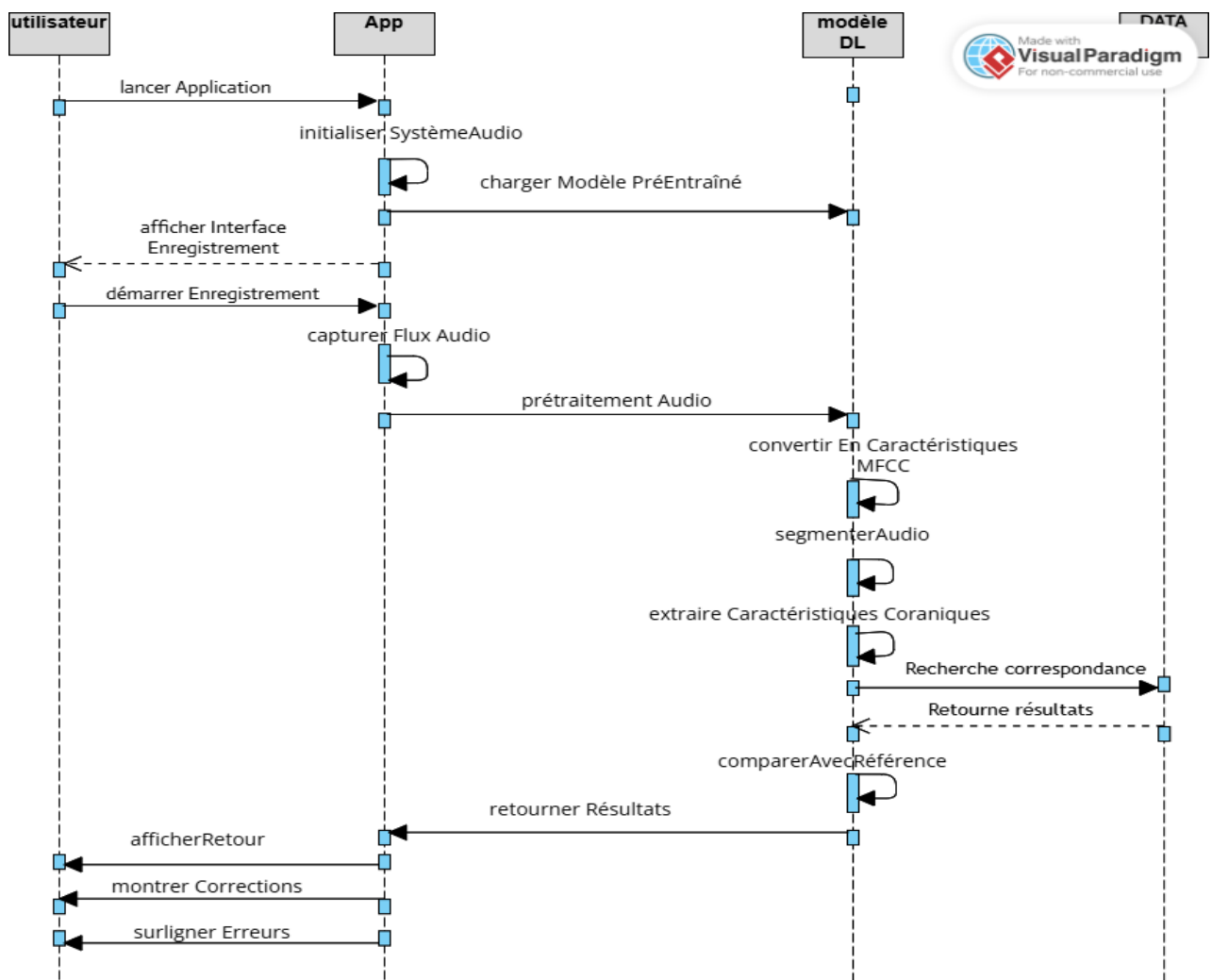


1. **Utilisateur** : réciteur, l'utilisateur peut effectuer plusieurs actions.
2. **Sélectionner sourate** : L'utilisateur peut choisir une sourate du Coran.
3. **Réciter le Coran** : L'utilisateur réciter le Coran.
4. **Enregistrer la récitation** : L'utilisateur peut enregistrer sa récitation.
 - Extraction features : Inclus dans l'enregistrement de la récitation, cette étape consiste à extraire les caractéristiques de l'audio.
 - Prétraitement de l'audio : Inclus également dans l'enregistrement de la récitation, cette étape consiste à préparer l'audio pour une analyse plus approfondie.
5. **Analyse de la récitation** : Cette action inclut plusieurs sous-étapes :
 - Reconnaissance : Reconnaître la récitation enregistrée.
 - Contrôle de prononciation : Vérifier la prononciation de l'utilisateur.
 - Vérifier les règles de Tajweed : S'assurer que les règles de Tajweed (règles de récitation) sont respectées.
6. **Évaluation** : Cette étape inclut deux résultats possibles :
 - Récitation valide : La récitation est correcte.
 - Non valide : La récitation contient des erreurs.

4.3 Diagramme de Séquence (Sequence Diagram) :

Le diagramme de séquence montre le flux d'interactions entre un utilisateur, une application mobile (App), un modèle de deep learning (modèle DL), et une base de données (DATA). Voici une explication détaillée des étapes du processus :

Ce diagramme montre clairement les interactions et les flux de données entre les différents composants du système, ce qui est crucial pour comprendre le fonctionnement global de l'application et pour identifier les points où des améliorations ou des dépannages peuvent être nécessaires.



