

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Master

Filière : Électronique

Spécialité : Electronique des systèmes embarqués

Présenté par

Berrami Merouane Fares

un système intelligent embarqué pour le suivi psychologique à distance

Proposé par : Nasr Eddine Khorissi

Année Universitaire 2024-2025

Remerciements

Je tiens à exprimer ma sincère gratitude à toutes les personnes qui ont contribué à la réalisation de ce mémoire de fin d'étude.

Tout d'abord, je ne remercierai jamais ma mère mon héroïne pour son soutien inconditionnel, ses sacrifices et sa patience tout au long de mes études.

Je remercie également mon frère, ma sœur, mes chers grand parents et mes tantes pour leur amour et leur encouragement.

Je voudrais remercier également mon directeur de mémoire M.NASR EDDINE KHORISSI professeur d'électronique à l'université de Blida 1 de m'avoir encadré, orienté, aidé et conseillé.

Un grand et spécial merci à madame NACEUR pour son soutien.

Enfin, j'adresse mes sincères remerciement aux membres du jury d'avoir accepté d'évaluer mon travail et pour leurs remarques constructives.

ملخص: ويهدف المشروع إلى توفير نظامين لتتبع المرضى، جهاز تتبع الأجهزة ونظام ذكي ومتقدم لمساعدة الأطباء النفسيين وعلماء النفس في علاج الإدمان، مجهزين بأحدث تقنيات الذكاء الاصطناعي (AI). ويستخدم الجهاز الأخير اللغة الطبيعية (NLP) في المراقبة والكشف عن الأنشطة باستخدام تقنيات الذكاء ومراقبة الحالة الصحية للمريض وحالته النفسية. كما تم تصميم مساعد صوتي متعدد اللغات لتطبيق أندرويد يتكون من منصة مفتوحة متاحة لجميع الأطباء.

كلمات المفاتيح: معالجة اللغات الطبيعية (NLP)، الذكاء الاصطناعي (AI)، إنترنت الأشياء (IOT)، ESP32، MPU6050، Android Studio، JAVA.

Résumé : Le projet vise à fournir deux systèmes de suivi de malades, un dispositif de suivi matériel et un système intelligent et avancé pour aider les psychiatres et les psychologues dans le traitement des addictions, équipé des dernières technologies de l'intelligence artificielle (AI). Ce dernier dispositif utilise le langage naturel (NLP) dans le suivi, la détection des activités avec des technologies d'intelligence et la surveillance de l'état de santé du malade ainsi que de son état psychologique. Un assistant vocal multilingue de l'application Android constituée par une plateforme ouverte disponible pour tous les médecins, a été également conçue.

Mots-clés : Traitement du langage naturel (NLP), Intelligence artificielle (AI) , Internet des objets (IOT), ESP32, MPU6050, Android Studio, JAVA

Abstract : The project aims to provide two patient tracking systems, a hardware tracking device and an intelligent and advanced system to assist psychiatrists and psychologists in the treatment of addictions, equipped with the latest artificial intelligence (AI) technologies. The latter device uses natural language (NLP) in monitoring, detecting activities with intelligence technologies and monitoring the patient's health status as well as his psychological state. A multilingual voice assistant of the Android application constituted by an open platform available to all doctors, was also designed.

Keywords : Natural Language Processing (NLP), Artificial Intelligence (AI), Internet of Things (IOT), ESP32, MPU6050, Android Studio, JAVA.

Listes des acronymes et abréviations

2FA : L'authentification à 2 facteurs

3D : Trois dimensions

AI : Artificielle intelligence

API : application programming interface

ARC : AI2 Reasoning Challenge

ASAM : American Society of Addiction Medicine

CBT : cognitive behavioral therapy

FPS : Frame per second

IoT : L'Internet des objets

LLM: Large Language Model

MMLU : Massive Multitask Language Understanding

NLP: Naturel langage processing

OLED : Organic Light Emitting Diode

RAM : Random access memory

TCC : La thérapie cognitivo-comportementale

TFLite: Tensorflow Lite

TTS :text to speech

UI : user interface

UX: user experience

VOSK: Vosk Speech Recognition

WebGL : Web Graphics Library

WebView: navigateur Web intégré à une application

WHO : Organisation mondiale de la santé

XML: Extensible Markup Language

Table des matières

<u>INTRODUCTION GENERALE</u>	1
<u>CHAPITRE 1 Généralités et introduction du système de suivi</u>	3
1.1 Introduction	3
1.2 Définition de la santé mentale.....	4
1.3 Définition de l'addiction	5
1.4 L'importance de la santé mentale et les dangers de sa négligence	5
1.5 Le besoin de système intelligents pour soutenir la psychothérapie	6
1.6 Thérapie comportementale et cognitive	7
1.7 L'importance de l'intervention précoce dans le traitement des troubles psychologiques	7
1.8 Les difficultés dans le traitement traditionnel en Algérie	8
1.9 Conclusion	8
<u>CHAPITRE 2 Systèmes et technologies utilisés par notre application</u>	10
2.1 Introduction	10
2.2 Intelligence Artificielle.....	10
2.3 l'internet des objets	11
2.4 La Reconnaissance vocale	12
2.5 Le moteur texte à voix	13
2.6 Analyse intelligente locale avec TensorFlow Lite	14
2.7 Firebase	15
2.8 Développement de système avec Android Studio	16

2.8.1 Environnement de développement Android Studio	16
2.8.2 Langage de programmation Java	16
2.8.3 Fichiers de conception XML	17
2.9 Traitement du langage naturel (NLP)	18
2.9.1 Qu'est-ce que le traitement du langage naturel (NLP)	18
2.9.2 Étapes de fonctionnement du système NLP dans le projet	19
2.9.3 Utilisation du modèle Mistral dans le système	20
2.9.4 Auto apprentissage (Self-Learning)	21
2.9.5 Formation personnalisée (réglage fin)	22
2.10 Techniques avancées pour améliorer l'interaction intelligente et la psychanalyse.....	23
2.10.1 Algorithme de distance Levenshtein	23
2.10.2 Analyse de l'activité physiologique à l'aide du MPU6050	24
2.10.3 Analyse du signal audio (INMP441)	24
2.10.4 Utiliser la détection de mot réveil (Porcupine)	24
2.11 Outils et technologies de soutien pour l'intégration de l'IA et l'interaction visuelle	25
2.11.1 Google Colab – Cloud Training Environment	25
2.11.2 Python – Le langage de programmation de base pour l'intelligence artificielle	25
2.11.3 Avatar humain réaliste en 3D – Personnage interactif réaliste	26
2.11.4 Lip Sync (LipSync) – pour créer une interaction audio-visuelle	27
2.12 Conclusion	28

CHAPITRE 3 Conception du système intelligent pour accompagner les

<u>patients psychiatriques</u>	29
3.1 Introduction	29
3.2 Structure générale du système	29
3.3 Composants logiciels	33
3.3.1 Application Android	33
3.3.2 Structuration des fichiers dans la demande	35
3.4 Intelligence artificielle et traitement du langage naturel (NLP)	36
3.4.1 Modèle Mistral – Puissance analytique du système	36
3.4.2 Fine-Tuning Custom	36
3.4.3 Algorithmes d'analyse supplémentaires	37
3.5 Analyse de la voix et du son	37
3.5.1 Ce que nous analysons	37
3.5.2 Algorithmes d'analyse	38
3.6 ESP32 et capteurs physiques (INMP441 + MPU6050)	38
3.6.1 INMP441 – Microphone numérique HD	38
3.6.2 MPU6050 – Capteur de mouvement et d'équilibre	39
3.6.3 Écran OLED	39
3.6.4 Envoi de données à Firebase	39
3.7 Personnage virtuel 3D (Avatar humain 3D) et synchronisation des lèvres (LipSync)	40
3.7.1 L'importance de l'interaction visuelle en psychothérapie	40

3.7.2 Outils de conception et d'animation des personnages	41
3.7.3 Caractéristiques du mouvement de caractère	41
3.7.4 Comment fonctionne le lip sync	42
3.7.5 L'impact de ce caractère sur l'expérience utilisateur	42
3.8 Système de sécurité et de protection de la vie privée	42
3.8.1 Authentification sécurisée de l'utilisateur et du médecin (Authentification)	42
3.8.2 Chiffrement des données	43
3.8.3 Politique de contrôle d'accès	43
3.8.4 Protéger les fichiers audio et les textes	43
3.8.5 Contrôle de la confidentialité des patients	43
3.8.6 Politique de mise à jour et de suivi	44
3.9 Diagrammes finaux et tableaux récapitulatifs	44
3.9.1 Diagramme de flux de données dans le système	44
3.9.2 Tableau du rôle de chaque composante	45
3.9.3 Structure de l'application Android	45
3.9.4 Dessin de connexion des composants physiques (schéma de circuit)	46
3.10 Contribution des appareils intelligents à la santé mentale	46
3.10.1 Présentation de l'ESP32	46
3.10.2 Le Capteur sonore haute résolution INMP441	47
3.10.3 Le Capteur de mouvement et d'équilibre MPU6050	48
3.10.4 L'Écran OLED d'affichage de l'information	49

3.11 Conclusion	50
-----------------------	----

CHAPITRE 4 Tests, résultats et analyse des performances..... 52

4.1 Introduction	52
------------------------	----

4.2 Environnement d'essai	52
---------------------------------	----

4.2.1 Les composants techniques	52
---------------------------------------	----

4.2.2 Les participants à l'essai	53
--	----

4.3 Tests de performance pour chaque composant du système	53
---	----

4.3.1 Performance des applications Android	53
--	----

4.3.2 Performances de l'assistant vocal (Wake Word + VOSK + TTS)	54
--	----

4.3.3 Performance du modèle d'intelligence artificielle (NLP Mistral)	54
---	----

4.3.4 Performances de la montre(bracelet) intelligente (ESP32 + INMP441 + MPU6050)	55
--	----

4.3.5 Performances des caractères 3D (3D + LipSync)	55
---	----

4.4 Expérience utilisateur (commentaires des utilisateurs)	56
--	----

4.4.1 Les Participants à l'expérience	56
---	----

4.4.2 Questions pour les utilisateurs	56
---	----

4.4.3 Résultats de l'enquête sur la satisfaction	57
--	----

4.4.4 Notes des médecins	57
--------------------------------	----

4.4.5 Suggestions des utilisateurs	57
--	----

4.5 Analyse des résultats	58
---------------------------------	----

4.5.1 Conformité du système aux objectifs généraux	58
--	----

4.5.2	Analyse de l'interaction des utilisateurs	59
4.5.3	Analyse de la performance technique	59
4.5.4	Intégration de systèmes multiples	59
4.5.5	Comment Surmonter les difficultés	59
4.6	Évaluation générale et comparaison avec d'autres solutions.....	60
4.6.1	Évaluation globale du système	60
4.6.2	Comparaison avec des solutions globales similaires	60
4.6.3	Valeur ajoutée réelle	61
4.7	Remarques, limites et défis.....	61
4.7.1	Limitations techniques	61
4.7.2	Limites humaines et comportementales	62
4.7.3	Défis lors du développement	62
4.7.4	Limites de la portée du projet.....	63
4.8	Conclusion	63
	<u>CONCLUSION GENERALE</u>	64
	<u>ANNEXE</u>	66
	<u>BIBLIOGRAPHIE</u>	70

Liste des figures

Figure 2.1: Intelligence Artificielle.....	11
Figure 2.2: Internet des Objet	11
Figure 2.3: Speech Recognition	12
Figure 2.4: Text to Speech	13
Figure 2.5: ElevenLabs IO	14
Figure 2.6: TensorFlow Lite	15
Figure 2.7: Google Firebase	15
Figure 2.8: Android Studio	16
Figure 2.9: Java Langage de programmation	17
Figure 2.10: XML.....	18
Figure 2.11: Naturel Language Processing	19
Figure 2.12: Fonctionnement de NLP	20
Figure 2.13: Mistral AI	21
Figure 2.14: Self-Learning.....	21
Figure 2.15:Levenshtein Algorithme.....	23
Figure 2.16: Pico voice Engine	24
Figure 2.17: Google Colab	25
Figure 2.18: Python	26
Figure 2.19: Avatar humain réaliste en 3D	26
Figure 2.20: : Lip Sync (LipSync).....	27
Figure 3.1: Diagramme du système 1.....	30
Figure 3.2: Diagramme du system 2.....	32
Figure 3.3: les interface Android App	33
Figure 3.4: les interface Android App	34
Figure 3.5: Structure des Fichier	35
Figure 3.6: Training Model Google Colab	37
Figure 3.7: Algorithmes classification Sound Capteur INMP441.....	38
Figure 3.8: Lien électronique avec ESP32.....	39
Figure 3.9: Send Data FireBase	40
Figure 3.10: caractère 3D	41
Figure 3.11: Diagramme de flux de données	44
Figure 3.12: schéma de circuit	46
Figure 3.13: ESP32.....	47
Figure 3.14: INMP441.....	48
Figure 3.15: MPU6050.....	49
Figure 3.16: Ecran OLED	50

Liste des tableaux

Tableau 3.1: conception et d'animation des personnages	41
Tableau 3.2: contrôle d'accès	43
Tableau 3.3: rôle de chaque composante	45
Tableau 4.1 :Performance des applications.....	53
Tableau 4.2: Performances de l'assistant vocal	54
Tableau 4.3: Performance du modèle d'intelligence	54
Tableau 4.4: Performances du système de montre intelligente	55
Tableau 4.5: Performances des caractères 3D	55
Tableau 4.6: Participants à l'expérience	56
Tableau 4.7: Résultats de l'enquête sur la satisfaction	57
Tableau 4.8: Conformité du système aux objectifs généraux	58
Tableau 4.9: Comparaison avec des solutions globales similaires	61

Introduction générale

La santé, telle que définie par l'Organisation mondiale de la santé, est un état de bien-être physique, psychologique et social complet, et non simplement l'absence de maladie ou d'invalidité. Cette définition englobe les aspects intégraux de la vie humaine et affirme que la santé exige un équilibre holistique entre le corps, l'esprit et la société.

La santé mentale dans notre monde contemporain est devenue l'une des principales préoccupations des sociétés modernes, suscitant l'attention des chercheurs.

De cette situation résulte l'augmentation des pressions sociales et économiques, notamment avec les changements des modes de vie, en particulier chez les jeunes en Algérie.

Malgré les progrès technologiques considérables, les services de santé mentale souffrent toujours d'une grave pénurie de soins, car l'aspect psychologique, qui est le plus important dans la vie des individus, est négligé.

La santé mentale ne se limite pas à l'absence de maladie, mais englobe également le véritable équilibre intérieur qui aide l'individu à faire face aux difficultés de la vie.

Etant donné que les troubles psychologiques peuvent rester cachés et invisibles, il devient indispensable de trouver des solutions précoces et une offre de soutien avant qu'il ne soit trop tard.

Parmi les meilleures méthodes thérapeutiques dans le domaine de la santé mentale qui se sont avérées efficaces, est la thérapie comportementale et cognitive (CBT).

Ce type de traitement nécessite une communication continue et une orientation personnelle, ce qui constitue un grand défi. Notre projet est venu comme une initiative scientifique visant à développer un système intelligent de surveillance à distance de

l'état psychologique. Dans ce système, nous avons misé sur l'intégration des dernières technologies, de l'intelligence artificielle (AI) et celle de l'internet des Object (IOT) avec l'application de thérapie comportementale cognitive pour Fournir un soutien personnel et efficace au patient d'une manière moderne et innovante .

Pour faciliter l'interaction avec ce système, nous avons créé et développé une application Android en utilisant les langages Java pour la programmation et Android Studio comme outil de développement .

Cela permet à l'utilisateur de communiquer facilement avec le médecin, de suivre les rendez-vous et d'utiliser l'assistant vocal, qui représente presque un ami fidèle et solidaire qui vient avec deux personnalités féminines et masculines.

Ces derniers écoutent et enregistrent vos données, les réponses sont immédiates avec des voix réalistes et une simulation proche des dialogues humains .

Le projet représente une plateforme ouverte pour tous les psychiatres, afin que tout médecin puisse créer son propre compte et suivre l'état de ses patients à distance d'une manière organisée et sécurisée, ce qui augmente les chances de traitements, réduit le nombre de patients non suivis et élargit l'accès aux soins psychologiques.

Le mémoire de notre projet est organisé en 4 chapitres :

Dans le chapitre 1 nous présentons une étude sur la santé mentale, la thérapie comportementale cognitive et les solutions numériques modernes. Le Chapitre 2 sera consacré à la Présentation des outils et des technologies adoptés pour le développement du Système. La conception du système avec ses composants logiciels et matériels fera l'objet du chapitre 3. Au Chapitre 4 nous présenterons les étapes de mise en œuvre ,les résultats obtenus ainsi que des suggestions d'améliorations.

Enfin, nous terminons notre étude par une conclusion générale, des suggestions et des perspectives.

1.1 Introduction :

La santé mentale est l'un des principaux piliers de la santé humaine en général, avec la santé physique.

Avec l'évolution des modes de vie et les difficultés sociales croissantes, l'attention portée à la santé mentale est devenue cruciale.

Les troubles mentaux, en particulier l'addiction, sont parmi les plus grands défis auxquels font face les individus et les sociétés en raison de leurs effets néfastes sur la qualité de vie, la productivité et la capacité de chaque personne à s'épanouir et à donner le meilleur d'elle-même.

Sur la base de cette observation, notre projet vise à proposer une solution technologique innovante pour soutenir la santé mentale et améliorer le télétraitement en intégrant l'intelligence artificielle et Internet.

Dans ce chapitre, nous présentons les concepts de base de la santé mentale et de l'addiction, en soulignant leur importance dans la vie humaine, ainsi que la nécessité d'adopter des solutions intelligentes et modernes pour faire face aux défis contemporains.

Avec le développement des technologies numériques pour la santé mentale, il ne suffit plus de se limiter à de simples interactions textuelles ou vocales.

A l'inverse, les dispositifs intelligents équipés de capteurs capables de surveiller le comportement physique et de biomarqueurs représentent une étape importante dans l'évaluation plus précise du statut psychologique, permettant une intervention précoce dans des situations critiques.

Dans ce contexte, un ensemble de capteurs électroniques a été intégré à notre projet pour recueillir des données réalistes sur l'activité physique et le statut vocal de l'utilisateur, pour fournir un soutien personnalisé et réactif.

1.2 Définition de la santé mentale :

L'organisation mondiale de la santé (WHO) définit la sante mentale comme un état de bien-être dans lequel l'individu prend conscience de ses capacités , peut gérer les stress ordinaires de la vie , Travailles de façon productive et contribuer à sa Communauté [1] .

La santé mentale ne se résume donc pas uniquement à l'absence de troubles psychiques mais concerne aussi la capacité de l'individu à penser de manière équilibrée, à gérer ses émotions, à réagir avec sagesse face aux aises et à établir des relations saines et stables avec autrui.