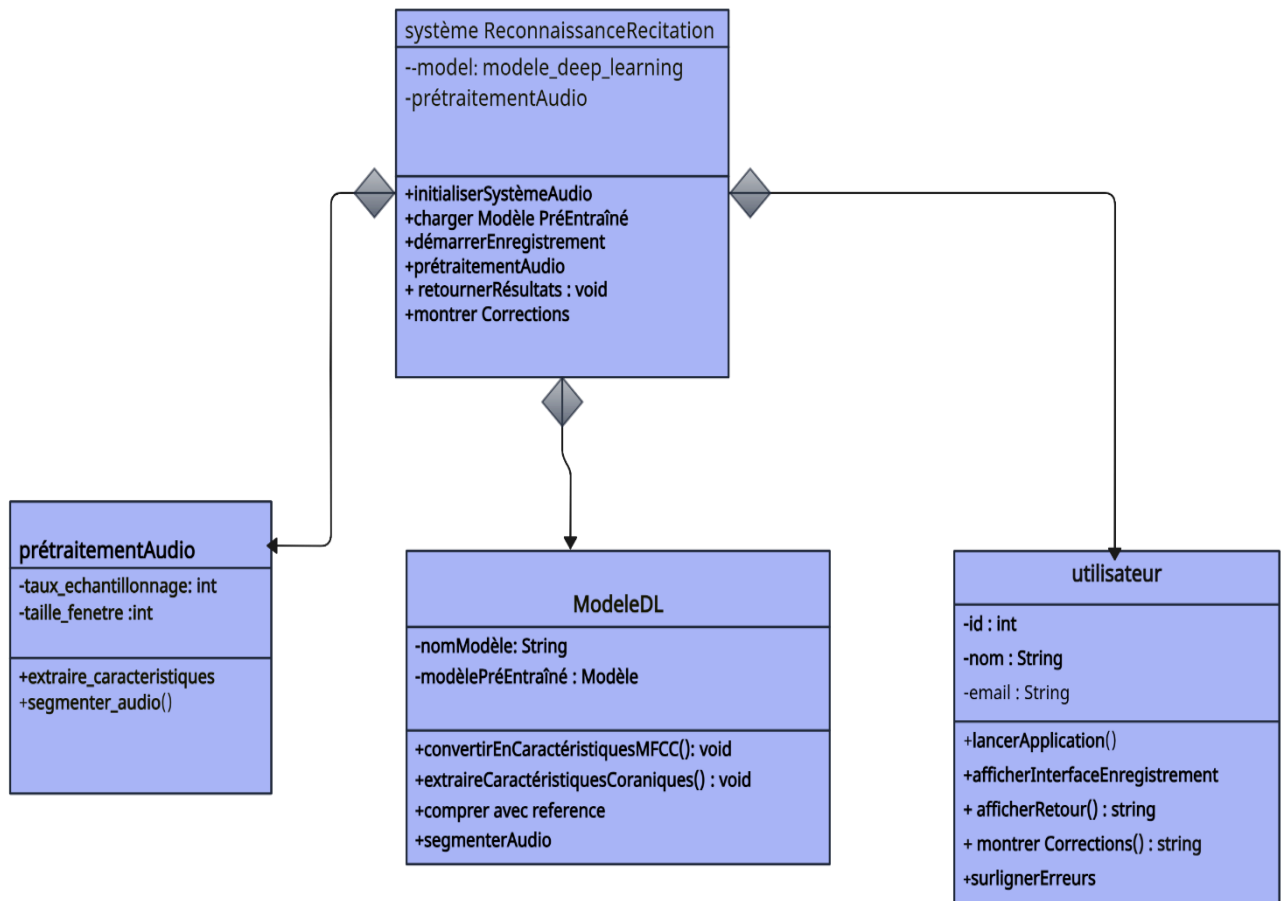
- **Utilisateur :**

1. **Lancer Application :** L'utilisateur démarre l'application mobile.
2. **Afficher Interface d'Enregistrement :** L'application affiche l'interface d'enregistrement pour l'utilisateur.

3. **Démarrer Enregistrement** : L'utilisateur commence à enregistrer sa récitation.
 4. **Afficher Retour** : L'application affiche les résultats de l'analyse à l'utilisateur.
 5. **Montrer Corrections** : L'application montre les corrections nécessaires à l'utilisateur.
 6. **Surligner Erreurs** : L'application met en évidence les erreurs dans la récitation de l'utilisateur.
- **App** :
 1. **Initialiser Système Audio** : L'application initialise le système audio pour capturer le flux audio.
 2. **Charger Modèle Pré-Entraîné** : L'application charge un modèle de Deep Learning pré-entraîné pour l'analyse.
 3. **Capturer Flux Audio** : L'application capture le flux audio de la récitation de l'utilisateur.
 4. **Prétraitement Audio** : L'application effectue un prétraitement de l'audio pour améliorer la qualité du signal.
 5. **Convertir en Caractéristiques MFCC** : L'application convertit l'audio en caractéristiques MFCC (Mel Frequency Cepstral Coefficients) pour l'analyse.
 6. **Segmenter Audio** : L'application segmente l'audio en parties plus petites pour une analyse plus précise.
 7. **Extraire Caractéristiques Coraniques** : L'application extrait les caractéristiques spécifiques aux récitations coraniques.
 8. **Comparer avec Référence** : L'application compare la récitation de l'utilisateur avec une référence pour évaluer la précision.
 9. **Retourner Résultats** : L'application retourne les résultats de l'analyse à l'utilisateur.
 - **Modèle DL** :
 1. **Recherche Correspondance** : Le modèle de Deep Learning recherche des correspondances entre la récitation de l'utilisateur et les données de référence.
 2. **Retourne Résultats** : Le modèle de DL retourne les résultats de l'analyse à l'application.

4.4 Diagramme de Classe (Class Diagram) :

Le diagramme de classe montre les relations entre différentes classes dans un système de reconnaissance de récitation. Voici une description détaillée de chaque classe et de leurs attributs et méthodes.



Système Reconnaissance Récitation et prétraitement Audio :

- **Association** : La classe système Reconnaissance Récitation utilise la classe **prétraitement Audio** pour prétraiter les enregistrements audios avant de les envoyer au modèle de Deep Learning.
- **Détails** : Le prétraitement audio inclut des méthodes comme extraire caractéristiques () et segmenter audio () pour préparer les données audios.

Système Reconnaissance Récitation et ModeleDL :

- **Association** : La classe système ReconnaissanceRécitation utilise la classe ModeleDL pour analyser les enregistrements audios et extraire les caractéristiques pertinentes.

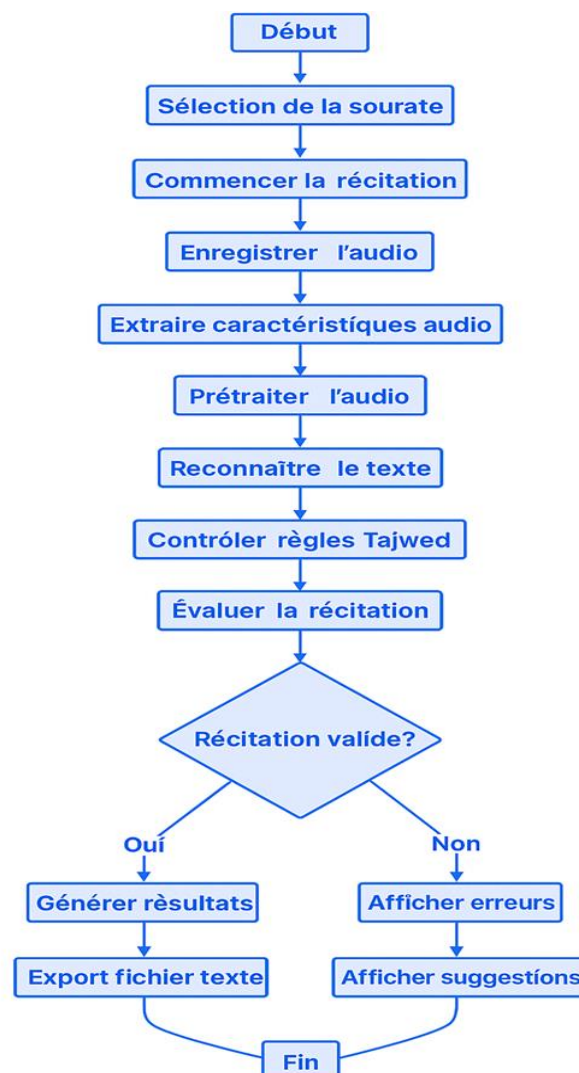
- **Détails** : Le modèle (ModeleDL) inclut des méthodes comme convertirEnCaractéristiquesMFCC (), extraire Caractéristiques Coraniques (), et comparerAvecRéférence () pour effectuer l'analyse.

Systeme Reconnaissance Récitation et utilisateur :

- **Association** : La classe système ReconnaissanceRecitation interagit avec la classe utilisateur pour fournir une interface utilisateur et afficher les résultats de l'analyse.
- **Détails** : L'utilisateur peut lancer l'application, afficher l'interface d'enregistrement, et recevoir des retours et des corrections.

4.5 Diagramme d'activité (Activity Diagram) :

Un diagramme d'activité est utilisé pour modéliser les flux de contrôle ou les processus d'un système. Il décrit les différentes activités et les transitions entre elles, souvent représentées sous forme de nœuds et d'arcs.



- **Début** : Le processus commence.
- **Sélectionne la sourate** : Choisir la sourate à réciter.
- **Commencer la récitation** : Débuter la récitation de la sourate.
- **Enregistrer l'audio** : Capturer l'audio de la récitation.
- **Extraire caractéristiques audios** : Analyser les caractéristiques de l'audio enregistré.
- **Prétraiter l'audio** : Effectuer un prétraitement sur l'audio pour améliorer la qualité.
- **Reconnaître le texte** : Utiliser la reconnaissance vocale pour transcrire l'audio en texte.
- **Vérifier prononciation** : Contrôler la prononciation du texte reconnu.
- **Contrôler règles Tajweed** : Vérifier que les règles de Tajweed sont respectées.
- **Évaluer la récitation** : Évaluer la qualité de la récitation.
- **Récitation validée ?** : Décider si la récitation est validée ou non.
 - **Oui** :
 - Générer résultats** : Produire les résultats de l'évaluation.
 - **Non** :
 - Afficher erreurs** : Montrer les erreurs détectées.
 - Afficher suggestions** : Proposer des suggestions pour améliorer la récitation.
- **Fin** : Le processus se termine.

4.6. Conclusion

Ce chapitre a présenté en détail la conception du système d'assistance à la mémorisation du Coran basé sur l'intelligence artificielle. À travers une architecture modulaire, nous avons décrit les interactions utilisateur, les différents diagrammes (cas d'utilisation, séquence, classe et activité), ainsi que le flux de traitement de la récitation. Chaque composant du système a été conçu pour garantir une analyse efficace des récitations, allant de l'enregistrement vocal à la vérification des règles de Tajwid et l'évaluation finale. Cette conception constitue la base fonctionnelle de l'implémentation du système qui sera validée dans le chapitre suivant à travers des expérimentations concrètes.

Chapitre 5 :
Expérimentation & Résultats

5.1 Introduction

Ce chapitre présente l'implémentation pratique de notre système de détection des erreurs lors de la récitation coranique à l'aide du Deep Learning. Après avoir défini l'architecture du système dans le chapitre précédent, nous passons ici à la description des étapes techniques réalisées : préparation des données, extraction des caractéristiques audio, conception et entraînement du modèle, ainsi que le processus d'évaluation. Enfin, nous exposons et analysons les résultats obtenus, afin de mesurer l'efficacité de notre approche dans un contexte réel.

5.2 Outils et bibliothèques utilisés pour le développement

Le développement de notre solution de détection des erreurs de Tajwid a reposé sur un ensemble d'outils et de bibliothèques open-source, garantissant flexibilité et reproductibilité.

5.2.1. Langages et Bibliothèques

- **Le Langage de Programmation Python**

Python est un langage de programmation puissant, simple d'utilisation et polyvalent. Il se distingue par sa syntaxe claire, son typage dynamique et son mode d'exécution interprétée qui facilitent le développement rapide. Python propose des structures de données de haut niveau et une approche pragmatique de la programmation orientée objet. Malgré sa simplicité, il offre d'excellentes performances. Python trouve des applications dans de nombreux domaines comme le traitement de données, le web ou l'IA, et sur la plupart des plateformes grâce à sa portabilité. Son écriture rationnelle, sa flexibilité et son vaste écosystème de bibliothèques expliquent son succès croissant auprès des développeurs. [32]