L'un des aspects de la santé mentale est l'aspect émotionnel qui concerne la façon dont l'individu exprime ses sentiments et apprend à les gérer efficacement.

L'aspect mental est la capacité de l'individu à raisonner correctement et à prendre des décisions éclairées. L'aspect social représente la capacité de construire et de maintenir des relations saines. Quant à l'aspect comportemental, il revoie à la capacité d'agir de manière appropriée, en accord avec les valeurs et les normes de la société. La santé mentale est également affectée par plusieurs facteurs, notamment les facteurs génétiques et biologiques, les expériences de vie traumatisantes telles que la perte d'un être cher, la violence, la pauvreté ... etc. ainsi que les conditions environnementales et sociales.

Pour préserver la santé mentale, il est essentiel de renforcer la résilience psychologique, de bénéficier d'un soutiens social solide ainsi que disposition de services thérapeutiques et de conseil appropriés.

Une santé mentale équilibrée permet à l'individu de surmonter les épreuves de la vie et de s'épanouir pleinement, tandis que sa dégradation peut conduire à des maladies et des troubles graves.

1.3 Définition de l'addiction :

L'addiction est un état de dépendance compulsive à une substance ou un comportement, dans lequel l'individu perd le contrôle de ses actions malgré les conséquences négatives graves. Selon l'association américaine de médecine des addictions (ASAM) , l'addiction est une maladie chronique qui affecte les circuits cérébraux liés à la motivation, à la récompense. Elle se caractérise par la poursuite de la consommation d'une substance ou l'adoption d'un comportement spécifique, malgré la connaissance des effets nocifs physiques ou psychiques qu'ils peuvent engendrer [2].

Parmi les types d'addiction, on distingue l'addiction aux substances, telles que les drogues, l'alcool et le tabac, ainsi que l'addiction comportementale comme celle liée aux jeux vidéo à l'usage excessif d'internet ou aux jeux d'argent.

Les caractéristiques d'une personne ad dicté incluent la perte de contrôle face aux désirs de poursuivre le comportement malgré ses effets nocifs, un besoin accru de consommer la substance ou de répéter l'acte, ainsi que des symptômes de sevrage lors de tentative d'arrêt.

Parmi les causes de l'addiction, on retrouve des facteurs génétiques, des troubles psychologiques associés, l'influence de l'environnement social, ainsi qu'un désir d'échapper des émotions négatives.

L'addiction mène à l'isolement social, et perte de travail et détérioration de la santé physique et mentale et, dans certains cas elle peut entraîner la mort.

1 .4 L'importance de la santé mentale et les dangers de sa négligence :

La bonne santé mentale représente la pierre angulaire pour le succès de l'individu et l'épanouissement de la société. Une personne qui a un équilibre mental peut apprendre, travailler, nouer des relations saines et participer efficacement à la

société. D'un autre Côté, le fait de négliger la santé mentale conduit à des résultats désastreux qui dépassent l'individu et affectent toute une société.

Parmi les effets négatifs résultant de la santé mentale pour l'individu, des troubles tels que l'anxiété, l'addiction, ou même le suicide. Pour les familles, la désintégration des liens familiaux, la violence domestique.

Quant à la société, l'augmentation de la criminalité et une diminution de la productivité économique.

Selon les études, chaque dollar investi dans la promotion de la santé mentale rapporte quatre fois plus en améliorant la productivité et en réduisant les couts de traitement.

Il est donc devenu nécessaire d'intégrer des programmes de prévention et de soutien psychologique dans les politiques générales de santé

1.5 Le besoin de système intelligents pour soutenir la psychothérapie :

En raison de l'augmentation croissante de nombre d'atteintes de troubles mentaux par rapport au nombre de médecin et de spécialistes disponibles, il est urgent d'adopter des solutions technologiques intelligentes pour renforcer le traitement classique.

Le rôle de la technologie moderne est de surveiller en permanence l'état psychologique analyser les données des patients et de prédire les changements comportementaux, Fournir un soutien immédiat adapté à chaque cas,

Améliorer la rapidité d'accès aux traitements et prédire les couts. Les systèmes intelligents contribuent également par le biais d'assistants virtuels (comme l'assistant vocal de notre projet) capables de parler aux patients et leur donner des conseils immédiats.

L'utilisation de l'intelligence artificielle pour analyser la parole, détecter Les symptômes précoces et stocker des données en toute sécurité grâce aux technologies Firebase. Ce qui permet aux médecins d'accéder immédiatement aux informations des patients peu importe leur localisation. Notre projet vise ces

technologies dans le but de créer une plateforme ouverte qui permet à chaque patient et médecin de bénéficier d'un soutien intelligent et fiable.

1.6 Thérapie comportementale et cognitive :

La thérapie comportementale et cognitive est l'une des méthodes thérapeutiques les plus efficaces utilisées pour aider les personnes souffrantes de troubles psychologiques tels que la dépression, l'anxiété, l'addiction et le trouble de stress post-traumatique. La thérapie comportementale et cognitive est basée sur l'idée fondamentale que les pensées négatives ou illogiques sont ce qui cause des émotions et des comportements négatifs, et donc changer les schémas de pensée conduit à améliorer les sentiments et les comportements.

Parmi ses principes l'identification des pensées négatives et des fausses croyances, l'analyse de la relation entre les pensées, les sentiments et les comportements, le remplacement des pensées réalistes et positives.

De plus, apprendre à s'adapter aux situations difficiles de façon saine [3].

La thérapie comportementale cognitive est généralement appliquée par des séances régulières avec un thérapeute spécialisé, où le patient est formé à reconnaître les schémas de pensées négatives, il teste la réalité de ces pensées et pratique des modes de pensée positives. Ainsi il apprend de nouveaux comportements pour faire face à des situations difficiles. Cette méthode est connue par son caractère à court terme et ciblé, car elle met l'accent sur des problèmes actuels au lieu de plonger dans les expériences passées, ce qui en fait un traitement pratique et efficace pour un grand nombre de troubles psychologiques.

1.7 L'importance de l'intervention précoce dans le traitement des troubles psychologiques :

La phase d'intervention précoce est considérée comme l'un des facteurs les plus importants et déterminants pour le succès du traitement des troubles psychologiques. De nombreuses recherches scientifiques confirment que le diagnostic précoce et

l'accompagnement psychologique immédiat réduisent la gravité des symptômes, et améliore grandement les chances d'un rétablissement complet.

L'intervention précoce permet de traiter le problème avant qu'il ne se développe en une condition complexe ou chronique, ce qui épargne au patient, à la famille et à la société un grand coût psychologique et économique.

Par exemple, le traitement des symptômes de la dépression au premier stade peut empêcher qu'elle ne dégénère en dépression sévère accompagnée de pensées suicidaires. Découvrir les prémices de la dépendance peut arrêter le développement d'une dépendance totale aux substances, En outre une intervention précoce peut renforcer la résilience psychologique de l'individu et améliorer sa capacité à s'adapter aux défis futurs, ce qui contribue à améliorer sa qualité de vie globale.

1.8 Les difficultés dans le traitement traditionnel en Algérie :

Malgré les efforts déployés pour améliorer les soins psychologiques en Algérie le traitement traditionnel est confronté à plusieurs défis qui rendent l'accès aux services de soutien psychologique limité et difficile pour de larges segments de la société.

Parmi ces difficultés, la plus importante est le petit nombre de psychiatres par rapport à la forte demande de services psychologiques ce qui entraîne de longues files d'attente, Tout comme les cliniques et les centres spécialisés sont concentrés dans les grandes villes, tandis que les zones rurales et éloignées souffrent d'une absence quasi totale de services psychologiques, Tous ces facteurs rendent nécessaire de réfléchir à des alternatives intelligentes et pratiques pour accéder au traitement psychologique d'une manière plus facile et confidentielle.

C'est pourquoi notre projet propose un système intelligent et ouvert qui fournit un soutien psychologique et des conseils à travers l'application Android et un assistant intelligent tout en réduisant les barrières normales [4].

1.9 Conclusion :

Dans ce chapitre, nous avons réalisé que les soins de santé mentale sont considérés comme une partie intégrante d'une vie saine complète pour l'individu. De nos jours,

l'aggravation du phénomène de la dépendance et la propagation des troubles psychologiques justifie l'importance de rechercher des solutions innovantes qui contribuent à améliorer les soins psychologiques.

À la lumière du développement technologique rapide, l'intelligence artificielle et l'Internet des objets sont devenus des outils prometteurs pour élargir le champ de soutien psychologique et fournir un traitement plus personnalisé et efficace.

Notre projet s'inscrit dans cette tendance visant à fournir un service intégré et complet pour améliorer la santé mentale en s'appuyant sur des packages développés par la technologie moderne.

2.1 Introduction :

Ces dernières années, nous avons assisté à un développement important des technologies de l'intelligence artificielle et de l'Internet des objets, qui ont permis de construire des systèmes intelligents capables de simuler les interactions humaines d'une manière naturelle et efficace.

Dans ce contexte, notre projet s'appuie sur un ensemble intégré de technologies modernes qui permettent la communication vocale, la psychanalyse, la réponse intelligente et la représentation visuelle réaliste.

Ce chapitre a pour objectif de présenter les outils, les environnements logiciels et les composants techniques utilisés dans le développement du système de soutien en santé mentale, Fournir une explication détaillée de la fonction de chaque technologie et de son rôle dans l'amélioration de l'expérience utilisateur et la réalisation des objectifs thérapeutiques du projet.

2.2 Intelligence Artificielle:

L'intelligence artificielle est la capacité des systèmes numériques à imiter l'intelligence humaine en apprenant des données, en les analysant et en prenant des décisions fondées sur celles-ci.

Dans le domaine psychologique, l'intelligence artificielle est utilisée pour analyser la parole, prédire les changements d'humeur et même fournir des conseils thérapeutiques initiaux à l'utilisateur.

Dans notre projet, l'intelligence artificielle est utilisée pour analyser la parole de l'utilisateur, comprendre le ton de la voix, déterminer s'il éprouve des symptômes d'anxiété ou de dépression, puis diriger une réponse appropriée via l'assistant intelligent [5].

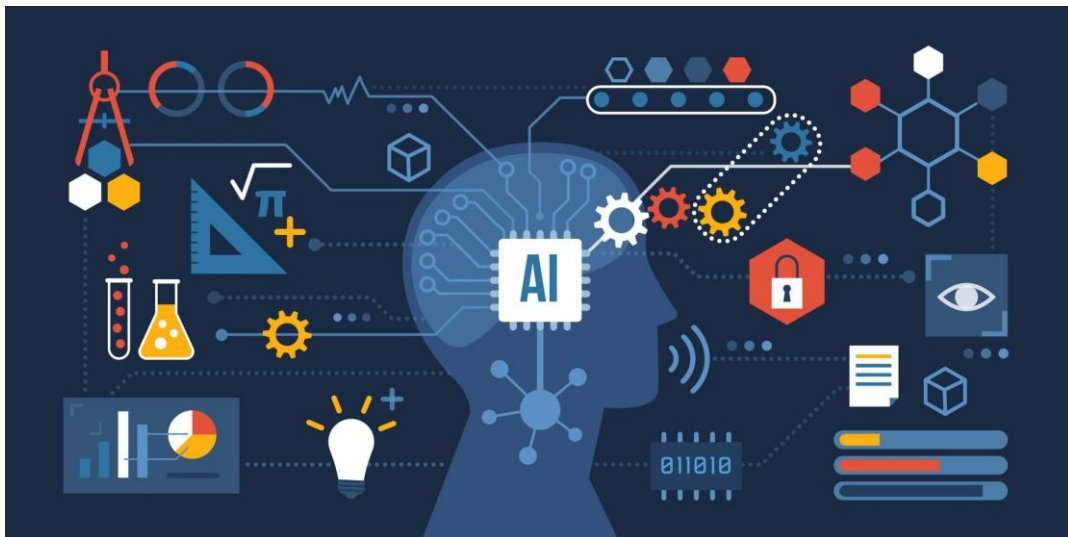


Figure 2.1 : Intelligence Artificielle

2.3 l'internet des objets :

L'Iot désigne le réseau d'appareils intelligents connectés à Internet, qui peuvent recueillir et échanger des données. Dans le domaine de la santé mentale, l'iot peut être utilisé pour relier des capteurs (un accéléromètre, un microphone, des moniteurs de sommeil) au système intelligent afin d'analyser le comportement du patient et les changements d'état psychologique.

Dans notre projet, nous utilisons l'iot pour connecter les appareils (les capteurs) à l'application Android tels qu'un microphone et un mpu6050, afin de comprendre le comportement du patient en termes de mouvement, de la voix et de vitesse de réaction, ce qui améliore la précision de l'évaluation psychologique [6].

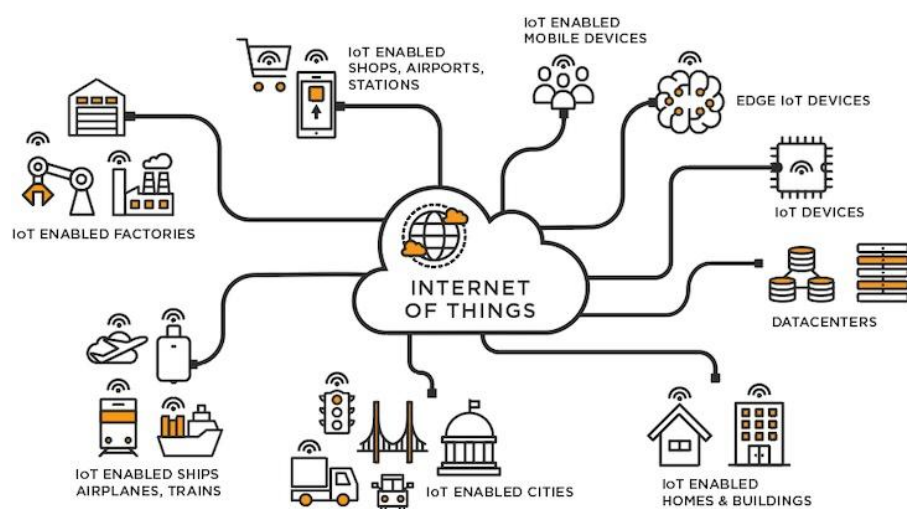


Figure 2.2 : Internet des Objet.

2.4 La Reconnaissance vocale :

La technologie de reconnaissance vocale est l'une des technologies les plus importantes dans les systèmes intelligents, car elle permet à l'utilisateur d'interagir par la voix sans avoir besoin de taper ou d'appuyer sur des boutons.

Dans notre projet, le moteur de reconnaissance vocale qui a été utilisé est Speech Recognizer et Une bibliothèque officielle dans l'api Android qui prend en charge la plupart des langues.

Ce moteur est utilisé pour convertir la parole du patient en texte qui peut être analysé par le système intelligent [7].

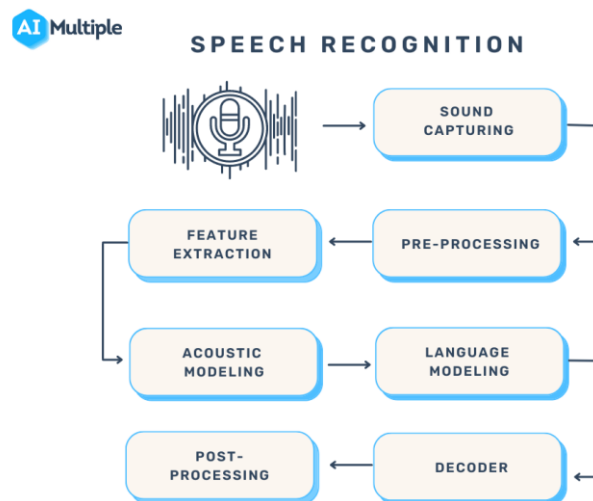


Figure 2.3: Reconnaissance de la parole

2.5 Le moteur texte à voix:

Text To Speech (TTS) est l'une des phases principales de la conception d'un assistant intelligent qui interagit avec les utilisateurs de manière naturelle.

Dans notre projet, le moteur ElevenLabs TTS a été utilisé, il est l'un des moteurs de texte-à-voix les plus puissants actuellement, qui se caractérise par des sons très réalistes qui ressemblent à la vraie voix humaine avec un pourcentage élevé.

Ce moteur prend en charge plusieurs langues et accents, avec la possibilité de personnaliser le ton de la voix (calme, rapide, émotionnel ...).

Possibilité de l'utiliser via API ou de l'intégrer dans une application Android en utilisant des solutions middleware telles que la connexion via un serveur local ou des services web [8].



Figure 2.4: Text to Speech

ElevenLabs a été choisi dans le cadre du projet en raison de ses capacités avancées pour convertir des textes en audio hautement réaliste, ce qui améliore l'expérience psychologique de l'utilisateur au sein du système intelligent. Ce moteur permet à l'assistant vocal de parler dans une voix humaine convaincante qui peut exprimer différentes émotions (telles que la réassurance, le sérieux, la joie) en fonction du contexte émotionnel de la réponse. Il prend également en charge plusieurs dialectes,

ce qui donne au dialogue un caractère personnel et renforce le sentiment de confiance et de confort de l'utilisateur. ElevenLabs est utilisé dans le projet envoyant du texte généré par l'analyse de la saisie utilisateur (après traitement NLP) à l'API ElevenLabs, où il sera converti en un fichier audio qui s'exécute directement dans l'application. Cet audio est automatiquement enregistré et stocké dans la base de données de Supabase dans le dossier du patient pour suivre et analyser les progrès psychologiques au fil du temps [9].



Figure 2.5: ElevenLabs IO

2.6 Analyse intelligente locale avec TensorFlow Lite :

TensorFlow Lite est une version allégée de la bibliothèque d'IA « TensorFlow », conçue pour fonctionner sur les appareils de smartphones sans avoir besoin d'une connexion Internet.

Dans notre projet, les modèles de TensorFlow sont utilisés pour analyser les données audio localement, identifier des mots-clés et reconnaître des modèles indicatifs de troubles psychologiques [10].