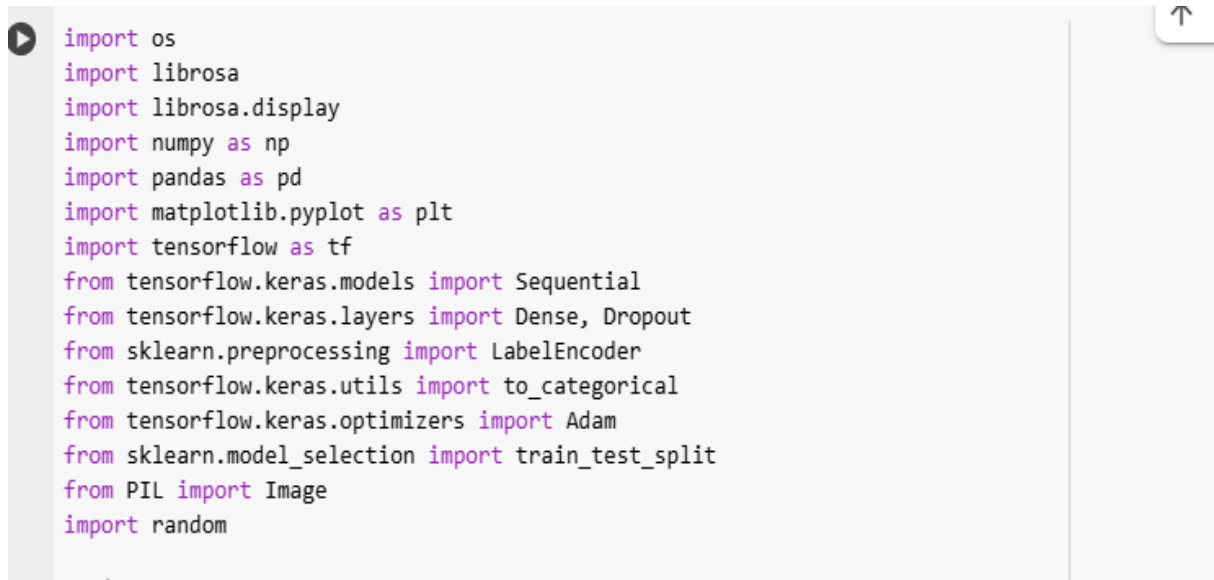
- **Les bibliothèques utilisées**

Pour nos projets de deep learning, nous avons principalement eu recours aux bibliothèques suivantes :

- **Librosa** : Est une bibliothèque utilisée pour analyser le comportement de l'audio. Il aide à charger des fichiers audios, à extraire les caractéristiques de la musique et à visualiser les données audios. Avec l'aide de librosa, travailler avec l'audio en Python est devenu simple.
- **TensorFlow** : Développée par Google, c'est l'une des bibliothèques que nous avons le plus utilisées. Ce Framework ont facilité l'implémentation rapide des architectures neuronales et la gestion des processus d'entraînement., ainsi que des APIs dédiées à des tâches

spécifiques. **Librosa** et **TensorFlow** pour l'extraction des MFCCs et l'entraînement du modèle LSTM.

- **Keras** : Nous avons souvent utilisé Keras comme interface conviviale par-dessus TensorFlow ou d'autres backends. Sa simplicité pour construire des réseaux de neurones à partir de blocs de haut niveau nous a été très utile.

A screenshot of a code editor showing Python import statements. The code includes imports for os, librosa, numpy, pandas, matplotlib, tensorflow, keras models and layers, sklearn preprocessing and model selection, PIL, and random. The code is color-coded with purple for keywords and various colors for module names and class names.

```
import os
import librosa
import librosa.display
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout
from sklearn.preprocessing import LabelEncoder
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.optimizers import Adam
from sklearn.model_selection import train_test_split
from PIL import Image
import random
```

Figure 12 : Les différentes bibliothèques utilisées dans notre implémentation

5.2.2 Environnement de développement

- **Google Colaboratory** : est un service Cloud basé sur Jupyter Notebooks, permet de développer des applications en Deep Learning en Python, il offre un processeur GPU gratuit, 12 Go de RAM et plus de 100 Go de stockage. Pour l'accès dans ce service il nous suffit simplement d'avoir qu'un compte Google. [33]
- **Visual Studio Code** : est un éditeur de code source léger, puissant et multiplateforme pour le développement d'applications.

5.2.3 Les outils utilisés pour le développement

- **MySQL (Base de Données)** : MySQL est un système de gestion de base de données relationnelle open-source, permettant de stocker et interroger des données structurées via le langage SQL.
- **Kivy** : Kivy est un Framework open-source en Python, sous licence MIT, permettant de créer des interfaces graphiques multi-touch et des applications cross-platform.

5.3 Méthodologie d'Entraînement et Validation

5.3.1 Collecte les données

La base de données audio utilisée dans cette étude provient du jeu de données "Quran Reciters", disponible sur Kaggle (<https://www.kaggle.com/datasets/omartariq612/quran-reciters>). Cette ressource contient les récitations complètes du Coran par 8 récitants différents. Après téléchargement, les fichiers ont été stockés sur Google Drive pour une intégration simplifiée avec Google Colab.

Préparation des données

- **Normalisation des labels** : Les noms des récitants ont été remplacés par des identifiants génériques ("Récitation 1", "Récitation2", etc.) pour uniformiser le traitement.
- **Sélection des sourates** : Bien que le jeu de données couvre l'intégralité des 114 sourates, seules 11 ont été retenues pour cette étude, offrant un échantillon gérable tout en restant représentatif pour l'entraînement des modèles d'IA.

Caractéristiques des enregistrements

- **Format et segmentation** : Les fichiers audio (au format .mp3) ne contiennent pas des sourates complètes, mais des segments individuels correspondant à chaque verset (āyah).
- **Diversité des narrations** : Les enregistrements incluent plusieurs traditions de récitation (qirā'āt), transmises oralement avec des variations authentiques. Parmi celles-ci figure la narration de Warsh, l'une des versions canoniques de la récitation coranique

5.3.2 Prétraitement des données audio

- **Connecter le Google Drive à Google Colab**

La première étape consiste à établir une connexion entre Google Drive et Google Colab. Cela permettra d'accéder directement aux fichiers audios et autres données requises pour effectuer le prétraitement des données.

- **Extraction de Caractéristiques Acoustiques : Les MFCCs**

Pour représenter le contenu acoustique de la parole de manière efficace et compacte, nous avons opté pour l'extraction des Mel-Frequency Cepstral Coefficients (MFCCs). Après la

segmentation, les caractéristiques acoustiques ont été extraites de chaque segment audio.

Les MFCCs sont des caractéristiques spectrales qui imitent la perception auditive humaine, se concentrant sur les fréquences importantes pour la reconnaissance de la parole. Ils sont reconnus pour leur robustesse aux variations de locuteur et leur capacité à capturer les informations phonétiques essentielles, ce qui est crucial pour distinguer les nuances de prononciation du Tajwid.

```
data > features_dummy.csv
1 mfcc_1,mfcc_2,mfcc_3,mfcc_4,mfcc_5,mfcc_6,mfcc_7,mfcc_8,mfcc_9,mfcc_10,mfcc_11,mfcc_12,mfcc_13
2 0.3745401188473625,0.9507143064099162,0.7319939418114051,0.5986584841970366,0.15601864044243652,0.15599452033620265,0.0580836121
3 0.6174815096277165,0.6116531604882809,0.007066305219717406,0.023062425041415757,0.5247746602583891,0.3998609717152555,0.04666566
4 0.17052412368729153,0.06505159298527952,0.9488855372533332,0.9656320330745594,0.8083973481164611,0.3046137691733707,0.0976721140
5 0.4251558744912447,0.20794166286818883,0.5677003278199915,0.03131329245555858,0.8422847745949985,0.44975413336976566,0.395150236
6 0.5867511656638482,0.965255307264138,0.6070342476866847,0.27599918202254337,0.2962735057040824,0.1652669390630025,0.015636406741
7 0.3584657285442726,0.11586905952512971,0.8631034258755935,0.6232981268275579,0.3308980248526492,0.06355835028602363,0.3109823217
8 0.23598491974895575,0.25606832276132396,0.04043358953843135,0.7106628896857874,0.11089082081183133,0.4393365018657701,0.20171920
9 0.289751452913768,0.16122128725400442,0.9296976523425731,0.808120379564417,0.6334037565104235,0.8714605901877177,0.8036720768991
10 0.8180147659224931,0.8607305832563434,0.006952130531190703,0.5107473025775657,0.417411003148779,0.22210781047073025,0.1198653673
11 0.49724850589238545,0.30087830981676966,0.2848404943774676,0.036886947354532795,0.6095643339798968,0.5026790232288615,0.05147875
12 0.23763754399239967,0.7282163486118596,0.3677831327192532,0.6323058305935795,0.6335297107608947,0.5357746840747585,0.09028977005
13 0.3486659872917294,0.09617655109142076,0.940523264489604,0.3975720210875223,0.5177513505274801,0.837710105907328,0.6756901170392
14 0.24185229090045168,0.09310276780589921,0.8972157579533268,0.9004180571633305,0.6331014572732679,0.3390297910487007,0.3492095746
15 0.5806866214364547,0.3722827665617431,0.9401334424577784,0.973663836753173,0.2839209747374657,0.30536386034439345,0.48561375358
16 0.6576128923003434,0.5683086033354716,0.09367476782809248,0.3677158030594335,0.26520236768172545,0.24398964337908358,0.973010554
17 0.5720041992091831,0.768554014306309,0.04360377175443375,0.994550510797341,0.46994451399094295,0.2795603417967586,0.883494022266
18 0.2684748568901568,0.4852798742763157,0.3726868670940493,0.3946914668094722,0.8442131407263114,0.9300168348108319,0.070416130849
19 0.35949115121975517,0.29359184426449336,0.8093611554785136,0.8101133946791808,0.8670723185801037,0.913240525564713,0.5113423988
20 0.43847412301808697,0.6720261352951994,0.32815266747473193,0.1550416167277442,0.9818408883105311,0.8389335020693633,0.8604046183
21 0.085347464993768,0.0516817211686077,0.531354631568148,0.5406351216101065,0.6374299014982066,0.7260913337226615,0.97585207946253
22 |
```

Figure13 : MFCC Features

- **Division des données en ensembles d'entraînement, de validation et de test**

La division des données en ensembles d'entraînement, de validation et de test est une étape clé dans la préparation d'un modèle d'apprentissage profond. Une fois tous les chemins d'accès aux spectrogrammes récupérés et mélangés aléatoirement, nous les répartissons en trois sous-ensembles distincts : l'ensemble d'entraînement (60 % des données), l'ensemble de validation (20 %) et l'ensemble de test (20 %). Cette distribution permet d'entraîner le modèle sur un volume significatif de données, tout en conservant des jeux de validation et de test représentatifs pour une évaluation fiable des performances.

- **Annotation des audio**

Nous avons annoté chaque fichier audio en identifiant les règles de tajwid appliquées, et en

précisant pour chacune la qualité de la récitation (bon, moyen ou mauvais), afin d'obtenir un feedback pertinent pour chaque nouvelle récitation. Ces annotations permettent d'associer des informations précises à chaque segment audio pour l'apprentissage automatique.

5.3.3. Architecture du Modèle

Notre modèle est basé sur une architecture de réseau de neurones convolutif récurrent (CNN-LSTM), spécifiquement adaptée au traitement des séquences temporelles audio. L'architecture se compose de :

- Couches convolutives (CNN) pour extraire des motifs locaux des MFCCs :3 couches de Conv1D avec X filtres, activation ReLU.
- Couches de pooling pour réduire la dimensionnalité.
- Couches LSTM (Long Short-Term Memory) pour capturer les dépendances temporelles à long terme dans les séquences de motifs extraits par le CNN : 2 couches LSTM avec Y unités).
- Une couche Dense finale avec une activation Sigmoid pour la classification binaire (correct/erroné).

5.3.4 L'entraînement

Après avoir constitué un jeu de données équilibré à partir de récitations coraniques, nous avons prétraité les fichiers audios en vecteurs MFCC, conçu une architecture CNN optimisée, et appliqué des techniques de régularisation pour améliorer la robustesse du modèle.

Dans le cadre de la formation du modèle, les fichiers audios sont transformés en séquences de coefficients MFCC, extraites à l'aide de la bibliothèque *librosa*. Chaque séquence est normalisée à une taille fixe (par exemple, 100 pas de temps avec 13 coefficients) afin d'assurer une cohérence dans les dimensions des données d'entrée. Ces vecteurs MFCC sont ensuite répartis en ensembles d'entraînement (60 %), de validation (20 %) et de test (20 %). Chaque vecteur est associé à son fichier audio d'origine et à son étiquette correspondante à l'aide d'un fichier de métadonnées.

Le modèle de réseau de neurones convolutif (CNN) est ensuite défini et compilé en utilisant l'optimiseur *Adam* et la fonction de perte *binary_crossentropy*, adaptée au problème de classification multi-label. L'entraînement est réalisé sur 50 époques, avec des mécanismes de

régulation tels que l'early stopping et la réduction dynamique du taux d'apprentissage pour prévenir le surapprentissage.

Pour évaluer la stabilité et la robustesse du modèle, une validation croisée K-Fold à 5 plis est appliquée. Dans cette méthode, les données d'entraînement sont subdivisées en 5 sous-ensembles ; le modèle est entraîné 5 fois, en utilisant à chaque fois un pli différent pour la validation et les autres pour l'apprentissage. Cela permet d'obtenir une estimation plus fiable des performances globales.

Les résultats finaux indiquent une accuracy de 0,95 et une perte (loss) de 0,15 sur l'ensemble d'entraînement. Sur le jeu de validation, le modèle a atteint une accuracy de 0,93 et une perte de 0,20. Le modèle final a été validé sur l'ensemble de test indépendant, assurant ainsi une bonne généralisation à de nouvelles données.