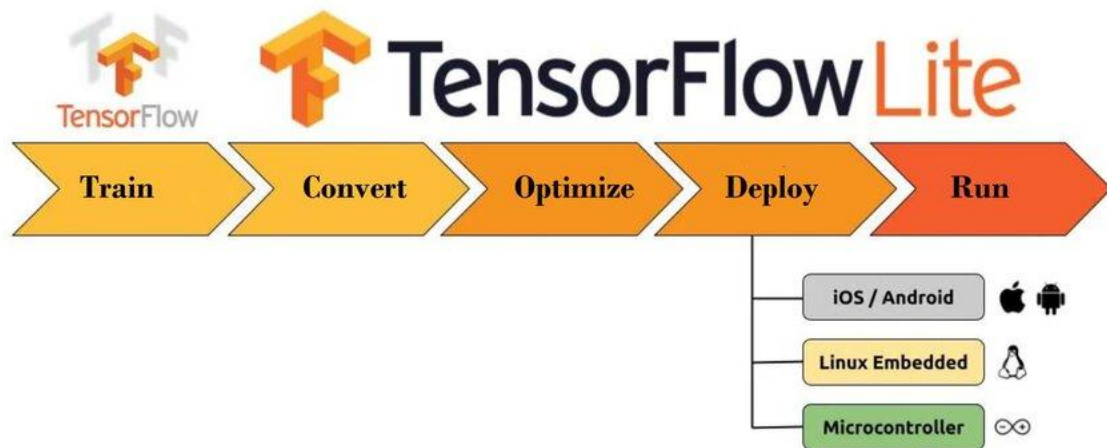


Figure 2.6: TensorFlow Lite

2.7 Firebase:

Firebase est une plate-forme complète de Google qui fournit un ensemble puissant d'outils et de services permettant aux développeurs de créer des applications hautes performances rapidement et en toute sécurité. Dans le cadre de notre projet, Firebase a été certifié pour soutenir l'infrastructure intelligente nécessaire à un système basé sur l'interaction en temps réel et l'auto-apprentissage. Firebase est utilisé dans le projet sous plusieurs aspects, notamment : le stockage des conversations entre le patient et l'assistant vocal dans le but de suivre l'état psychologique et d'améliorer la qualité du soutien fourni, gérer les comptes utilisateurs des patients et des médecins via un système d'authentification sécurisé, envoyer des alertes immédiates et des notifications pour rappeler les sessions ou les alertes importantes, et également stocker de nouvelles données que l'assistant apprend au cours des conversations pour soutenir apprendre et développer ses réponses à l'avenir [11].



Figure 2.7 : Google Firebase.

2.8 Développement de système avec Android Studio :

2.8.1 Environnement de développement Android Studio :

Android Studio est l'environnement de développement officiel de Google pour le développement d'applications Android. Il s'agit d'un environnement intégré qui permet aux développeurs de concevoir des interfaces utilisateur, de tester des applications et de déboguer des erreurs facilement et efficacement. Dans ce projet, Android Studio a été adopté pour créer l'interface interactive de l'assistant intelligent, reliant l'application à la plate-forme cloud Firebase, ainsi que l'intégration des technologies d'IA via des bibliothèques externes telles que VOSK pour le traitement audio. Nous avons également profité des fonctionnalités d'Android Studio telles qu'un émulateur (Emulator) pour tester l'application sur des machines virtuelles avec plusieurs spécifications, le système Gradle pour les gérer et relier de façon transparente les bibliothèques requises dans un environnement de développement flexible et robuste [12] .



Figure 2.8: Android Studio.

2.8.2 Langage de programmation Java :



Le langage Java a été choisi pour la programmation d'applications en raison de sa position bien établie comme un langage officiellement approuvé pour le développement d'applications Android depuis le lancement du système, en plus de sa grande capacité à s'intégrer avec la plupart des bibliothèques d'intelligence artificielle et des outils techniques modernes. Java offre une grande flexibilité dans la gestion des données, le traitement de la voix, l'exécution d'opérations réseau et le contrôle transparent des interfaces utilisateur. Dans le cadre de ce projet, Java a été utilisé pour contrôler entièrement l'interaction vocale entre l'utilisateur et l'assistant intelligent et s'intégrer à la plateforme Firebase en termes de lecture et d'écriture de données, en plus d'analyser les résultats extraits du modèle d'intelligence artificielle, et également dans le contrôle des capteurs associés au système intelligent, tels que les capteurs de mouvement et de son [13].



Figure 2.9 : Java Langage de programmation

2.8.3 Fichiers de conception XML :

XML (Extensible Markup Language) est un composant essentiel dans le développement des interfaces d'applications Android, car il est utilisé pour définir la structure visuelle de chaque écran dans l'application d'une manière organisée et personnalisable. Dans notre projet, XML a été utilisé de manière pivotante pour concevoir des interfaces utilisateur conformes aux exigences du système intelligent, telles que l'interface d'interaction avec l'assistant vocal, et organiser l'affichage des conversations sous forme de bulles de dialogue pour distinguer les messages entre le patient et l'assistant. Les interfaces ont également été personnalisées pour changer en fonction du personnage assistant choisi (Jack ou Emily), en termes de couleurs et de distribution visuelle. D'autres utilisations comprenaient l'incorporation d'éléments graphiques avancés tels que le déplacement du microphone avec Lottie Animations, ce qui a contribué à une expérience utilisateur plus interactive et réaliste [14].



Figure 2.10 :XML

2.9 Traitement du langage naturel (NLP) :

À l'ère moderne, l'interaction avec les appareils n'est plus limitée aux interfaces graphiques ou aux commandes logicielles, mais plutôt la communication naturelle utilisant le langage humain est devenue un objectif intrinsèque des technologies d'intelligence artificielle. D'où l'importance du Natural Language Processing (NLP), qui est devenu le cœur battant de nombreuses applications telles que les assistants vocaux, les chatbots, la traduction automatique et les systèmes de soutien psychologique.

Dans notre projet, nous avons intégré les technologies NLP au sein d'un système intelligent qui vise à analyser et comprendre la parole faite par un utilisateur ayant des problèmes psychologiques ou comportementaux tels que l'anxiété ou la dépendance, et fournir une réponse audio naturelle et de soutien basée sur les règles de la thérapie comportementale cognitive (TCC). Ce système se distingue par sa dépendance à un modèle moderne connu sous le nom de Mistral, qui est l'un des modèles forts dans ce domaine, en plus d'introduire le concept d'auto-apprentissage continu et d'appliquer des techniques de formation personnalisées (Fine-Tuning) adapter le système aux dialectes locaux, notamment au dialecte algérien.

2.9.1 Qu'est-ce que le traitement du langage naturel (NLP) :

Le traitement du langage naturel (NLP) est un domaine interdisciplinaire qui combine l'informatique, la linguistique et la psychologie, dans le but de permettre aux machines de comprendre le langage humain d'une manière qui va au-delà de la simple conversion de la parole en texte. Il cherche à comprendre les significations, les sentiments et les intentions exprimées par les mots dans leurs différents contextes, ce qui en fait un outil efficace pour analyser plus profondément la communication humaine.

Dans le cadre de notre projet, il ne suffit pas de savoir "ce que le patient a dit", il faut analyser cette énonciation pour découvrir si elle indique une situation de risque, reflète une tension psychologique ou représente un simple discours éphémère. Il est également nécessaire d'être conscient de la possibilité de pensées suicidaires ou de signes de dépendance, dans le but d'une intervention précoce et de fournir un soutien approprié au moment opportun [15].

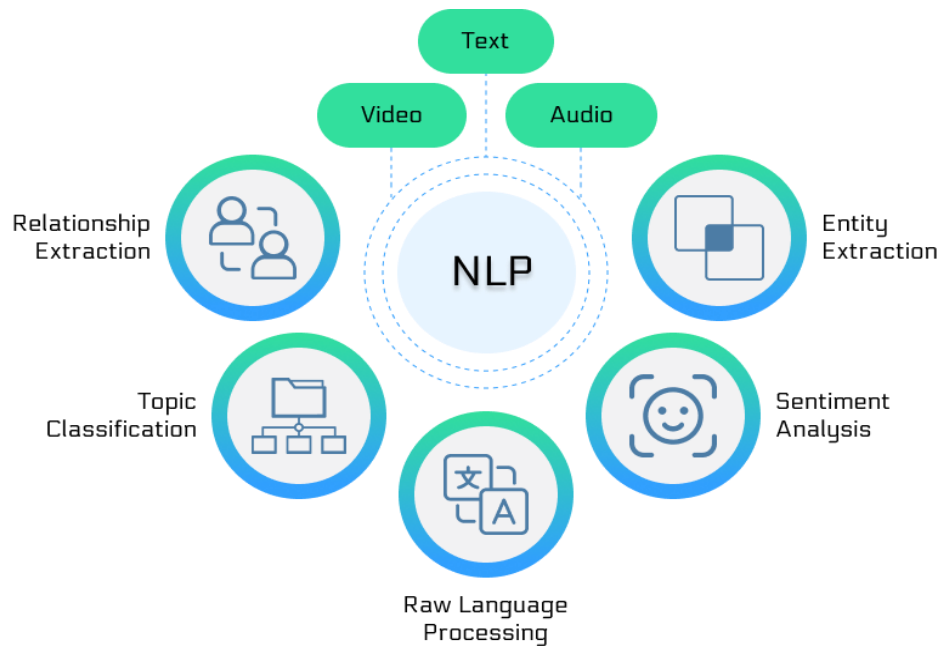


Figure 2.11 : Naturel Language Processing

2.9.2 Étapes de fonctionnement du système NLP dans le projet :

Le texte provenant de l'utilisateur dans notre système passe par plusieurs étapes précises visant à comprendre l'état psychologique avec précision et efficacité. Le processus commence au moment où les données sont saisies, lorsque le système est activé par la voix via un "mot de réveil" tel que « Emily », pour commencer à écouter la voix de l'utilisateur et la convertir en texte à l'aide du moteur de reconnaissance vocale Regonizer. Le texte est ensuite traité en supprimant les symboles inutiles, en corrigeant les erreurs courantes et en normalisant les formules linguistiques pour assurer la clarté de l'analyse.

Le texte est ensuite décomposé en petites unités linguistiques (tokenisation) pour faciliter l'analyse, et l'intention derrière la parole est alors déterminée à l'aide d'un modèle spécialisé de détection des intentions. L'utilisateur demande-t-il de l'aide? Exprime-t-il des sentiments? Y a-t-il des signes de pensées suicidaires?

L'analyse des sentiments suit, où les mots sont comparés à une base de données pré-catégorisée d'émotions positives, négatives ou sérieuses. Enfin, les résultats sont combinés pour aboutir à une évaluation globale de l'état psychologique, classée en quatre catégories principales : stable, stress léger, dépression ou danger urgent. Ces étapes assurent une réponse précise et sensible aux besoins psychologiques de l'utilisateur.

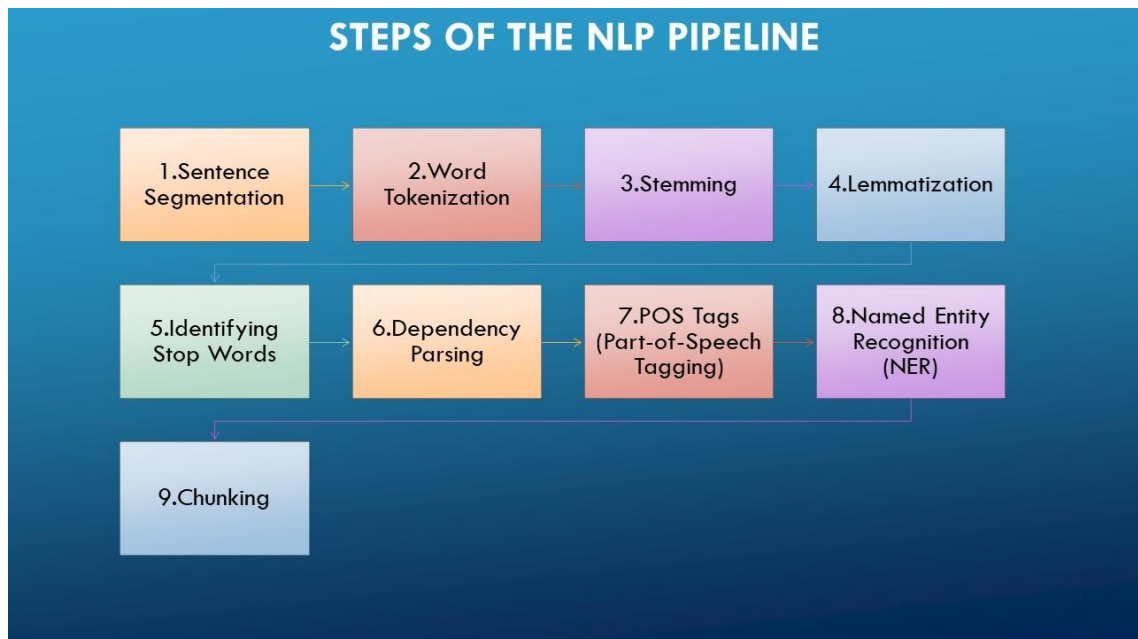


Figure 2.12 : Fonctionnement de NLP

2.9.3 Utilisation du modèle Mistral dans le système :

Mistral est un modèle linguistique avancé open source qui appartient à la génération moderne de modèles d'intelligence artificielle. Il se distingue par sa grande capacité à comprendre de longs contextes et à analyser des phrases complexes avec précision. Son excellence dans les tests tels que MMLU et ARC en fait un choix idéal pour les projets qui nécessitent une compréhension approfondie du langage et des émotions, comme dans notre. Nous avons choisi Mistral parce qu'il est personnalisable, fonctionne efficacement sur les appareils locaux sans infrastructure cloud et prend en charge plusieurs langues et dialectes, y compris le français, l'arabe et le dialecte algérien. Dans notre système, Mistral joue le rôle central d'analyser le contexte de la parole de l'utilisateur, d'interpréter les sentiments et les intentions, et de générer des réponses appropriées basées sur les principes de la thérapie cognitivo-comportementale (TCC), avec un potentiel d'amélioration continue basé sur l'interaction antérieure, assurant une expérience de soutien psychologique intelligente et humaine [16].



Figure 2.13:Mistral AI

2.9.4 Auto apprentissage(Self-Learning):

Notre système ne se limite pas à fournir des réponses, mais s'appuie plutôt sur un mécanisme d'apprentissage continu pour améliorer son rendement au fil du temps. Chaque interaction avec l'utilisateur est stockée dans Firebase, ce qui permet au système d'examiner et d'analyser les phrases qui n'ont pas été correctement comprises, dans le but d'en tirer des leçons et d'éviter de répéter la même erreur plus tard. La base de données est progressivement mise à jour avec l'ajout de nouveaux mots, des réponses améliorées et des classifications plus précises, ce qui rend le modèle plus adapté à la nature des utilisateurs et à l'évolution de leur style d'expression. De cette façon, le système se transforme en un assistant intelligent qui est en constante évolution, devenant plus compréhensif et efficace pour fournir un soutien psychologique [17].

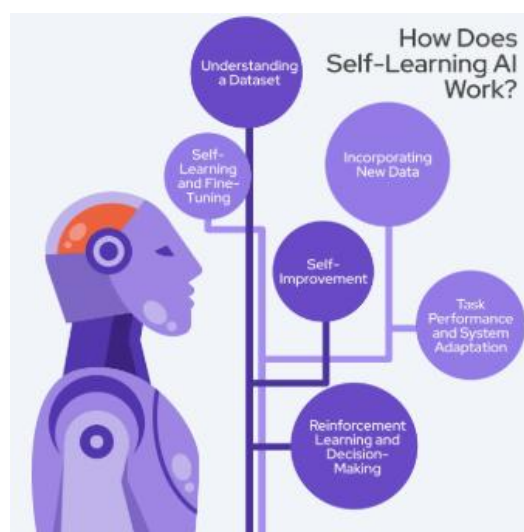


Figure 2.14 :Self-Learning

2.9.5 Formation personnalisée (réglage fin) :

Le perfectionnement, ou la formation personnalisée, est un processus par lequel un modèle linguistique prêt à l'emploi tel que Mistral est modifié et réformé sur des données spécifiques liées à un domaine particulier, dans le but d'améliorer sa précision et sa compréhension des contextes de ce domaine. Dans notre projet, nous nous appuyons sur le Fine-Tuning pour rendre le modèle plus sensible et comprendre les états psychologiques complexes et les expressions locales. Le processus commence par la collecte de centaines de phrases réalistes associées à des sentiments tels que l'anxiété, la dépression ou les remords, puis la traduction de ces phrases en dialecte algérien pour convenir à l'environnement de l'utilisateur. Chaque phrase est ensuite catégorisée par son état psychologique, et Mistral est réformé sur ces données catégorisées. Le résultat est un modèle capable d'interagir de manière intelligente et réaliste dans l'application, avec une plus grande capacité à comprendre et à répondre aux émotions dans le contexte culturel et linguistique local [18].

Le système de traitement du langage naturel dans notre projet représente la colonne vertébrale qui permet à l'assistant intelligent de comprendre l'utilisateur, de sentir ce qu'il ressent et d'interagir avec lui de manière intelligente et humaine.

Grâce à l'utilisation du modèle avancé Mistral, l'auto-apprentissage et la formation personnalisée dans des accents et des contextes réels, ce système peut se transformer d'un simple assistant technique en un véritable compagnon psychologique qui aide l'utilisateur à surmonter ses crises et à partager ses sentiments en toute sécurité.

2.10 Techniques avancées pour améliorer l'interaction intelligente et la psychanalyse :

Dans le contexte de l'amélioration de la précision de l'interaction entre l'assistant intelligent et l'utilisateur, et en augmentant la capacité du système à comprendre plus profondément l'état psychologique, un ensemble d'algorithmes et de modèles intelligents supplémentaires ont été utilisés. Ces techniques permettent une interaction naturelle et une analyse psychologique basée sur plus que des mots.

2.10.1 Algorithme de distance Levenshtein :

L'algorithme de Levenshtein Distance est utilisé dans notre projet comme un moyen efficace pour mesurer la similarité entre les énoncés des utilisateurs et les questions stockées dans la base de données, en comptant le nombre de modifications requises (ajout, suppression ou substitution) pour convertir une phrase en une autre. Ce mécanisme est nécessaire pour permettre à l'assistant intelligent de comprendre différentes formulations du même sens, même si elles ne sont pas littéralement identiques. Par exemple, si l'utilisateur dit : « Je suis très triste », et qu'il y a une question similaire dans la base de données, comme « Je me sens très triste », l'algorithme détermine un pourcentage élevé de similarité entre les deux énoncés, ce qui permet au système de choisir la réponse appropriée la plus proche. De cette façon, une interaction plus intelligente et plus flexible avec les utilisateurs est assurée, en particulier dans des contextes psychologiques où les expressions sont très diverses

[19].

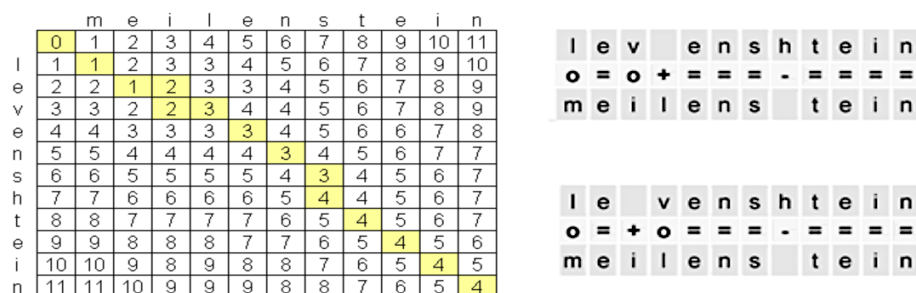


Figure 2.15 : Levenshtein Algorithmme

2.10.2 Analyse de l'activité physiologique à l'aide du MPU6050 :

Le capteur MPU6050 est un élément clé du projet, car il sert à suivre avec précision les mouvements de l'utilisateur en mesurant l'accélération et la rotation. Ces données permettent au système d'analyser le comportement moteur de l'utilisateur dans le but d'identifier des indicateurs liés à l'état psychologique, tels que le manque de mouvement qui peut refléter des symptômes de dépression ou d'isolement, ou une hyperactivité inhabituelle qui peut indiquer des états d'anxiété ou un trouble. Ces données sont envoyées en temps réel à Firebase, où elles sont stockées dans le journal personnel de chaque utilisateur, ce qui permet au psychiatre de suivre avec précision la progression de l'état et de prendre des décisions éclairées en fonction des mouvements quotidiens.

2.10.3 Analyse du signal audio (INMP441) :

Le microphone numérique INMP441 est utilisé dans ce projet non seulement comme un outil d'analyse de la parole, mais aussi comme un composant intelligent qui contribue à analyser les caractéristiques audio de l'utilisateur avec une grande précision. Des données importantes sont extraites, telles que le ton de la voix (aigu, bas, perturbé), le niveau du volume (haut, bas, oscillant) et la durée de la parole (lent, rapide, interrompu), qui sont des indicateurs comportementaux utilisés pour améliorer la compréhension par le système de l'état psychologique de l'utilisateur. Ces données sont combinées aux résultats du traitement du langage naturel (NLP), ce qui permet au clinicien de prendre des décisions plus précises et d'offrir des réponses conformes à l'état psychologique réel du patient.

2.10.4 Utiliser la détection de mot réveil (Porcupine) :

Pour améliorer l'expérience d'interaction avec l'assistant intelligent, le système Porcupine a été intégré pour détecter les "mots de réveil" tels que "Emily" ou "Jack", permettant à l'utilisateur d'activer automatiquement l'assistant vocal sans avoir à appuyer sur un bouton. Porcupine est un moteur open source qui fonctionne de manière native sur l'appareil, améliorant la confidentialité et assurant une réactivité même hors ligne. Il permet également de personnaliser les mots d'alerte en fonction de la personnalité de l'assistant, ce qui ajoute un caractère personnel et efficace à l'expérience utilisateur [20].

The logo for PICOVOICE, featuring the word "PICO" in a standard blue font and "VOICE" in a stylized blue font where the 'V' is formed by two overlapping shapes.

Figure 2.16: Pico voice Engine

2.11 Outils et technologies de soutien pour l'intégration de l'IA et l'interaction visuelle :

2.11.1 Google Colab – Cloud Training Environment :

Dans le développement d'un modèle de langage personnalisé basé sur l'architecture Mistral, la plate-forme Google Colab a été utilisée comme environnement de formation en nuage qui fournit des capacités élevées sans avoir besoin d'utiliser de puissantes ressources locales. Les fichiers de données d'entraînement (ensembles de données) ont été préparés et téléchargés sur Google Drive, puis le processus de réglage fin du modèle a été mis en œuvre au sein de Colab à l'aide des capacités GPU librement disponibles. La plate-forme se distingue par son support complet pour les bibliothèques d'intelligence artificielle telles que Transformers, TensorFlow et PyTorch, ce qui en fait un outil idéal pour expérimenter, la recherche et le développement de modèles d'intelligence artificielle dans un environnement éducatif ou de recherche flexible et intégré [21].



Figure 2.17 : Google Colab

2.11.2 Python – Le langage de programmation de base pour l'intelligence artificielle :

Aujourd'hui, Python est considéré comme le langage numéro un dans le domaine de l'intelligence artificielle grâce à sa facilité et ses puissantes bibliothèques telles que NumPy, Pandas, TensorFlow, PyTorch et Scikit-learn, en plus de son support avancé pour le traitement du langage naturel (NLP) et les techniques d'apprentissage profond. Dans notre projet, Python a été utilisé à toutes les étapes du travail, depuis le traitement des données, en passant par la formation aux modèles et l'analyse des résultats, jusqu'à la génération de réponses textuelles et la préparation de fichiers de prononciation phonétique à l'aide d'ElevenLabs, en plus d'écrire les scripts nécessaires pour lier le modèle de formation au service Firebase [22].



Figure 2.18 : Python

2.11.3 Avatar humain réaliste en 3D – Personnage interactif réaliste :

Afin d'offrir une expérience d'interaction plus humaine et naturelle, un personnage virtuel 3D (3D Human Avatar) a été intégré dans l'application, apparaissant pendant que l'assistant intelligent Jack ou Emily parle. Ce personnage se caractérise par des traits du visage réalistes qui conviennent au type d'assistant (homme ou femme), des expressions faciales qui sont en phase avec l'état émotionnel tel qu'un sourire ou une tristesse, en plus de bouger la tête et les mains en harmonie avec la voix. Ces personnages ont été créés à l'aide d'outils avancés tels que Motion Portait et Unity Avatar Systems. Cette intégration vise à améliorer le sentiment d'intégration psychologique avec l'assistant et à réduire le sentiment d'isolement ou d'interaction avec une machine, qui contribue à construire une plus grande confiance et un réel sens de la présence d'un personnage interagissant avec l'utilisateur et l'éco



Figure 2.19: Avatar humain réaliste en 3D.

2.11.4 Lip Sync (LipSync) – pour créer une interaction audio-visuelle :

Pour obtenir une interaction visuelle réaliste avec l'utilisateur, la technologie LipSync a été adoptée, ce qui signifie que le mouvement des lèvres du personnage virtuel correspond à la voix parlée. Dans notre projet, lorsque vous prononcez une phrase de l'assistant Jack ou d'Emily (par exemple : "Je suis ici avec toi..."), les lèvres 3D du personnage se déplacent précisément et de manière identique à la voix. Des outils tels que WavToLip ou Rhubarb Lip Sync ont été utilisés avec Unity Animator et les scripts de synchronisation audio dans Unity, avec un contrôle direct depuis l'application Android via 3D WebView. Cette intégration améliore le réalisme de l'interaction, car l'utilisateur sent que le personnage parle réellement, ce qui améliore l'expérience de l'utilisateur et renforce la réponse émotionnelle à l'assistant [24].



Figure 2.20 : Lip Sync (LipSync).

2.12 Conclusion:

Ce chapitre représente la base technique du projet, où tous les outils et méthodes utilisés pour construire le système intelligent ont été examinés.

Dans un premier temps, les concepts d'Intelligence Artificielle (IA) et d'Internet des Objets (IOT) sont introduits, et comment cette technologie contribue à la conception de solutions réalistes pour accompagner les patients psychiatriques.

Le projet s'appuyait sur Firebase en tant que plateforme de stockage basée dans le cloud pour gérer les conversations et les données vitales, et Android Studio pour développer une application mobile interactive utilisant Java et XML.

Du côté du traitement audio, VOSK et ElevenLabs ont été utilisés pour convertir la parole en texte et le texte en une voix humaine réaliste.

Le cœur du système analytique est le module de traitement du langage naturel (PNL) basé sur le modèle avancé Mistral, qui a été formé via Google Colab en utilisant Python et un ensemble de données réelles dérivées de cas psychologiques et de phrases colloquiales algériennes.

Pour améliorer l'efficacité, un système d'auto-apprentissage a été intégré qui permet au modèle d'évoluer progressivement en fonction des données de Firebase, tout en utilisant la technologie de réglage fin pour améliorer la compréhension linguistique et contextuelle.

Sur le plan visuel, un personnage 3D réaliste (3D Avatar) avec des mouvements faciaux naturels tels que des clignements d'œil et un léger sourire, ainsi qu'une synchronisation labiale (LipSync) avec la parole ont été intégrés, humanisant l'assistant vocal et rendant l'interaction plus chaleureuse et plus réaliste.

Ce chapitre montre comment de multiples technologies ont été intégrées dans un système qui fonctionne bien, donne à l'utilisateur une expérience riche et psychologiquement favorable, et ouvre de nouveaux horizons pour traiter les cas d'anxiété, la dépression et la dépendance en utilisant des moyens technologiques qui sont à la fois efficaces et humains.

Chapitre 3 Conception du système intelligent

pour accompagner les patients psychiatriques

3.1 Introduction :

Avec les avancées technologiques en matière d'intelligence artificielle (IA) et de l'Internet des objets (IoT), il est devenu possible de construire des systèmes intelligents qui fournissent de vraies solutions dans des domaines sensibles tels que la santé mentale. Dans ce chapitre, nous détaillerons tous les aspects appliqués développés dans le cadre de notre projet, du côté logiciel au côté physique, en fournissant des exemples, des codes, des images et des diagrammes qui expliquent comment fonctionne l'ensemble du système.

Cette conception vise à soutenir les patients psychiatriques d'une manière innovante basée sur l'interaction audio et vidéo et l'analyse en temps réel des conditions psychologiques. Il s'agit d'un système intégré qui combine des outils d'analyse audio, une intelligence artificielle, une interface graphique, des capteurs de mouvement et le stockage en nuage du statut de l'utilisateur, en plus d'une interface dédiée pour le médecin.

3.2 Structure générale du système :

Le système proposé est basé sur une structure architecturale composée de deux unités principales intégrées, qui fonctionnent en parfaite harmonie pour fournir un soutien psychologique complet à l'utilisateur, à savoir:

un appareil portable intelligent:(système1)

Cet appareil est porté au poignet sous forme d'un bracelet qui recueille des données physiologiques importantes relatives à l'utilisateur. Il comporte :

- Un Microphone numérique (INMP441) pour l'analyse du ton de voix, de l'intensité et des périodes silencieuses.
- Un Capteur de mouvement (MPU6050) pour surveiller l'activité motrice telle que la marche, la longue position assise, les secousses ou l'inactivité

-Un Écran OLED pour afficher les alertes rapides (exemple : « Votre activité est faible aujourd’hui »).

-Un Module ESP32 pour lire et envoyer des données à Firebase via WiFi.

Ce système fonctionne en continu en arrière-plan, sans interaction directe avec l'utilisateur, et envoie des données à chaque période de temps spécifiée, pour être analysées plus tard et reliées à l'état psychologique général.

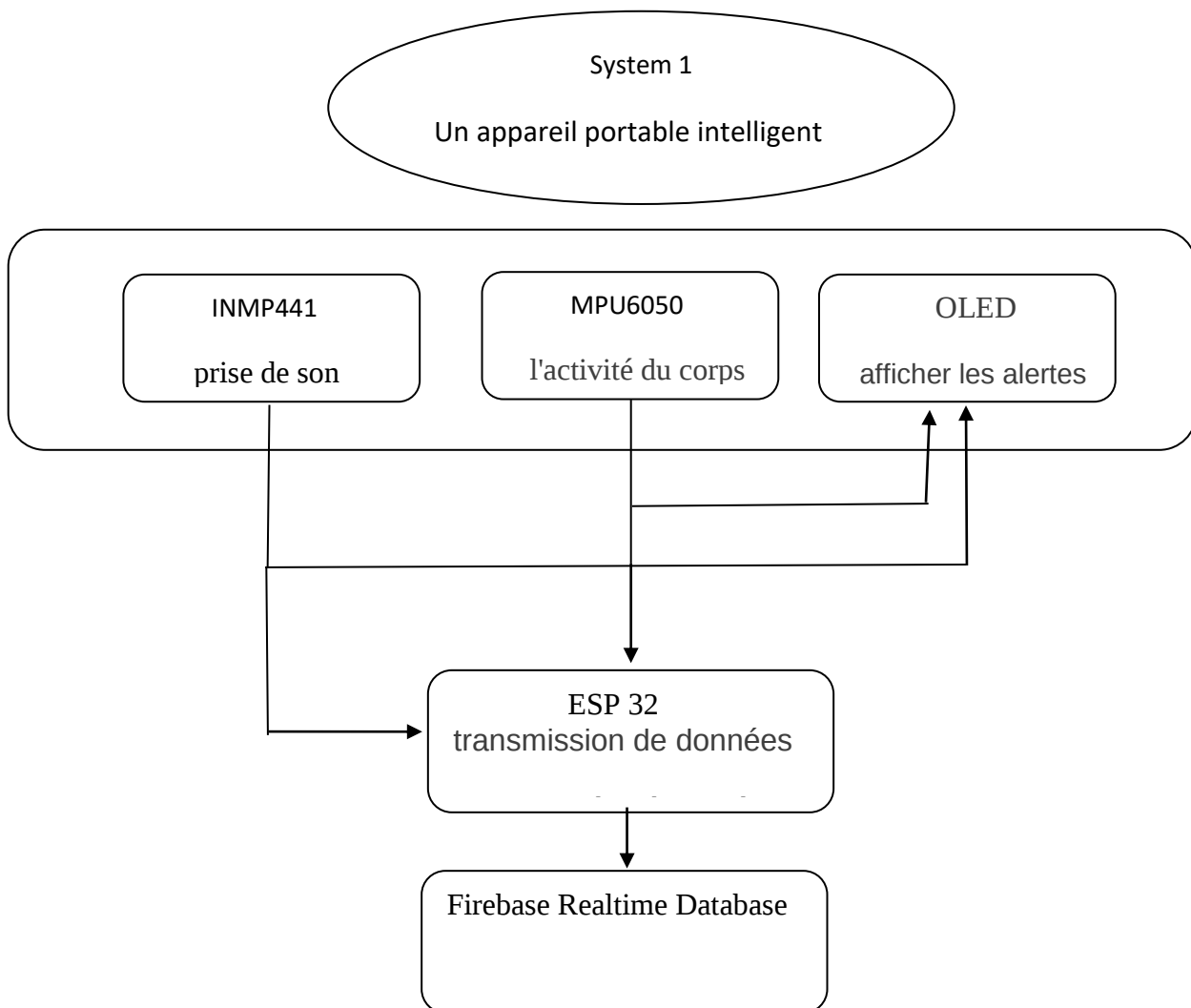


Figure 3.1 : Diagramme du System 1.

une application intelligente sur le téléphone (Android App : système2)

Il s'agit de l'interface utilisateur principale et contient toutes les fonctions interactives, notamment:

- Réveillez-vous en utilisant le mot de réveil (« Jack », « Emily ») ou un bouton.
- Convertir l'audio en texte à l'aide de VOSK et SpeechRecognizer.
- Analyse de texte via le modèle Mistral formé dans le dialecte local.
- Analyse du ton de la voix basée sur les données de Firebase.
- Afficher les réponses de manière réaliste via un personnage 3D (3D Avatar) avec LipSync.
- Fournir des exercices comportementaux cognitifs intégrés (TCC).
- La capacité d'envoyer un message ou texte au médecin.
- Une interface spéciale pour la prise de rendez-vous avec le médecin.
- Affichage de l'état psychologique quotidien (vert, jaune, rouge) basé sur des données combinées.

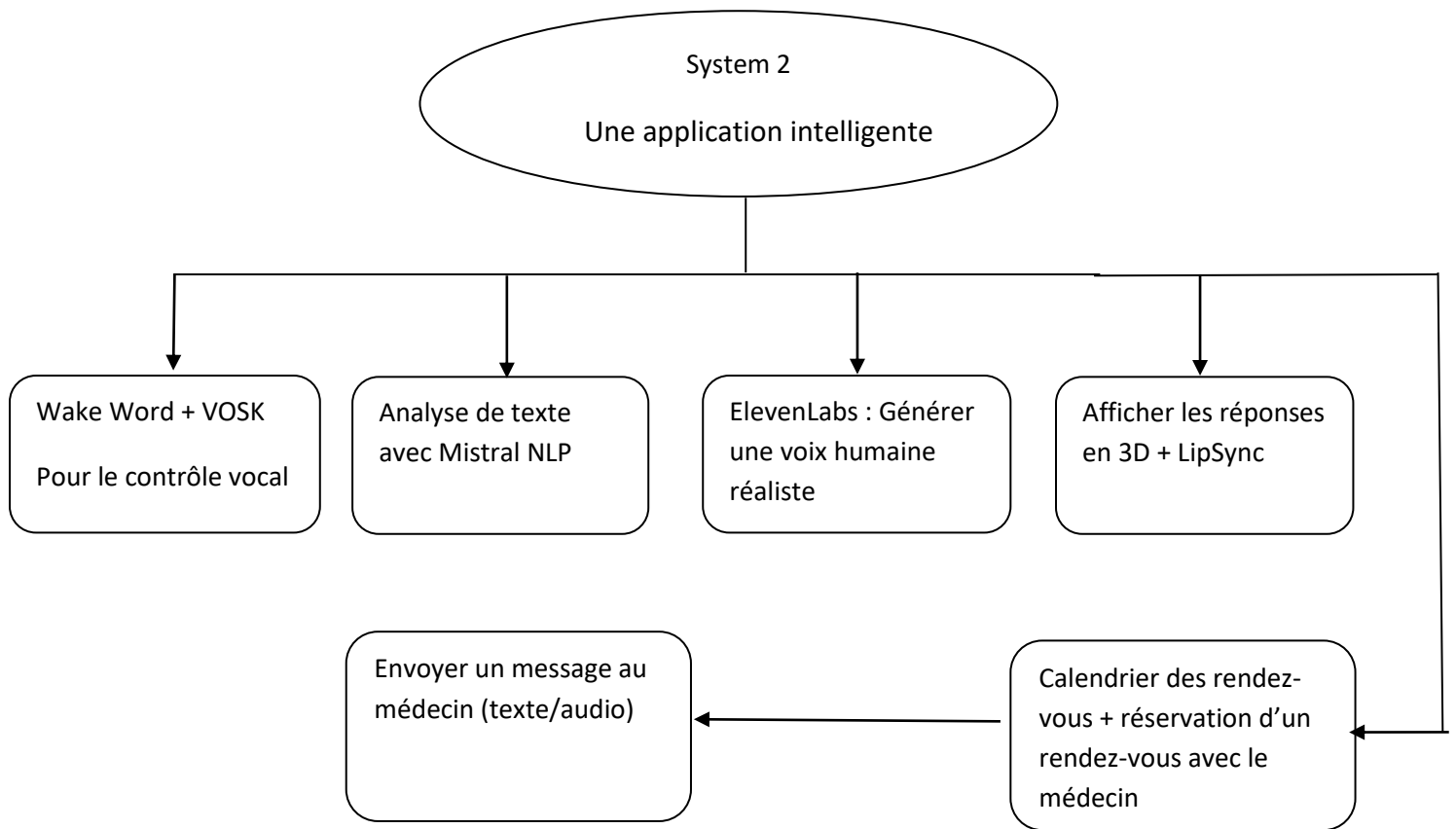


Figure 3.2 : Diagramme du System 2

Couche intermédiaire : Firebase Realtime Database :

Firebase représente l'épine dorsale de cette structure elle permet de :

- Stocker toutes les conversations et données en toute sécurité
- Recevoir des données du système portable
- Réception et envoi d'analyses de cas psychologiques provenant du PNL et de l'IA
- Activer les alertes au besoin (par exemple : une situation de danger récurrent)
- Faciliter la communication du médecin avec le patient

3.3 Composants logiciels :

3.3.1 Application Android :

L'interface de l'application a été développée en utilisant Android Studio, avec le langage Java utilisé pour programmer les fonctions et la logique, et des interfaces conçues via des fichiers XML. L'écran est également divisé en modules interchangeables à l'aide de Fragments. L'application propose la réception de commandes vocales et leur conversion en texte à l'aide de la technologie VOSK, avec affichage audio et vidéo des réponses, enregistrement des conversations, ainsi que l'envoi et la synchronisation des données avec la base de données Firebase.