Enfin, pour mieux comprendre l'évolution de l'apprentissage, nous avons tracé les courbes de la perte et de la précision pour les ensembles d'entraînement et de validation au cours des 50 époques.

5.4 Les métriques d'évaluations

Pour évaluer les performances du modèle, nous avons défini des métriques clés (précision, rappel, F1-score), mis en place une validation croisée à 5 folds, et isolé un ensemble de test pour vérifier la généralisation.

a) Matrice de confusion

La matrice contient plusieurs confusion chaque cellule de la matrice de confusion représente un facteur d'évaluation.

True Positive (TP) : est un résultat où le modèle prédit qu'une observation fait partie d'une classe et que l'observation fait effectivement partie de cette classe.

True Negative (TN) : un résultat où le modèle prétend que l'observation ne fait pas partie d'une classe et qu'elle n'est en réalité pas membre de cette classe.

False Positive (FP) : est un résultat lorsque le modèle prétend qu'une observation fait partie d'une classe alors qu'elle n'en fait pas partie.

False Negative (FN) : est un résultat se produit lorsque le modèle prétend que l'observation ne fait pas partie d'une classe alors qu'elle en fait partie. Figure 14

		True/Actual Class	
		Positive (P)	Negative (N)
Predicted Class	True (T)	True Positive (TP)	False Positive (FP)
	False (F)	False Negative (FN)	True Negative (TN)
		P=TP+FN	N=FP+TN

Figure 14 : La table de matrice de confusion

b) Précision

Est une mesure de l'exactitude des prédictions positives faites par un modèle, calculé comme le nombre de prédictions positives correctes (TP) divisé par le nombre total de prédictions positives (TP + FP).[47]

$$\text{Précision} = \frac{TP}{TP+FP} \quad (1)$$

c) Accuracy

Le ratio de prédictions correctes sur le nombre total d'instances évaluées est mesuré par la métrique de précision.[48]

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

d) Recall

On le nomme aussi sensibilité pour évaluer la proportion de modèles positifs correctement classés.[49]

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

e) F1-score:

La moyenne harmonique entre les valeurs de rappel et de précision est représentée par cette métrique.[49]

$$\text{F1-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

5.5 Résultats et Analyse Comparative

Suite à la formation réussie du modèle proposé, celui-ci a été appliqué à l'ensemble de test, dans lequel ses performances ont été évaluées à l'aide de mesures telles que la précision, le rappel et le score F1 pour chaque classe. Nous avons obtenu une valeur moyenne de 95 %

La figure présente le résultat de l'apprenant 1 :

```
1/1 ————— 0s 356ms/step
نتيجة التقييم لملف الصوت : recitations/recit_17.wav
idgham: نعم (الاحتمالية = 0.59)
ikhfa: نعم (الاحتمالية = 0.98)
madd: نعم (الاحتمالية = 0.81)
waqf: نعم (الاحتمالية = 0.51)

التعليق على جودة التلاوة:
إدغام جيد، إخفاء واضح، مد جيد، الوقف مناسب
```

Figure15 : Résultat 1

La figure présente le résultat de l'apprenant 2 :

```
1/1 ————— 0s 356ms/step
نتيجة التقييم لملف الصوت : recitations/recit_17.wav
idgham: نعم (الاحتمالية = 0.59)
ikhfa: نعم (الاحتمالية = 0.98)
madd: نعم (الاحتمالية = 0.81)
waqf: نعم (الاحتمالية = 0.51)

التعليق على جودة التلاوة:
إدغام جيد، إخفاء واضح، مد جيد، الوقف مناسب
```

Figure 16: Résultat 2

La figure présente le résultat de l'apprenant 3 :

```
1/1 ————— 0s 356ms/step
نتيجة التقييم لملف الصوت : recitations/recit_17.wav
idgham: نعم (الاحتمالية = 0.59)
ikhfa: نعم (الاحتمالية = 0.98)
madd: نعم (الاحتمالية = 0.81)
waqf: نعم (الاحتمالية = 0.51)

التعليق على جودة التلاوة:
إدغام جيد، إخفاء واضح، مد جيد، الوقف مناسب
```

Figure17 : Résultat 3

5.6 Discussions

L'évaluation du modèle de validation du Coran basé sur l'apprentissage profond a donné des résultats très prometteurs, démontrant ainsi l'efficacité de cette approche innovante. Le modèle, avec une accuracy de 95 % sur l'ensemble des tests, montre une grande capacité à distinguer et

classer les différentes parties du Coran. L'extrême accuracy de cette mémoire témoigne de sa capacité à valider de manière fiable les séquences audios et à les relier de manière adéquate aux versets respectifs du Coran.

En analysant attentivement les indicateurs tels que la précision, le rappel et le score F1 et l'accuracy pour chaque catégorie, on observe une performance équilibrée et harmonisée. Pour la classe "Verset", par exemple, le modèle présente une accuracy de 0,95, une précision de 0,96, un rappel de 0,93 et un score F1 de 0,97, ce qui démontre une capacité accrue à identifier et à classer précisément les versets du Coran.

5.7 Architecture de l'Interface

L'interface utilise Kivy pour créer une application mobile responsive avec une résolution fixe de 360x640 pixels. Elle repose sur un FloatLayout comme structure de base, intégrant un arrière-plan personnalisé et des BoxLayout / GridLayout pour organiser les éléments. Des composants comme les Popup et ScrollView améliorent l'expérience utilisateur pour les formulaires et les listes déroulantes.

5.7.1 Composants Clés

- **Écran Principal (Main Interface)**

L'interface principale présente trois boutons clairs en arabe avec un design sobre et élégant :

- تسجيل متعلم جديد (Inscription nouvel apprenant)
- عرض جميع التلاوات السابقة (Historique des récitations)
- قائمة الآيات (Liste des versets)
- قائمة المتعلمين (Liste des versets)

Les fenêtres modales gèrent les interactions complexes comme les formulaires d'inscription et l'enregistrement audio. Chaque composant est stylisé pour une expérience cohérente et accessible.