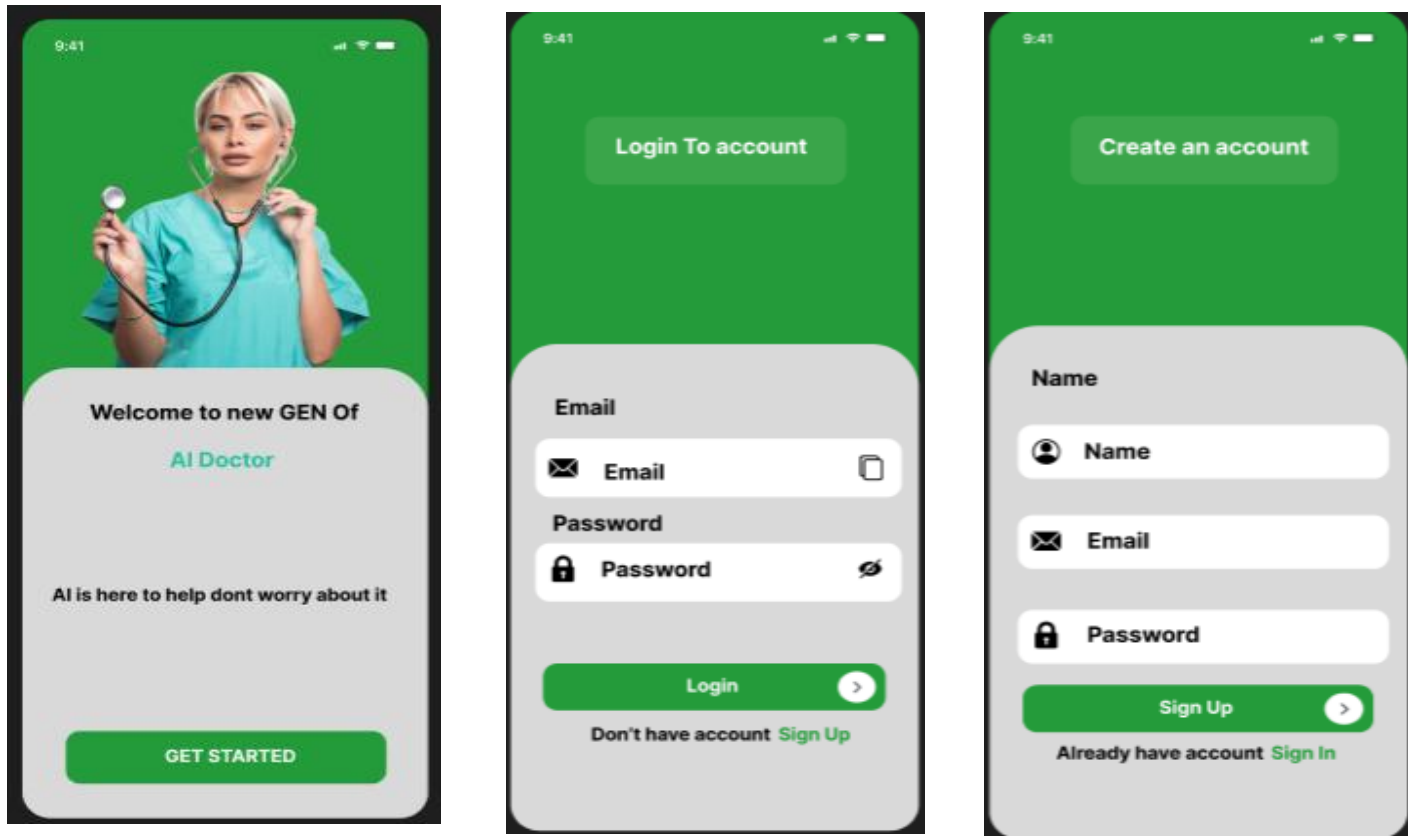


Figure 3.3: les interface Android App.

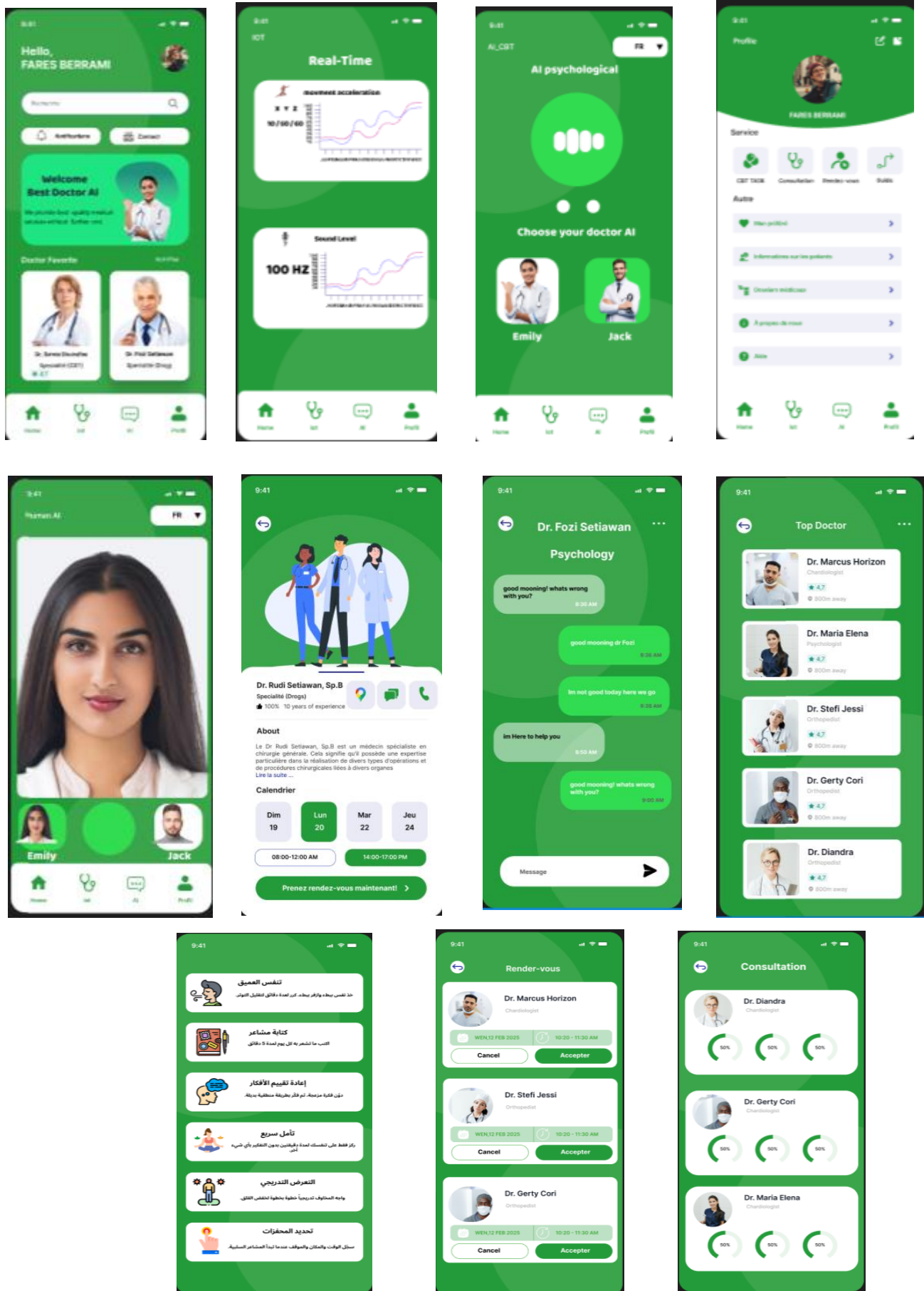


Figure 3.4 : les interfaces de l'Application

3.3.2 Structuration des fichiers dans la demande :

L'application intègre plusieurs classes clés qui assurent son bon fonctionnement et ses performances en temps réel. La classe **AudioStorageManager** est responsable de l'enregistrement et du stockage des fichiers audio localement sur l'appareil.

WakeWordManager permet la détection du mot d'activation (wake word) utilisé pour appeler l'assistant vocal à tout moment. Pour garantir la réactivité du système, même en cas d'urgence ou après un redémarrage du téléphone, la classe **BackgroundTaskWorker** permet à l'application de se relancer automatiquement en arrière-plan. Enfin, la classe **AvatarView** exploite les technologies **OpenGL**, **WebGL** et **SurfaceView à 30 FPS** pour afficher un avatar 3D réaliste, fluide et interactif, améliorant ainsi l'expérience utilisateur visuelle.

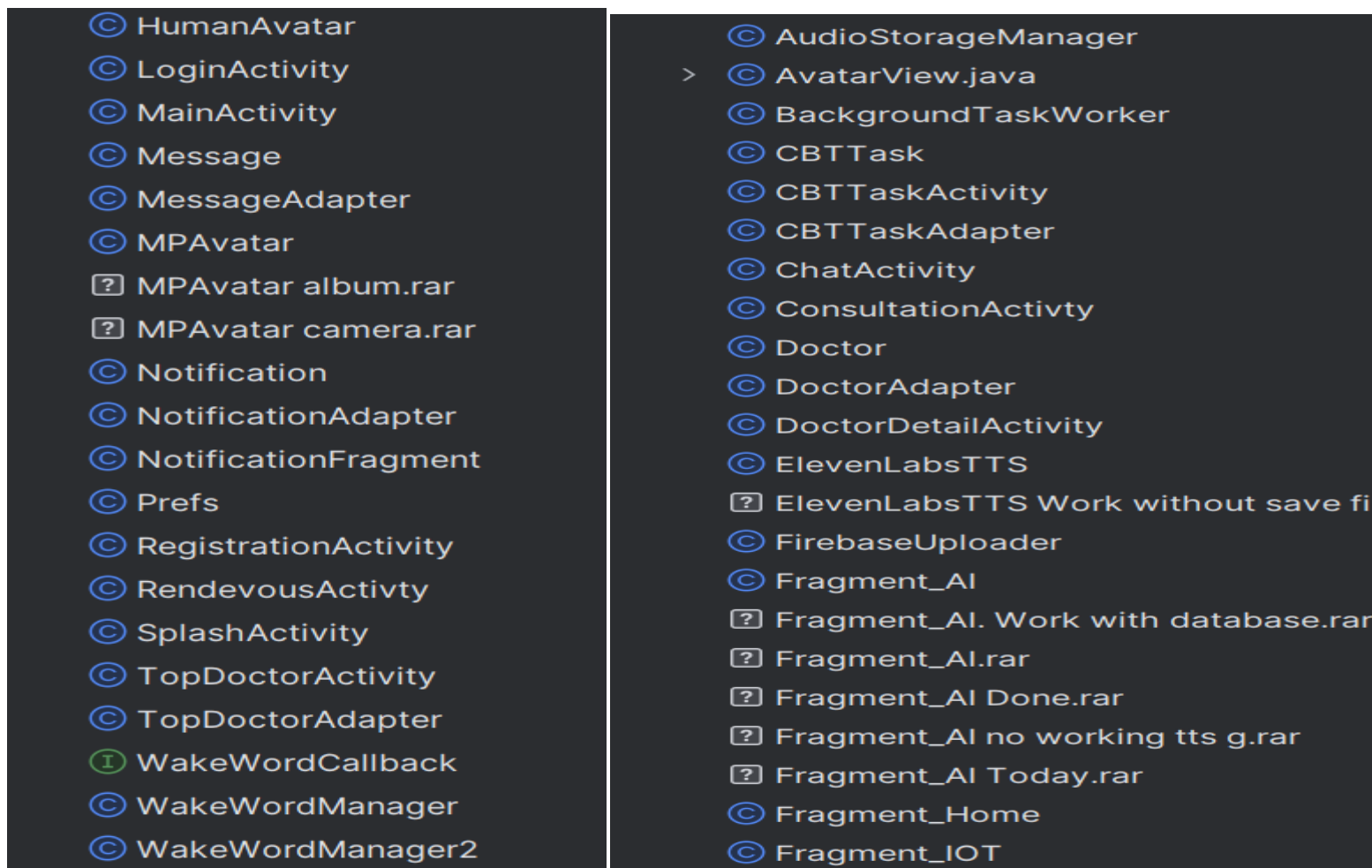


Figure 3.5 :Structure des Fichier

3.4 Intelligence artificielle et traitement du langage naturel (NLP) :

Au cœur du système intelligent se trouve un puissant module d'analyse de la parole humaine appelé Natural Language Processing (NLP), qui vise à transformer la parole ou le texte en données compréhensibles pouvant être classées et interprétées par une machine. C'est l'unité responsable de comprendre l'intention, les sentiments et le contexte du discours de l'utilisateur, puis de générer une réponse appropriée.

3.4.1 Modèle Mistral – Puissance analytique du système :

Le modèle Mistral 24B a été utilisé dans ce projet, et il s'agit de l'un des modèles avancés de LLM (Large Language Model) open source, formé sur des millions de phrases humaines pour assurer une grande efficacité et précision dans la compréhension et la réponse. L'une de ses principales caractéristiques est sa forte capacité à comprendre le contexte avec précision, en plus de son soutien au processus de réglage Fine-Tuning, ce qui lui permet d'être modifié et personnalisé en fonction des besoins du projet. Il est également possible de l'exécuter localement sur des appareils utilisant le format TF Lite, ce qui le rend adapté aux applications hors ligne. Le modèle est capable de comprendre et de produire l'arabe, le français ainsi que le dialecte algérien, ce qui le rend idéal pour les applications d'IA locales axées sur l'utilisateur.

3.4.2 Fine-Tuning Custom :

Nous avons créé un fichier de données personnalisé contenant un ensemble de phrases factuelles tirées des expériences de patients psychiatriques, soigneusement catégorisées à la main pour couvrir un large éventail de conditions psychiatriques. Des phrases dans le dialecte algérien ont également été incluses pour s'assurer que le modèle est adapté au contexte local, en plus des réponses thérapeutiques suggérées basées sur la méthode CBT (Cognitive Behavioral Therapy). En utilisant l'environnement Google Colab et la bibliothèque HuggingFace, le modèle Mistral est formé sur ce contenu pour être capable de comprendre et de comprendre des situations sensibles telles que anxiété, dépression, tristesse profonde, demander de l'aide, et exprimer un effondrement, En plus de traiter avec des appels d'intervention critiques tels que : "Nous ne pouvons pas continuer."

Timestamp (UTC)	Step	Train	Log Level	Step	Done	Acc (%)	Loss	Lr	Peak Alloc Mem (GB)	Alloc Mem (GB)
2024-05-24 19:22:20	000074	train	INFO	step: 000074	done	74.0	0.828	1.7e-05	peak_alloc_mem	22.2
2024-05-24 19:22:38	000075	train	INFO	step: 000075	done	75.0	0.908	1.6e-05	peak_alloc_mem	22.2
2024-05-24 19:22:56	000076	train	INFO	step: 000076	done	76.0	0.881	1.5e-05	peak_alloc_mem	22.2
2024-05-24 19:23:14	000077	train	INFO	step: 000077	done	77.0	0.819	1.4e-05	peak_alloc_mem	22.2
2024-05-24 19:23:32	000078	train	INFO	step: 000078	done	78.0	0.867	1.3e-05	peak_alloc_mem	22.2
2024-05-24 19:23:50	000079	train	INFO	step: 000079	done	79.0	0.913	1.2e-05	peak_alloc_mem	22.2
2024-05-24 19:24:08	000080	train	INFO	step: 000080	done	80.0	0.826	1.1e-05	peak_alloc_mem	22.2
2024-05-24 19:24:26	000081	train	INFO	step: 000081	done	81.0	0.835	9.5e-06	peak_alloc_mem	22.2
2024-05-24 19:24:44	000082	train	INFO	step: 000082	done	82.0	0.854	8.6e-06	peak_alloc_mem	22.2
2024-05-24 19:25:02	000083	train	INFO	step: 000083	done	83.0	0.772	7.7e-06	peak_alloc_mem	22.2
2024-05-24 19:25:20	000084	train	INFO	step: 000084	done	84.0	0.813	6.8e-06	peak_alloc_mem	22.2
2024-05-24 19:25:38	000085	train	INFO	step: 000085	done	85.0	0.836	6.0e-06	peak_alloc_mem	22.2
2024-05-24 19:25:56	000086	train	INFO	step: 000086	done	86.0	0.866	5.3e-06	peak_alloc_mem	22.2
2024-05-24 19:26:14	000087	train	INFO	step: 000087	done	87.0	0.845	4.6e-06	peak_alloc_mem	22.2
2024-05-24 19:26:32	000088	train	INFO	step: 000088	done	88.0	0.777	3.9e-06	peak_alloc_mem	22.2
2024-05-24 19:26:50	000089	train	INFO	step: 000089	done	89.0	0.813	3.3e-06	peak_alloc_mem	22.2
2024-05-24 19:27:08	000090	train	INFO	step: 000090	done	90.0	0.841	2.7e-06	peak_alloc_mem	22.2
2024-05-24 19:27:26	000091	train	INFO	step: 000091	done	91.0	0.807	2.2e-06	peak_alloc_mem	22.2
2024-05-24 19:27:44	000092	train	INFO	step: 000092	done	92.0	0.807	1.7e-06	peak_alloc_mem	22.2
2024-05-24 19:28:02	000093	train	INFO	step: 000093	done	93.0	0.827	1.3e-06	peak_alloc_mem	22.2
2024-05-24 19:28:20	000094	train	INFO	step: 000094	done	94.0	0.816	9.8e-07	peak_alloc_mem	22.2
2024-05-24 19:28:38	000095	train	INFO	step: 000095	done	95.0	0.825	6.8e-07	peak_alloc_mem	22.2
2024-05-24 19:28:56	000096	train	INFO	step: 000096	done	96.0	0.794	4.4e-07	peak_alloc_mem	22.2

Figure 3.6 : Training Model Google Colab

3.4.3 Algorithmes d'analyse supplémentaires:

Nous avons ajouté au système un ensemble d'algorithmes de soutien pour améliorer la compréhension du modèle du contexte et de l'utilisateur, parmi lesquels Levenshtein Distance est utilisé pour corriger les erreurs dans des mots écrits ou prononcés de manière inexacte, ce qui contribue à améliorer la précision de l'interaction. Un algorithme Classification des intentions a également été inclus pour déterminer avec précision l'intention d'une phrase, qu'il s'agisse d'une demande de soutien, d'un avertissement, d'une plainte ou autre, ainsi que Analyse des sentiments pour analyser les émotions et les classer en émotions positives, des situations neutres, négatives ou même dangereuses. Toutes ces opérations sont effectuées localement sur l'appareil (On-Device AI) sans aucune connexion à des serveurs externes, ce qui garantit le plus haut niveau de confidentialité et une vitesse de réponse instantanée.

3.5 Analyse de la voix et du son :

Une voix transmet non seulement des mots, mais aussi des émotions... Par conséquent, une unité spéciale d'analyse de tonalité vocale a été développée en utilisant le microphone numérique INMP441.

3.5.1 Ce que nous analysons :

Le système est également basé sur l'analyse des signaux audio dans le cadre de l'évaluation de l'état psychologique de l'utilisateur. Amplitude est mesurée, lorsqu'une intensité sonore indique la possibilité de anxiété ou de stress, tandis que calme excessif est considéré comme un indicateur possible de tristesse ou de dépression. La durée de la parole et les pauses entre les phrases sont également prises en compte, car ces données sont utilisées comme indicateurs pour évaluer l'état psychologique

avec plus de précision, ce qui contribue à fournir des réponses plus sensibles et réalistes à l'état de l'utilisateur.

3.5.2 Algorithmes d'analyse :

Les résultats sont automatiquement envoyés à Firebase, et combinés avec l'analyse de texte pour une évaluation psychologique complète.

```
usage
private void classifySound(float peakFrequency, int noiseLevel) {
    String state;

    if (peakFrequency < 80) {
        state = "صامت";
    } else if (peakFrequency >= 100 && peakFrequency <= 2500) {
        state = "يتكلم";
    } else if (peakFrequency > 2600 && peakFrequency < 3900) {
        state = "يصرخ";
    } else if ((peakFrequency > 4000 && peakFrequency < 5000)) {
        state = "ضحك";
    } else if ((peakFrequency > 3500 && peakFrequency < 3900)) {
        state = "بكاء";
    } else if (peakFrequency > 9000) {
        state = "اضطراب نفسي";
    } else {
        state = "صامت";
    }

    updateSoundState(state);
}
```

Figure 3.7 : Algorithmes classification Sound Capteur INMP441

3.6 ESP32 et capteurs physiques (INMP441 + MPU6050) :

En plus de l'analyse textuelle et audio, des composants physiques ont été inclus pour lire les indicateurs physiologiques qui contribuent à une évaluation plus complète de l'état psychologique de l'utilisateur. Une carte ESP32 a été utilisée comme commande principale, et a été couplée à deux micro capteurs : INMP441 et MPU6050.

3.6.1 INMP441 – Microphone numérique HD :

Un microphone I2 S (Inter-IC Sound) a été utilisé pour capturer le son sous sa forme brute et en haute fidélité, tout en minimisant le bruit. Ce microphone recueille des informations importantes telles que l'intensité du son, la fréquence et le temps, qui sont des indicateurs essentiels pour comprendre l'état psychologique et émotionnel de l'utilisateur. Ces données sont envoyées directement à Firebase, où elles sont ensuite stockées et analysées dans le système d'analyse intelligente pour déterminer avec

précision l'état émotionnel, ce qui contribue à améliorer la réponse de l'assistant intelligent et à l'adapter à l'état de l'utilisateur.

```
// Microphone I2S (INMP441) Settings
#define I2S_WS 5
#define I2S_SD 4
#define I2S_SCK 18
```

Figure 3.8 : Lien électronique avec ESP32.

3.6.2 MPU6050 – Capteur de mouvement et d'équilibre :

Le capteur MPU6050 a été adopté dans le projet, qui est un capteur à 6 axes qui combine un gyroscope et un accéléromètre, permettant un suivi précis du mouvement de l'utilisateur. Dans notre contexte, ce capteur est utilisé pour mesurer activité motrice quotidienne et déterminer si l'utilisateur est actif ou inactif, en plus de surveiller les mouvements rapides qui peuvent indiquer un trouble du mouvement ou des conditions d'anxiété. Il est également utilisé pour détecter les périodes de rémission prolongée, qui sont des indicateurs possibles de dépression ou d'une baisse de l'état psychologique.

3.6.3 Écran OLED :

Un petit écran OLED de 128x64 pixels a été ajouté au système, afin d'afficher les données capturées par les capteurs directement et instantanément. Les valeurs du capteur de mouvement MPU6050 et du microphone INMP441 sont affichées sur cet écran, qui fournit une interface visuelle rapide pour les indicateurs suivants tels que l'activité motrice, l'intensité sonore et les changements temporels, sans avoir besoin d'ouvrir l'application, qui facilite le processus de suivi sur le terrain et d'évaluation de l'état psychologique.

3.6.4 Envoi de données à Firebase :

Toutes les 1 secondes, de nouvelles valeurs sont mises à jour dans la base de données en temps réel de Firebase, permettant une analyse immédiate et un suivi par le médecin.

```

void sendDataToFirebase() {
    String path = "/sensor_data";
    FirebaseJson json;
    json.set("acceleration_x", sensorData.acceleration_x);
    json.set("acceleration_y", sensorData.acceleration_y);
    json.set("acceleration_z", sensorData.acceleration_z);
    json.set("gyro_x", sensorData.gyro_x);
    json.set("gyro_y", sensorData.gyro_y);
    json.set("gyro_z", sensorData.gyro_z);
    // json.set("speed", sensorData.speed);
    json.set("Temp", sensorData.temperature);
    json.set("peakFrequency", sensorData.peakFrequency);
    json.set("noiseLevel", sensorData.noiseLevel);

    if (Firebase.RTDB.setJSON(&fbdo, path, &json)) {
        Serial.println("✅ Donner envoyer ");
    } else {
        Serial.print("❌ failed envoyer des donner ");
        Serial.println(fbdo.errorReason());
    }
}
}

```

Figure 3.9 : Send Data FireBase .

3.7 Personnage virtuel 3D (Avatar humain 3D) et synchronisation des lèvres (LipSync) :

Pour offrir une expérience humaine interactive proche de la réalité, un personnage virtuel en 3D représentant l'assistant intelligent (Jack ou Emily) a été adopté, qui interagit avec l'utilisateur non seulement vocalement, mais aussi visuellement par le biais d'expressions faciales, de mouvements des lèvres, de clignements et même d'un sourire.

3.7.1 L'importance de l'interaction visuelle en psychothérapie :

Des études psychologiques indiquent que l'expression visuelle contribue de façon significative à améliorer les sentiments de réconfort et de satisfaction, en particulier pendant les séances de thérapie et de soutien psychologique. En conséquence, un caractère visuel interactif a été adopté dans l'application, car cet élément contribue à approfondir le sentiment d'inclusion de l'utilisateur et à renforcer la relation émotionnelle avec l'assistant intelligent, créant ainsi un environnement plus chaleureux et plus humain. La présence de ce personnage stimule également une interaction et un engagement accrus dans le dialogue, et encourage l'utilisateur à s'exprimer et à participer plus ouvertement et confortablement.

3.7.2 Outils de conception et d'animation des personnages :

Le personnage a été créé à l'aide des outils suivants :

Outie	Function
Own APP SDK 3D	Pour créer une apparence réaliste du personnage (masculin ou féminin).
Fares SDK MontionPortrit	Pour importer et animer le personnage.
LipSync	Pour synchroniser le mouvement des lèvres avec le texte parlé.
Animator	Pour contrôler le sourire, les clignements et les expressions faciales.
WebGL	Pour afficher le caractère dans l'application.

Tableau 3.1 : conception et d'animation des personnages.

3.7.3 Caractéristiques du mouvement de caractère :

Le personnage visuel de l'application a été soigneusement conçu pour incarner une interaction humaine réaliste et précise, ses mouvements étant parfaitement synchronisés avec les événements audio et contextuels. Ces mouvements comprennent clignement des yeux spontané toutes les 5 à 7 secondes pour simuler un comportement normal, et un léger sourire qui accompagne des phrases rassurantes pour favoriser un sentiment de confort. LipSync a également été adopté pour assurer le mouvement des lèvres avec une précision extrême avec le son, en plus de une légère inclinaison de la tête pendant l'interaction, qui ajoute une dimension émotionnelle qui renforce le réalisme de l'expérience et renforce le lien entre l'utilisateur et l'assistant.

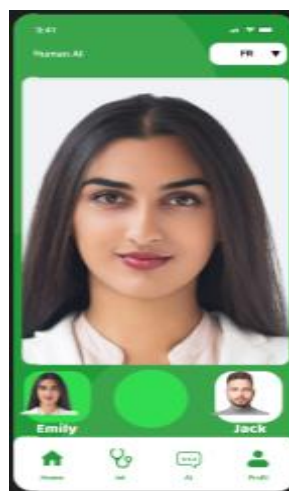


Figure 3.10 : caractère 3D

3.7.4 Comment fonctionne le lip sync :

Lors de la génération d'audio via le service ElevenLabs, le clip est d'abord enregistré au format MP3 puis transmis à un script de synchronisation des lèvres, où la fréquence audio est analysée et reliée aux commutateurs de mouvement des lèvres appropriés (phonèmes). La synchronisation est ensuite appliquée avec précision dans l'application en utilisant Audio Source pour s'assurer que le son correspond au mouvement de la bouche. La hauteur et la vitesse de la voix peuvent également être ajustées en fonction du contexte de la phrase... Par exemple, le tempo est élevé et la voix est chargée d'énergie lorsqu'elle exprime de la joie, ou ralentie et adoucie lorsqu'elle fait preuve de sympathie et de calme, ce qui donne au personnage une voix dynamique capable de transmettre des sentiments plus réalistes.

3.7.5 L'impact de ce caractère sur l'expérience utilisateur :

Le système a été piloté sur un échantillon de 10 utilisateurs afin d'évaluer l'impact du caractère virtuel sur l'expérience d'interaction. Les résultats ont montré que 90 % des participants estimaient que la présence du personnage rendait l'interaction "plus réaliste", et 80 % ont exprimé qu'ils se sentaient plus à l'aise psychologiquement lorsqu'il y avait un élément visuel interagissant. Ce qui est remarquable, c'est que 60% des utilisateurs ont signalé qu'ils oubliaient parfois que le système n'est pas un être humain réel, ce qui reflète l'efficacité de la conception visuelle et comportementale du personnage. Par conséquent, la personnalité virtuelle ne se limite pas à l'aspect esthétique, mais représente plutôt un élément central dans l'expérience thérapeutique, contribuant à créer un environnement humain chaleureux qui améliore l'efficacité de la communication et du soutien psychologique.

3.8 Système de sécurité et de protection de la vie privée :

Étant donné que ce projet traite des informations hautement sensibles liées à l'état psychologique et de santé de l'utilisateur, une série de mesures techniques et éthiques ont été adoptées pour assurer la protection de ces données et prévenir toute violation ou utilisation illégale.

3.8.1 Authentification sécurisée de l'utilisateur et du médecin (Authentification) :

Le système s'appuie sur Firebase Authentication pour sécuriser le processus de connexion tant pour le patient que pour le médecin, où l'e-mail et un mot de passe chiffré avec l'algorithme SHA256 sont utilisés pour assurer la sécurité. La vérification à deux facteurs (2FA) a également été activée pour le médecin au besoin, afin d'améliorer la protection, en particulier dans les cas d'accès à des données sensibles. Les sessions sont gérées par Expiration du jeton qui expire automatiquement après une période donnée pour empêcher l'accès non autorisé. Si une tentative d'accès suspecte ou non autorisée se produit, elle est rejetée immédiatement et automatiquement documentée dans les journaux Firebase, ce qui permet de suivre l'activité et d'assurer l'intégrité des données.

3.8.2 Chiffrement des données :

Les données du système sont protégées par plusieurs couches de protection. Toutes les informations stockées dans la base de données en temps réel de Firebase sont soumises au cryptage automatique, ce qui garantit que le contenu est protégé même si l'on accède à la base de données elle-même. Le protocole HTTPS est utilisé pour sécuriser la communication entre l'application et Firebase, empêchant toute interception ou falsification lors de la transmission des données. Les fichiers audio ne sont pas non plus stockés localement sur les appareils des utilisateurs ou des médecins, ce qui réduit le risque d'accès non autorisé. Ainsi, toutes les conversations, évaluations, statuts et activités sont transmis et stockés en toute sécurité, et ne peuvent être consultés par des tiers en aucune circonstance.

3.8.3 Politique de contrôle d'accès :

Role	Permission
Patient	Voir seulement ses conversations et résultats
doctor	Uniquement sur ses patients inscrits
développeur	Il n'a pas accès au contenu ou aux conversations psychologiques

Tableau 3.2 : contrôle d'accès.

3.8.4 Protéger les fichiers audio et les textes :

Le système est basé sur le principe de la confidentialité maximale, car aucune voix ou texte n'est stocké sur le téléphone en permanence. Les réponses vocales sont chargées dans le cache uniquement et sont supprimées immédiatement après l'affichage pour éviter que des données inutiles ne restent sur l'appareil. Dans les cas où la mise en cache est nécessaire (comme une mauvaise connectivité), Encrypted SharedPreferences, un mécanisme crypté et sécurisé alimenté par Android, est utilisé pour assurer que les données restent complètement protégées même pendant un court moment sur le téléphone.

3.8.5 Contrôle de la confidentialité des patients :

L'utilisateur (patient) a le plein contrôle sur la confidentialité de ses données, conformément aux principes du "droit à l'oubli". Où il peut à tout moment demander que toutes ses données soient définitivement supprimées de Firebase sans aucune

possibilité de récupération. Il peut également mettre en pause l'analyse en temps réel, arrêtant instantanément la collecte et l'analyse de données sensorielles ou audio. Pour assurer son confort psychologique, on lui donne la possibilité de cacher son identité au médecin en utilisant un code d'identification à la place de son vrai nom lors des séances ou en interagissant avec le système, ce qui renforce son sentiment de sécurité et encourage la libre expression.

3.8.6 Politique de mise à jour et de suivi :

Le système s'appuie sur Firebase Logs pour documenter automatiquement chaque modification ou mise à jour effectuée dans l'application, assurant ainsi un haut niveau de transparence. Le médecin et le patient peuvent consulter le journal des modifications par l'intermédiaire de ce que l'on appelle les journaux de transparence, qui permettent de suivre toute activité ou modification apportée aux données ou aux sessions. Il est important de noter que toutes les données restent dans le système et ne sont pas partagées avec une plate-forme externe, ni utilisées à des fins publicitaires, ce qui renforce la confiance et souligne l'engagement du projet à protéger la vie privée des utilisateurs.

3.9 Diagrammes finaux et tableaux récapitulatifs :

Dans cette section, nous examinons un ensemble de diagrammes et de tableaux récapitulatifs qui aident à comprendre le fonctionnement du système, la structure de ses composants et l'interaction des éléments entre eux.

3.9.1 Diagramme de flux de données dans le système :

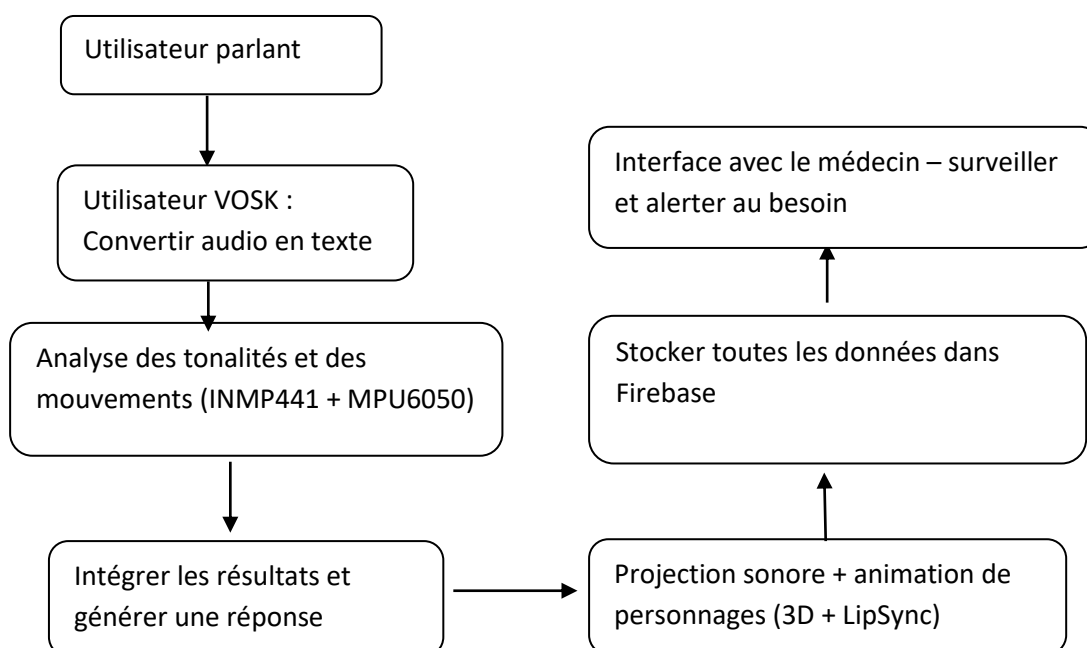


Figure 3.11 : Diagramme de flux de données .

3.9.2 Tableau du rôle de chaque composant :

Composant	Role
ESP32	Il recueille les données des capteurs et les envoie à Firebase
INMP441	Capte le ton et l'intensité de la voix
MPU6050	Mesure l'activité physique de l'utilisateur
Firebase	Stocke tout (texte, son, mouvement, analyse)
Android App	Interagit avec l'utilisateur et affiche les résultats
VOSK	Convertit l'audio en texte
Mistral	Analyse le texte et identifie les sentiments et l'intention
ElevenLabs	Transforme la réponse en une voix humaine réaliste
3D Avatar	Il montre le caractère et capture l'interaction visuelle
Docteur	Il examine la situation et intervient au besoin

Tableau 3.3 : rôle de chaque composante .