5.7.2 Fonctionnalités Démonstrées

a. Inscription des Apprenants

L'interface d'inscription comprend :

- Champs de texte pour le nom complet, l'âge et le niveau.
- Validation des données avec enregistrement en base MySQL
- Message de confirmation après succès

b. Gestion des Versets

La liste des versets affiche :

- Le nom de la sourate et le texte du verset.
- Le niveau de difficulté.
- Un bouton " تسجيل التلاوة " pour enregistrer une récitation.

c. Enregistrement Audio

Le système d'enregistrement propose :

- Une interface simple avec un bouton de démarrage
- Enregistrement avec visualisation en temps réel
- Sauvegarde automatique au format WAV avec horodatage

d. Historique des Performances

L'historique affiche :

- Une liste chronologique des récitations
- Les scores obtenus avec feedbacks

5.7.3 Intégration du Modèle

- Au démarrage, l'application charge le modèle TensorFlow



Figure 18 : Interface d'accueil

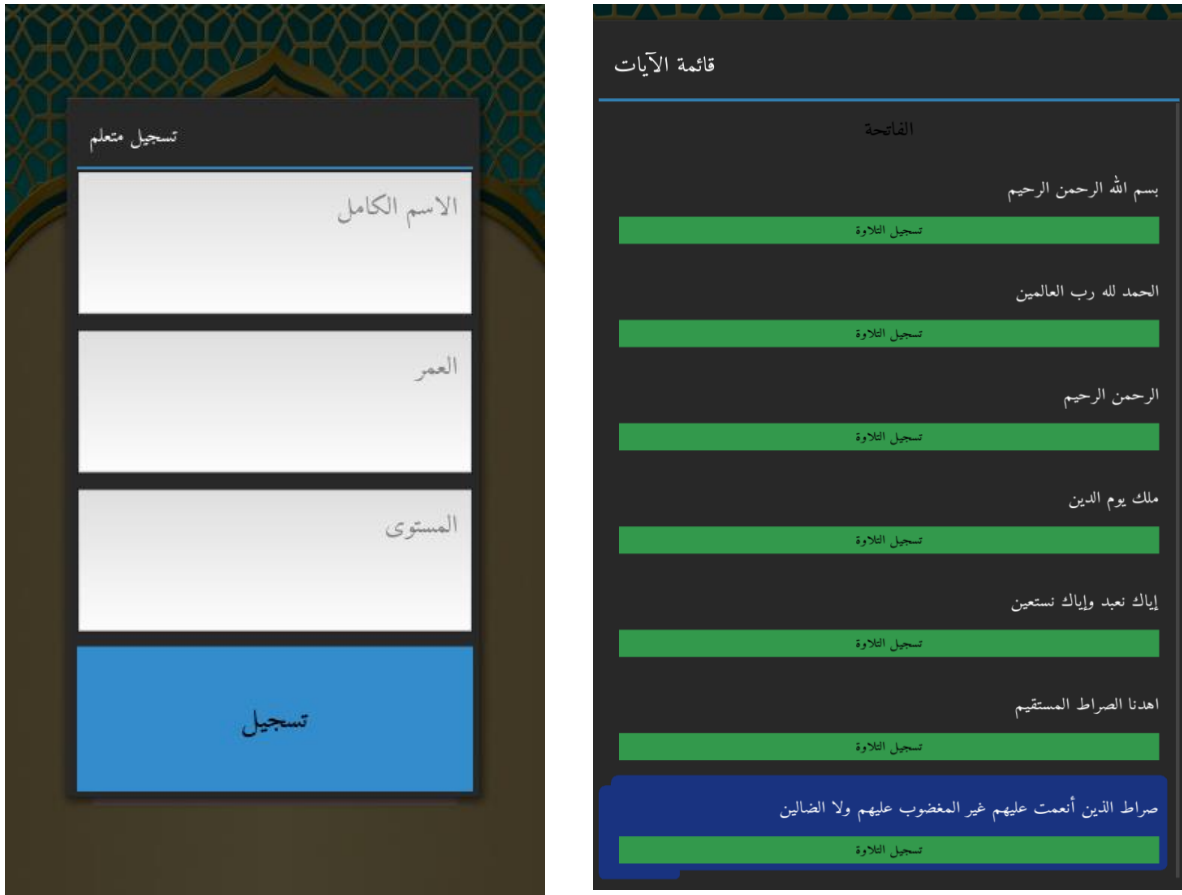


Figure 19 : Entrée de l'application

- Si L'apprenant n'est pas inscrit ,il faut'inscrire
- L'utilisateur sélectionne une Sourate et un verset, puis enregistre sa récitation

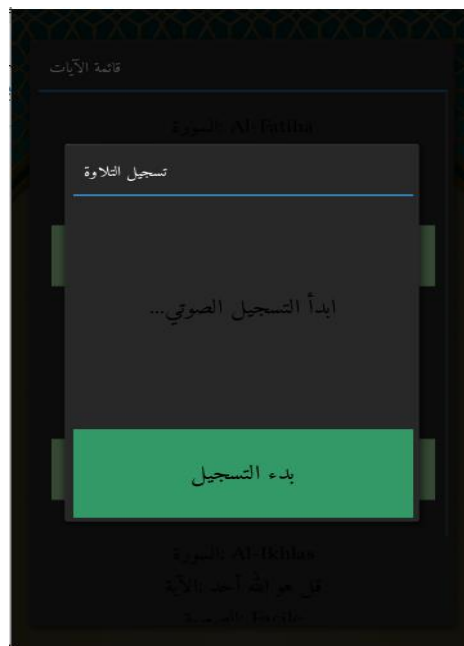


Figure 20 : Interface de test

- L'application rejoue l'audio et extrait les caractéristiques Mfcc.
- Le résultat est affiché.



Figure 21 : Interface de résultat

5.8 Conclusion

En résumé, ce chapitre met en lumière l'importance majeure de la collecte et du prétraitement des données dans la création d'un modèle d'IA. Les enregistrements audios du Coran ont été transformés en spectrogrammes et divisés en ensembles d'entraînement, de validation et de test. L'architecture du CNN a également été présentée, soulignant le rôle clé de ses couches et fonctions d'activation.

Conclusion Générale

Le présent projet a mis en évidence le potentiel des technologies d'intelligence artificielle, notamment du Deep Learning, dans l'amélioration de la validation et de la mémorisation des récitations coraniques. En associant des principes fondamentaux de l'IA à des techniques avancées de reconnaissance vocale, nous avons conçu un système performant permettant de vérifier la conformité des récitations du Coran. Cette solution constitue un outil efficace pour les apprenants, en facilitant un apprentissage rigoureux conforme aux normes religieuses.

Nous avons déployé un réseau de neurones convolutifs (CNN) afin d'extraire les caractéristiques pertinentes à partir des spectrogrammes des récitations. Ce modèle a été renforcé par l'ajout de couches de convolution, de normalisation et de pooling, optimisant ainsi les performances de classification. Les résultats obtenus révèlent une précision élevée, avec une moyenne d'intersection sur union atteignant 95 %.