3.9.3 Structure de l'application Android :

■ app

└─ AudioManager.java	└─ MainActivity.java
└─ AvatarView.java	└─ NotificationFragment.java
└─ BackgroundTaskWorker.java	└─ RegistrationActivity.java
└─ CBTTTaskActivity.java	└─ SplashActivity.java
└─ ChatActivity.java	└─ TopDoctorActivity.java
└─ DoctorDetailActivity.java	└─ RendezvousActivity.java
└─ ElevenLabsTTS.java	└─ LoginActivity.java
└─ Fragment_AI.java	└─ WakeWordManager.java
└─ Fragment_Home.java	
└─ Fragment_IOT.java	
└─ Fragment_Setting.java	
└─ Human.java	

3.9.4 Dessin de connexion des composants physiques (schéma de circuit) :

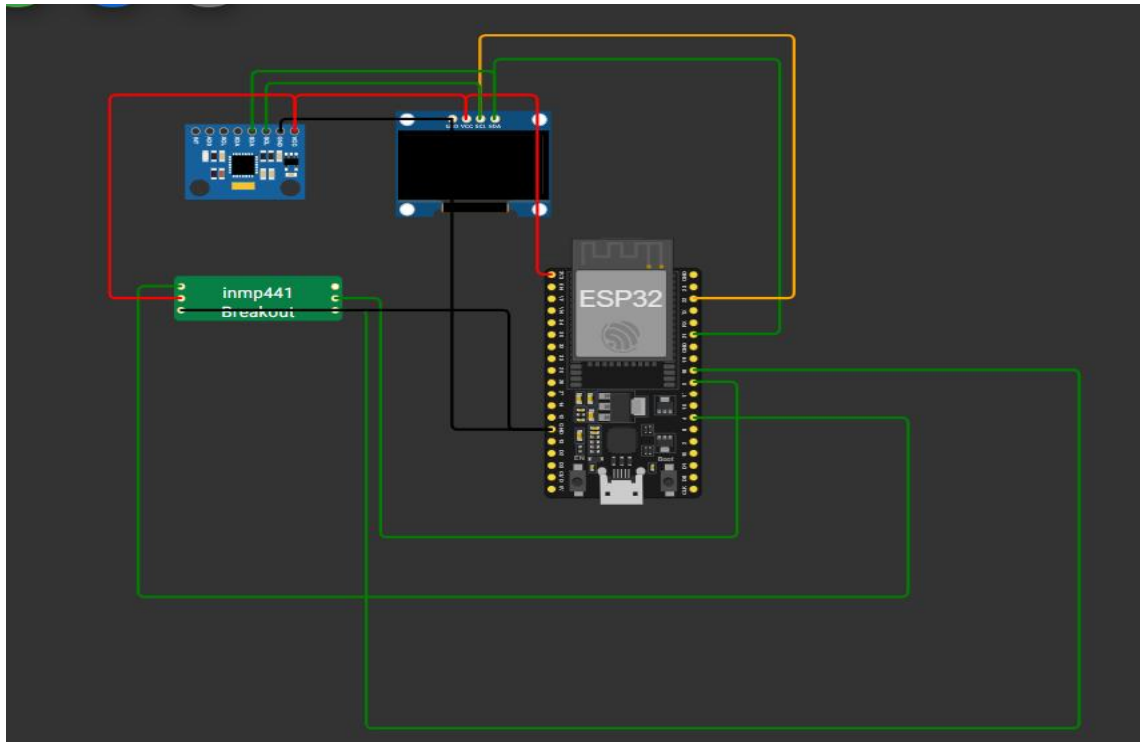


Figure 3.12 : schéma de circuit.

3.10 Contribution des appareils intelligents à la santé mentale :

Au fil des années, les appareils intelligents sont devenus un outil efficace dans le domaine des soins psychologiques, en particulier grâce à l'avènement des micro capteurs et des microprocesseurs. L'intégration d'un matériel intelligent avec une analyse intelligente des données est une étape révolutionnaire pour comprendre la psychologie des utilisateurs de manière plus précise et réaliste.

3.10.1 Présentation de l'ESP32 :

L'ESP32 est un microcontrôleur puissant et de faible puissance qui prend en charge la connectivité sans fil via Wi-Fi et Bluetooth, ce qui le rend idéal pour connecter des capteurs et envoyer des données à des systèmes intelligents.

Dans ce projet, l'ESP32 connecte le capteur de son INMP441 et le capteur de mouvement MPU6050, puis envoie les données à Firebase pour analyse par le système intelligent, permettant ainsi de surveiller instantanément et avec précision l'état psychologique de l'utilisateur.

ESP32 Wroom DevKit Full Pinout

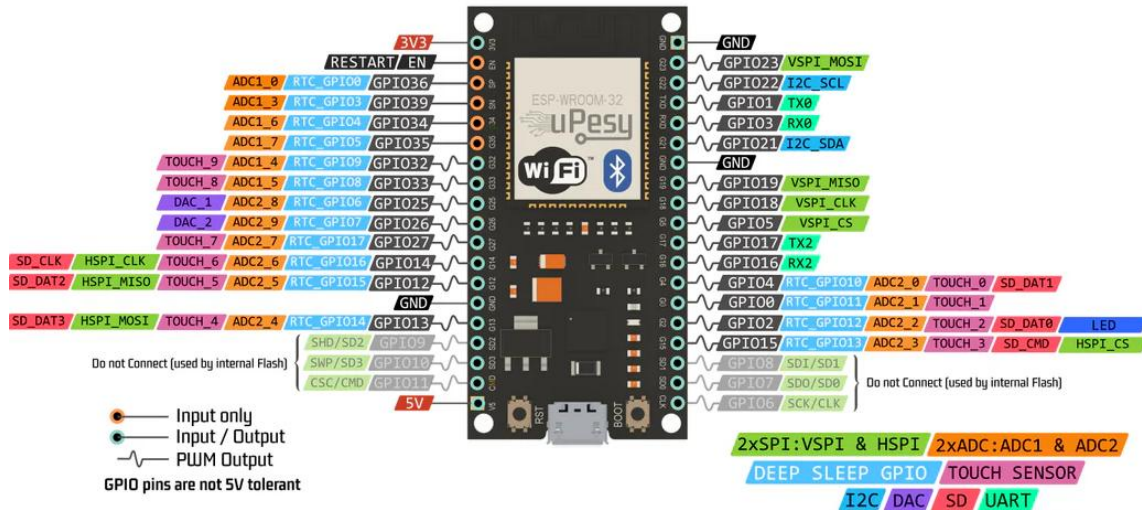


Figure 3.13 :ESP32

3.10.2 Le Capteur sonore haute résolution INMP441 :

Le capteur INMP441 est un microphone numérique haute résolution utilisé dans les applications nécessitant une analyse audio précise, et envoie des données directement via l'interface I2S, permettant une excellente qualité de détection et une intégration efficace avec les systèmes embarqués tels que ESP32. Dans ce contexte, le microphone est utilisé pour analyser l'état émotionnel de l'utilisateur à travers le ton, l'intensité et la fréquence de sa voix. Les caractéristiques vocales peuvent être liées à divers indicateurs psychologiques :

Une voix faible peut indiquer des cas de dépression

Bien qu'un son aigu et rapide puisse indiquer de l'anxiété,

Un son interrompu peut refléter une tension ou une hésitation dans l'expression.

Cette analyse contribue à fournir des réponses intelligentes et personnalisées au sein d'applications de santé mentale et d'assistants vocaux intelligents.

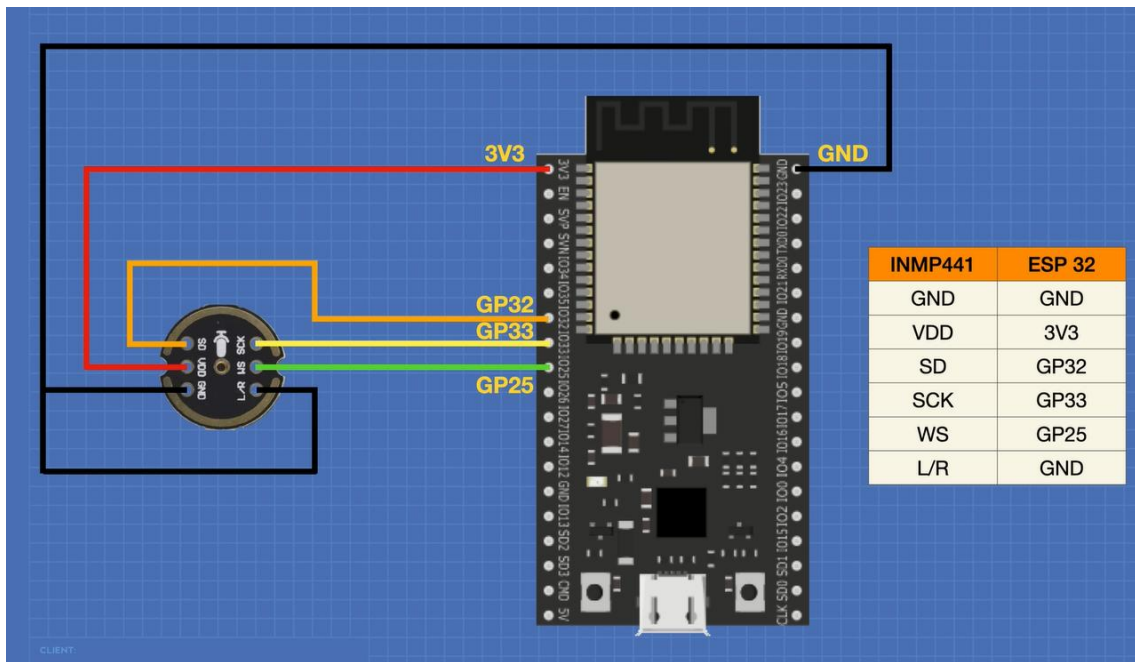


Figure 3.14: INMP441

3.10.3 Le Capteur de mouvement et d'équilibre MPU6050 :

Le capteur MPU6050 est une unité de détection multifonctionnelle qui comprend un accéléromètre et un gyroscope, ce qui lui permet de suivre les mouvements du corps et d'équilibre avec une grande précision. Ce capteur est utilisé pour analyser le comportement moteur de l'utilisateur, qui est un indicateur important de l'état psychologique. Par exemple, le manque de mouvement peut indiquer la léthargie ou la dépression, tandis que les mouvements excessifs et soudains indiquent l'anxiété ou le stress. Le capteur permet également de détecter des positions anormales en position assise ou endormie qui peuvent être liées à des troubles psychologiques ou physiques. Lorsque l'analyse des mouvements est combinée avec les données audio extraites du microphone INMP441, le système peut créer une perception plus complète et précise de l'état psychologique de l'utilisateur, améliorant ainsi l'efficacité de l'évaluation intelligente et du soutien comportemental fourni.

MPU6050 Hardware Connection with ESP32

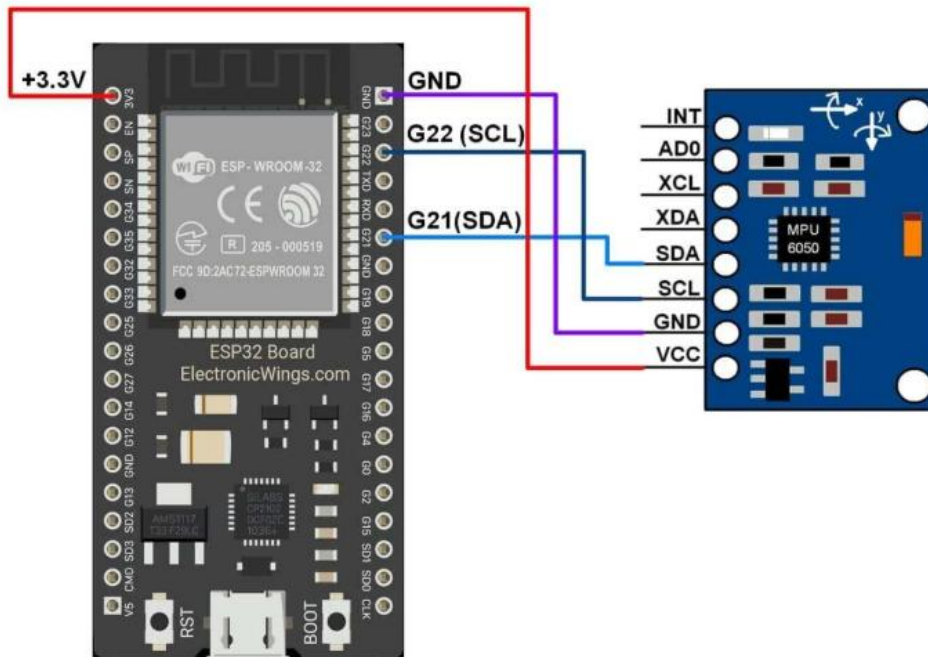


Figure 3.15 : MPU6050

3.10.4 L'Écran OLED d'affichage de l'information :

Le petit écran OLED est un élément essentiel de l'interface utilisateur, fournissant un affichage visuel direct et facile à comprendre des informations importantes. Avec sa clarté élevée et sa faible consommation d'énergie, cet écran est utilisé pour afficher l'état psychologique récemment détecté basé sur des analyses de voix et de mouvement, ainsi que le temps, les alertes et l'état de la connexion réseau. Il permet également l'affichage de messages instantanés importants qui dirigent ou alertent l'utilisateur lorsque cela est nécessaire, ce qui rend l'interaction avec le système plus fluide et plus efficace dans les environnements quotidiens.

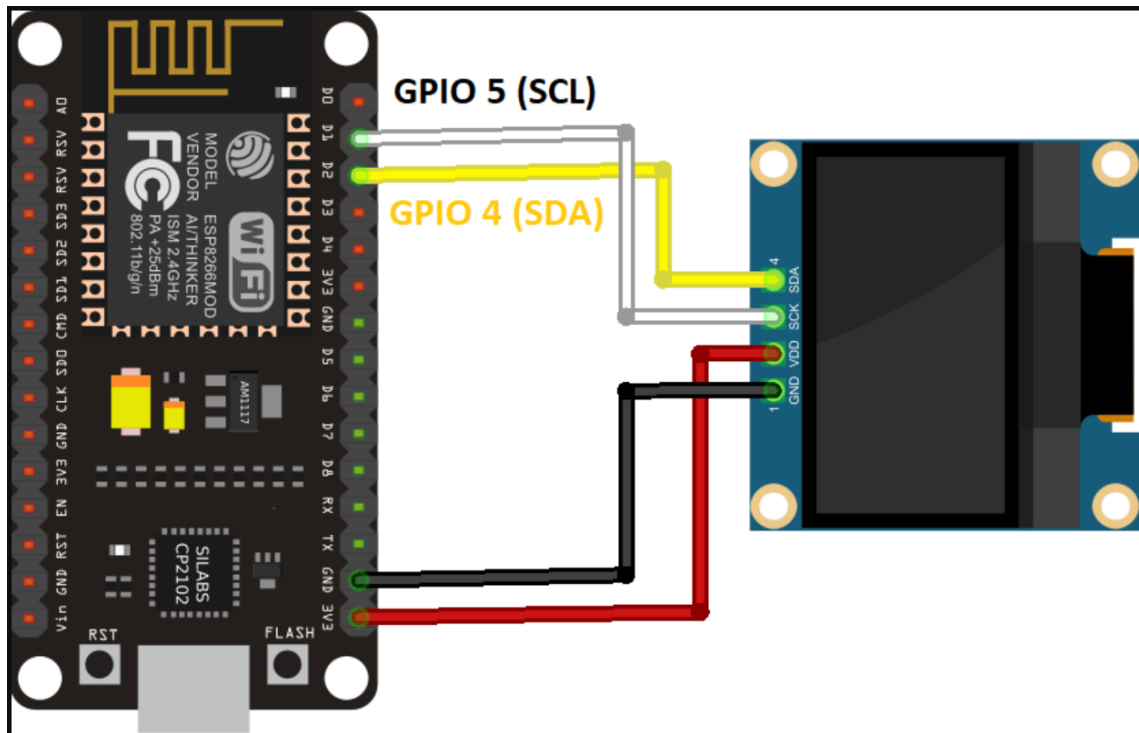


Figure 3.16 : Ecran OLED.

3.11 Conclusion :

Dans ce chapitre, nous avons présenté en détail les aspects appliqués et techniques du projet, tout en abordant la conception pratique des composants physiques et logiciels du système. Il explique comment les capteurs physiques tels que l'INMP441 et le MPU6050 s'intègrent au module ESP32 pour lire et analyser les données physiologiques de l'utilisateur, telles que la voix et l'activité motrice, puis les envoyer à la base de données cloud Firebase en temps réel.

Nous avons également discuté de la construction d'une application Android professionnelle qui représente l'interface interactive pour le patient, et comprend toutes les fonctions intelligentes de l'éveil vocal, la conversion de la parole en texte, l'analyse des phrases dans la langue de l'utilisateur via le modèle PNL Mistral, en plus de générer une voix humaine réaliste à l'aide d'ElevenLabs. Nous avons ajouté une couche interactive visuelle représentée par un personnage tridimensionnel (3D Avatar) équipé de la technologie LipSync pour améliorer la qualité de la communication humaine entre l'utilisateur et l'assistant.

En outre, l'interface du médecin a été présentée, ce qui lui permet de suivre l'état du patient, d'examiner ses conversations et d'analyser les indicateurs de voix et de mouvement, ce qui ouvre de nouveaux horizons pour le diagnostic à distance et le suivi.

Chaque composante de ce système remplit une fonction claire au sein d'une structure architecturale intégrée, ce qui fait de ce projet une plate-forme prometteuse pour des solutions intelligentes de soins psychologiques, surtout dans des contextes où il y a pénurie de personnel médical ou difficulté d'accès à un soutien psychologique direct.

Cette combinaison de l'intelligence artificielle, de l'Internet des objets et de l'interaction humaine fait du projet une véritable étape vers l'avenir de la psychothérapie numérique.

Chapitre 4 Tests, résultats et analyse des performances

4.1 Introduction :

Ce chapitre consiste à évaluer l'efficacité du système proposé en effectuant une série de tests appliqués et réalistes sur tous ses composants, que ce soit en termes de performance, de précision, de rapidité de réponse ou d'interaction avec l'utilisateur. Les résultats des tests du système dans de multiples conditions seront également affichés, en plus d'une évaluation réelle par différents utilisateurs, y compris les médecins et les patients. Nous analysons ces résultats pour comprendre les forces et les faiblesses, et suggérer des améliorations futures.

4.2 Environnement d'essai :

Les tests ont été effectués dans un environnement quasi réaliste simulant les conditions réelles d'utilisation de l'utilisateur, en utilisant une combinaison de dispositifs techniques et de services intégrés:

4.2.1 Les composants techniques:

Les composants suivants ont été utilisés :

Un Smartphone : Samsung S9 fonctionnant sous Android 14.

Un Contrôleur : ESP32 DevKit V1.

Un Microphone : INMP441 connecté via l'interface I2S pour la capture l'audio brut.

Un Capteur moteur : MPU6050 connecté via I2C pour mesurer l'accélération et la rotation.

Une Base de données : Firebase Realtime Database pour le stockage et la récupération instantanés des données.

Une Connectivité : 20 Mbit/s WiFi à domicile.

Un Modèle intelligent : Mistral 24 B formé avec Google Colab et converti en TFLite pour l'exploitation locale.

La Prononciation vocale : Intégrer ElevenLabs audio via des fichiers MP3 locaux.

L’Affichage visuel : Un personnage 3D (3D Avatar) intégré dans WebGL à l’aide du module SDK .

4.2.2 Les participants à l’essai :

6 patients ont participé volontairement pour évaluer l’expérience psychologique.
3 psychiatres observent la réaction et fournissent une rétroaction thérapeutique.
2 ingénieurs techniques pour tester la stabilité de l’application et vérifier les performances du logiciel.