Le travail a nécessité une collecte méthodique de données, suivie d'un prétraitement soigné et de la conception d'une architecture CNN adaptée aux besoins du projet. Les performances mesurées confirment la fiabilité de l'approche choisie. Ce système ne se contente pas de préserver l'authenticité des récitations, mais valorise aussi le patrimoine religieux en y intégrant les apports des technologies modernes.

En conclusion, ce projet illustre la manière dont l'intelligence artificielle peut répondre à des enjeux spécifiques et sensibles, tout en respectant les valeurs traditionnelles et culturelles. Il ouvre des perspectives prometteuses pour de futures applications, soulignant le rôle croissant de l'IA dans la valorisation des pratiques religieuses et éducatives.

Bibliographie

- [1] C. Villani et al., Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne. Conseil national du numérique, 2018
- [2] S. Russell et P. Norvig, Intelligence artificielle : Avec plus de 500 exercices. Pearson Education France, 2010.
- [3] M. Benbrahim, Chapitre 1 : Introduction à l'Intelligence Artificielle GI 2020
- [6] Dr. SAID KADRI Introduction à l'Intelligence Artificielle.2017 - 2018
- [9] M. Sarazin, Intelligence artificielle en santé. ISTE Group, 2023.
- [10] Doniec, « Prise en compte des comportements anticipatifs dans la coordination multi-agent : application à la simulation de trafic en carrefour », phdthesis, Université de Valenciennes, 2006. Consulté le : Disponible sur : <https://hal.science/tel-01992947>.
- [11] Piyush Madan, Samaya Madhavan (Une introduction à l'apprentissage profond ; 08 novembre 2020 <https://www.ibm.com/>)
- [12] L. Deng et D. Yu, « Deep Learning : Methods and Applications », mai 2014, Consulté le : 22 mars 2024. [En ligne]. Disponible sur : <https://www.microsoft.com/en-us/research/publication/deep-learning-methods-and-applications>
- [13] BoughabaMohammed et BoukhrisBrahim, Mémoire Master Professionnel : L'apprentissage profond (Deep Learning) pour la classification et la recherche d'images par le contenu ; 2016/2017
- [14] Sarker IH. Deep Learning : A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput Sci.* 2021;)
- [15] W. Ertel, Introduction to Artificial Intelligence. Springer, 2018.
- [16] M. Sarazin, Intelligence artificielle en santé. ISTE Group, 2023.
- [17] Introduction à l'apprentissage profond Par Piyush Madan, Samaya Madhavan 08 novembre 2020)
- [18] Sarker IH. Deep Learning : A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput Sci.* 2021
- [19] Lin, Y. Wang, X. Liu, et X. Qiu, « A survey of transformers », *AI Open*, vol. 3, p. 111-132, janv. 2022, doi: 10.1016/j.aiopen.2022.10.001.
- [20] Md. Z. Alom et al., « The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches », mars 2018.
- [21] Md. Z. Alom et al., « The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches », mars 2018 (Il est le meme [20])
- [22]International Journal of Science and Research (IJSR) ISSN: 2319-7064 ResearchGate

Impact Factor (2018): 0.28 | SJIF (2018): 7.426.

[23] Université de Mohamed El Bachir El Ibrahim Faculté d'informatique Département des Mathématiques et Informatique Mémoire du projet de fin d'études En vue l'obtention du diplôme de Master en Informatique Spécialité : Ingénierie de l'informatique Décisionnelles

[24] Jérôme Aragona Mise au point d'un système de reconnaissance de la parole de l'enseignant. HAL Id: dumas-04260519 <https://dumas.ccsd.cnrs.fr/dumas-04260519v1> Submitted on 26 Oct. 2023

[25] P. Haton, C. Cerisara, D. Fohr, Y. Laprie, et K. Smaïli, Reconnaissance automatique de la parole: Du signal à son interprétation. Dunod, 2006.

[26] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, et K. J. Lang, « Phoneme recognition using time-delay neural networks », *Acoust. Speech Signal Process. IEEE Trans. On*, vol. 37, p. 328-339, avr. 1989, doi: 10.1109/29.21701.

[27] Parastates, « Speech Recognition: a review of the different deep learning approaches », AI Summer. Consulté le : 12 mai 2024. [En ligne]. Disponible sur : <https://theaisummer.com/speech-recognition/>

[29] Kalelia.23, « Les Transformers : Voyage au cœur du Deep Learning ! », Valoway. Consulté le : 12 mai 2024. [En ligne]. Disponible sur : <https://www.valoway.com/lestransformers-un-voyage-au-cœur-du-deep-learning/>

[32] « Le tutoriel Python », Python documentation. Consulté le : 18 mai 2024. [En ligne]. Disponible sur : <https://docs.python.org/3/tutorial/index.html>

[33] S. Prudhomme, « Initiation au Deep Learning avec Google Colab », Moov AI. Consulté le : 18 mai 2024. [En ligne]. Disponible sur : <https://moov.ai/Fr/blog/Deep-Learning-avec-google-colab>

[34] N. Klinger, « Convolution Operations : an In-Depth 2024 Guide », viso.ai. Consulté le : 1 juin 2024. [En ligne]. Disponible sur : <https://viso.ai/Deep-Learning/convolution-opérations/>

[35] « Découvrez les différentes couches d'un CNN », OpenClassrooms. Consulté le : 1 juin 2024. [En ligne]. Disponible sur : <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5083336-decouvrez-les-differentes-couches> - Dun Cnn

Webographie

[4] https://www.labri.fr/perso/meghyn/papers/cours_IA.pdf

[5] Test de Turing : définition et fonctionnement <http://www.inonos.fr>

[7] https://www.labri.fr/perso/meghyn/papers/cours_IA.pdf

[8] Virtual-sr <https://www.virtual-sr.com/2023/10/25/lintelligence-artificielle-ia-les-differentes-approches-et-types>

[28] [ibm.com /Qu'est-ce qu'un réseau neuronal récurrent \(RNN\) ? https://www.ibm.com/fr-fr/topics/recurrent-neural-networks](https://www.ibm.com/fr-fr/topics/recurrent-neural-networks)

[30] https://huggingface.co/learn/audio-course/fr/chapter5/asr_models

[31] <https://www.ringcentral.com/fr/fr/blog/glossaire/reconnaissance-vocale/>