4.3 Tests de performance pour chaque composant du système :

Dans cette section, la performance de chaque composant est évaluée individuellement, en termes de vitesse, précision, stabilité et réponse du système pendant l’utilisation réelle. L’objectif est de s’assurer que chaque partie fonctionne efficacement et contribue à fournir une expérience de traitement intelligente et sécuritaire.

4.3.1 Performance des applications Android :

Le test a comporté sur plusieurs aspects techniques pour assurer une performance efficace et une utilisation fluide, l’accent étant mis sur le temps de réponse de l’interface utilisateur en termes d’ouverture des écrans et de déplacement entre les différentes sections, en plus de surveiller la consommation de mémoire pendant l’exécution de l’application pendant de longues périodes. La stabilité de l’application a également été évaluée en cas de travail continu, afin de surveiller tout dysfonctionnement ou fuite dans les ressources. Enfin, la performance du module WebView responsable de la présentation du personnage animé en 3D a été analysée, afin d’assurer la fluidité du mouvement et sa synchronisation avec le son sans clipping ni délai.

Indice	valeur
Temp de démarrage Initiale	2.3 seconde
Temps d’ouverture interface de communication	< 1 seconde
Capacité d’utilisation Ram Moyen	148 MB
Stabilité après 30 min Marche	Le crash ne s’est pas produit
Temps d’affichage de la voix et des caractères	1,2 seconde (moyenne)

Tableau 4.1 : Performances des applications.

4.3.2 Performances de l'assistant vocal (Wake Word + VOSK + TTS) :

Le temps de réponse global a été testé à partir du moment où l'utilisateur a prononcé le mot d'alerte "Emily" ou "Jack", jusqu'au début de la réponse vocale de l'assistant, pour assurer une interaction instantanée et naturelle. La précision de la conversion audio-texte a également été mesurée à l'aide du modèle VOSK, en mettant l'accent sur les trois langues prises en charge : le dialecte algérien, le français et l'anglais, dans le but d'assurer une compréhension correcte des différents dialectes et structures. En outre, la vitesse de génération de la voix dans le service ElevenLabs a été évaluée pour s'assurer que les réponses vocales sont fournies rapidement et sans retards notables qui affectent l'expérience utilisateur.

Indice	Valeur
temps de réponse Wake Word	0.8 Secound
Résolution VOSK pour convertir l'audio en texte	91% pour le français/87% pour l'algérien/100% pour l'anglais
Temps de conversion texte-voix (TTS)	1,4 seconde (moyenne)
LipSync + Avatar temps pour déplacer les lèvres	Instantanée (dans WebView)

Tableau 4.2 : Performances de l'assistant vocal.

4.3.3 Performance du modèle d'intelligence artificielle (NLP Mistral) :

Le système a été testé à des niveaux d'analyse de phrases et d'extraction d'émotion en utilisant des modèles TFLite locaux, garantissant la vitesse et l'absence de connexion à des serveurs externes. L'exactitude de la cote d'intention a également été mesurée pour déterminer si l'utilisateur demande du soutien, exprime une plainte ou vit une situation d'urgence. De plus, les tests comprenaient l'évaluation de la précision de la cote émotionnelle (sentiment) pour déterminer si les sentiments sont positifs, neutres, négatifs ou présentent un risque, afin d'assurer une compréhension précise du contexte émotionnel dans chaque interaction.

indice	Valeur
Temps d'analyse de la phrase (texte local)	300 ms (TFLite au téléphone)
Exactitude de la classification des intentions des utilisateurs	93%
Exactitude de l'analyse des sentiments	88%

Tableau 4.3 : Performances du modèle d'intelligence.

4.3.4 Performances de la montre(bracelet) intelligente (ESP32 + INMP441 + MPU6050) :

Un certain nombre d'indicateurs essentiels ont été surveillés afin d'évaluer l'efficacité du système, en suivant la fréquence de transmission à Firebase pour s'assurer que les données n'étaient pas dupliquées ou sur transmises. La précision des mesures de mouvement du capteur MPU6050 a également été évaluée afin de surveiller avec précision l'activité physique. La sensibilité du microphone INMP441 a été testée lors de l'enregistrement des fréquences et des intensités sonores, même dans des environnements peu bruyants. Enfin, la surveillance comprenait le taux de consommation d'énergie de la batterie pour s'assurer que le système fonctionnait pendant des périodes prolongées sans perte de puissance induite.

Indice	valeur
Mise à jour vers Firebase	Toutes les 1 secondes
Précision des capteurs de mouvement	Bon (3 %)
Réponse du micro	Efficace même à partir de 1,5 m
la consommation de la batterie	45 mA/h environ
connexion après une panne	Automatique en 5 secondes

Tableau 4.4 : Performances de la montre intelligente.

4.3.5 Performances des caractères 3D (3D + LipSync) :

La synchronisation entre le son et les mouvements des lèvres a été soigneusement testée à l'aide de fichiers ElevenLabs et LipSync pour s'assurer que les phonèmes correspondent aux mouvements visuels. Le temps de chargement du caractère 3D dans WebView a également été mesuré pour assurer l'apparition rapide et la continuité de l'interaction sans délai. De plus, des mouvements expressifs subtils tels que le clignement d'œil spontané, le sourire léger et l'inclinaison de la tête ont été évalués en raison de leur rôle dans l'amélioration du sens du réalisme et du confort psychologique de l'utilisateur.

Indice	Valeur
temps de téléchargement du personnage pour la première fois	1.8 seconde
Compatibilité audio des lèvres (LipSync)	Correspondance réaliste à 95
FPS dans WebView	28 à 30 images par seconde
Cils et expressions du visage	fonctionnement périodique et normale

Tableau 4.5 : Performances des caractères 3D.

4.4 Expérience utilisateur (commentaires des utilisateurs) :

L'évaluation de l'expérience utilisateur (UX) est un élément essentiel pour mesurer l'efficacité des systèmes intelligents, en particulier dans le domaine de la santé, où les aspects humains et psychologiques sont à la base de l'acceptation de la technologie par l'utilisateur. De ce point de vue, un essai pilote réaliste a été réalisé qui comprenait un groupe de patients et de spécialistes pour évaluer plusieurs axes de base, notamment : la facilité d'utilisation, la qualité de l'interaction audio et visuelle, et l'utilité du contenu thérapeutique basé sur CBT, en plus du sentiment de sécurité et de confiance de l'utilisateur dans l'utilisation du système, et le degré de satisfaction avec le caractère virtuel qui représente l'interface interactive du système.

4.4.1 Les Participants à l'expérience :

L'expérience du système a été menée avec un groupe de participants qui ont donné leur plein consentement volontaire à participer, en veillant au respect des principes éthiques des expériences psychologiques. L'expérience a eu lieu dans deux environnements différents : soit à l'intérieur de la maison pour assurer le confort et l'intimité de l'utilisateur, soit dans des cliniques professionnelles sous la supervision de médecins, afin d'assurer une diversité de contextes et de piloter le système dans des conditions réalistes qui reflètent l'utilisation réelle prévue.

catégorie	Nombre	remarques
psychopathes	6	Ils ont entre 20 et 40 ans
des psychologues	3	Ils ont analysé l'efficacité des réponses et de la thérapie comportementale
Ingénieurs et concepteurs	2	Évaluer les performances et l'interface

Tableau 4.6 : Participants à l'expérience .

4.4.2 Questions pour les utilisateurs :

Un questionnaire de 10 questions a été utilisé pour évaluer avec précision l'expérience de l'utilisateur sous divers aspects. Les questions portaient sur la facilité d'interaction avec l'assistant intelligent, le caractère humain et confortable des réponses, ainsi que sur la mesure dans laquelle la personnalité virtuelle affecte le soulagement du stress. Des questions ont également été posées au sujet de l'amélioration notable de l'état psychologique après la conversation, et du niveau de confiance dans ce type d'assistance psychologique numérique, en plus de surveiller les difficultés techniques rencontrées par l'utilisateur lors de son utilisation.

4.4.3 Résultats de l'enquête sur la satisfaction :

Les réponses au questionnaire ont été analysées et converties en pourcentages, et les résultats ont montré que plus de 85 % des participants exprimaient une satisfaction totale à l'égard du système. Ils ont également souligné que l'expérience leur a fourni un soutien psychologique réel et rapide, dans un environnement exempt de la pression ou de la stigmatisation sociale qui peut accompagner certaines méthodes traditionnelles de traitement. Ce qui reflète l'efficacité du système pour améliorer le confort et la confiance des utilisateurs.

Question	Taux de satisfaction
facilité d'utilisation	92%
Les réponses sont réalistes et humaines	88%
Caractère 3D impressionnant et réaliste	90%
Confort psychologique après utilisation	85%
la confiance dans le système	81%
Qualité sonore (TTS + Avatar)	93%
Assistant compréhension du dialecte/de la langue	89%
Accès facile aux rendez-vous et aux médecins	86%
Désir de réutiliser le système	84%
Pas de dysfonctionnements techniques ou problèmes	78%

Tableau 4.7 : Résultats de l'enquête sur la satisfaction .

4.4.4 Notes des médecins :

Lors de l'évaluation du système, certains cliniciens et utilisateurs ont exprimé leurs opinions qualitatives, l'un notant que "Emily donne des réponses très semblables à la TCC", ce qui reflète la qualité du contenu psychologique incorporé. D'autres ont également fait remarquer que "la personnalité interactive encourage les patients à être sans vergogne", ce qui est un aspect important de l'environnement thérapeutique. Malgré les éloges généraux, on a noté que « le système nécessite des mises à jour périodiques et une formation continue du modèle, mais qu'il a un excellent départ », ce qui reflète la compréhension de la nature des systèmes intelligents et la nécessité de leur développement continu.

4.4.5 Suggestions des utilisateurs :

Grâce aux commentaires des utilisateurs, plusieurs suggestions importantes sont apparues pour améliorer l'expérience à l'avenir, notamment : la possibilité de choisir la voix de l'assistant entre un mode calme ou actif en fonction des préférences de

l'utilisateur, ce qui améliore le sentiment de personnalisation et de confort. L'idée d'ajouter des rapports hebdomadaires qui résument les développements psychologiques de manière simplifiée a également été proposée, ce qui aiderait à suivre les progrès. Enfin, certains ont demandé une meilleure prononciation de certains mots du dialecte algérien afin de rendre l'interaction plus naturelle et plus proche de la réalité locale.

Les utilisateurs se sont dits très satisfaits de l'expérience, considérant que le caractère virtuel leur apportait un soutien psychologique tangible et les aidait à s'exprimer librement. Les médecins participants ont également confirmé que le système a des capacités prometteuses qui le qualifient pour être un véritable outil de soutien dans les séances de psychothérapie, à condition que son développement et sa modernisation continuent d'assurer son efficacité et sa sécurité psychologique.

4.5 Analyse des résultats :

Après avoir effectué des tests de performance et interagi avec les utilisateurs, nous analysons dans cette section les résultats académiquement et objectivement pour déterminer les points forts et les points d'amélioration du système, et relier les données expérimentales aux objectifs théoriques du projet.

4.5.1 Conformité du système aux objectifs généraux :

L'objectif du projet est de créer un système intelligent, interactif et disponible qui aide à soutenir les patients psychiatriques et à analyser leurs cas à distance.

Dimension	Vérification	explication
Soutien à la psychothérapie TCC	OUI	Les réponses vocales et personnelles sont pensées de manière thérapeutique
Réponse vocale instantanée	OUI	Utilisation de Wake Word + TTS + LipSync
Analyse en temps réel de l'état moteur	OUI	Au moyen de ESP32, MPU6050 et analyse d'activité via Firebase
Analyse en temps réel du ton de la voix	OUI	Utiliser INMP441 et le lier à la détection de sentiment
Soutien multilingue (Darija/français/anglais)	OUI	Via un modèle Mistral préformé
Permettre au médecin de faire un suivi sur la condition	OUI	Via le tableau de bord Firebase
Une expérience humaine avec une interface réaliste	OUI	Personnage 3D interactif réaliste et percutant

Tableau 4.8 : Conformité du système aux objectifs généraux .

4.5.2 Analyse de l'interaction des utilisateurs :

D'après les résultats de la section 4.4, l'expérience a montré que le personnage « Emily » s'engageait plus rapidement que « Jack », ce qui est attribué à son ton silencieux et à son expression émotionnelle qui augmentait le sentiment de sécurité des utilisateurs. Il a également été démontré que les patients qui utilisaient la montre intelligente pendant des périodes plus longues obtenaient des analyses psychologiques plus précises grâce à la disponibilité continue de données vitales. En ce qui concerne le style de réponse, les réponses brèves et directes étaient plus acceptables et efficaces dans l'interaction par rapport aux réponses longues ou ramifiées.

4.5.3 Analyse de la performance technique :

Les expériences ont montré que tous les composants du système fonctionnaient en temps quasi réel, ce qui permet une interaction continue et fluide entre l'utilisateur et l'assistant. Aucun plantage critique ou plantage n'a été observé pendant la période de test, ce qui indique la stabilité de l'architecture logicielle. Firebase s'est également avéré très efficace comme point de stockage et de synchronisation des données entre les différentes parties du système. Cependant, il y a eu petits problèmes de performance avec WebView, comme des retards dans le chargement du caractère interactif sur certains appareils vulnérables ou limités en mémoire ont été observés.

4.5.4 Intégration de systèmes multiples :

L'une des réalisations les plus remarquables de ce projet est l'intégration efficace de quatre composants principaux : un dispositif physique (une montre intelligente avec capteurs), une application intelligente interactive fonctionnant sur Android, un serveur de données en nuage fiable (Firebase), et un système d'analyse interne basé sur l'intelligence artificielle (en utilisant un modèle Mistral converti en TFLite). Cette intégration complète entre le matériel, les logiciels et l'analyse en nuage confère au système un caractère professionnel proche des systèmes médicaux approuvés dans les pays développés, ce qui améliore sa fiabilité et sa valeur dans l'environnement de la psychothérapie intelligente.

4.5.5 Comment Surmonter les difficultés :

Un équilibre délicat entre vitesse de réponse et analyse de phrases instantanée a été atteint grâce à l'utilisation du modèle TFLite localement sur la machine, ce qui a permis de réduire le temps de traitement sans sacrifier la précision. Le darija algérien a également été soutenu par une formation dédiée qui reflète les particularités du dialecte local et assure une compréhension naturelle des besoins de l'utilisateur. D'autre part, le son, l'image et les expressions faciales interactives ont été harmonieusement combinés en utilisant WebGL SDK pour la présentation des personnages, ElevenLabs pour la génération du son et WebView pour la présentation

dans l'application, ce qui a grandement amélioré la qualité de l'interaction humaine et le sens du réalisme.

4.6 Évaluation générale et comparaison avec d'autres solutions :

Après avoir effectué des tests et analysé les résultats, nous pouvons maintenant évaluer de manière exhaustive le système en termes d'innovation, d'efficacité et de faisabilité pratique, et également le comparer à certaines solutions similaires actuellement sur le marché ou la recherche scientifique.

4.6.1 Évaluation globale du système :

Le système combine les dernières technologies d'intelligence artificielle et d'Internet des objets appliqué au domaine de la santé. Il fournit un assistant intelligent qui réagit vocalement en darija et en français, soutenu par un personnage réaliste 3D qui parle et s'exprime en mouvement, apportant une dimension humaine à l'expérience. À cela s'ajoute sa capacité de réaliser une psychanalyse instantanée basée sur le son et le mouvement, avec un soutien direct pour les méthodes de thérapie comportementale cognitive (TCC). Le système comprend également une interface dédiée aux médecins pour un suivi précis des cas, en plus d'un système de capteurs portable (smart watch) pour collecter les données vitales et une intégration complète avec une base de données cloud mise à jour en temps réel. Toutes ces composantes font du projet une innovation intégrée qui combine l'informatique, la santé mentale et l'IOT pour offrir une solution pratique et efficace adaptée au contexte local.

4.6.2 Comparaison avec des solutions globales similaires :

Le tableau 4.9 comparatif montre que notre système offre une intégration complète entre plusieurs fonctions de base sur une même plateforme, notamment : interaction vocale, psychanalyse momentanée, soutien de la thérapie cognitivo-comportementale (TCC), personnage parlant en trois dimensions, interface dédiée Pour le médecin, et un système de capteurs intelligent via la montre. La caractéristique la plus importante réside dans le soutien à la langue algérienne ,qui n'est fourni par aucune des autres solutions, en plus d'adopter un modèle local d'intelligence artificielle spécialement formé pour ce contexte. Cela rend le système non seulement techniquement avancé, mais socialement adapté aux besoins de l'utilisateur algérien.

Norme	Notre système	Wysa App	Replika	Fitbit + Calm
Prise en charge la TCC	✓	✓	✗	✗
Fonctionnement sans internet (en partie)	✓	✗	✗	✓
Caractère visuel en 3D.	✓	✗	en partie	✗
Soutient le dialecte algérien.	✓	✗	✗	✗
Analyse physiologique (voix + mouvement).	✓	✗	✗	✓ en partie
Possibilité de communiquer avec un vrai médecin.	✓	✗	✗	✗
Open source + entièrement local.	✓	✗	✗	✗

Tableau 4.9 : Comparaison avec des solutions globales similaires .

4.6.3 Valeur ajoutée réelle :

Ce système est considéré comme une solution efficace au problème de la pénurie de psychiatres en Algérie, car il offre une expérience de soutien psychologique réaliste et professionnel 24 heures sur 24, dans le plein respect de la vie privée de l'utilisateur grâce à son travail local sans dépendance importante sur des serveurs externes. Sa structure actuelle permet également de la développer ultérieurement en dispositif médical approuvé s'il est soumis à l'approbation des autorités sanitaires compétentes.

4.7 Remarques, limites et défis :

Bien que le système proposé ait fait preuve de rendement élevé et d'efficacité démontrée, tous les projets de recherche ou techniques ne sont pas exempts de certains défis et limitations qui peuvent avoir une incidence sur l'échelle ou la performance dans d'autres environnements. Dans cette section, nous présentons les observations techniques et humaines les plus importantes que nous avons rencontrées au cours du développement et de l'expérimentation.

4.7.1 Limitations techniques :

Bien que le système soit efficace et novateur, il y a des défis techniques qui se sont posés pendant le développement et les essais. La plus notable est l'adoption de certains composants sur Internet, car certaines fonctions comme la synchronisation des données avec Firebase ou la conversion du texte en audio via ElevenLabs

nécessitent une connexion réseau, bien que le système prenne en charge un travail partiel sans Internet à l'aide du modèle TFLite. En outre, l'utilisation de WebView pour exécuter le caractère 3D entraîne une consommation de mémoire élevée sur certains appareils faibles, ce qui affecte les performances.

En outre, un léger retard dans le téléchargement de l'audio résultant du TTS d'une durée allant jusqu'à 1,5 seconde a été observé dans certains cas, ce qui peut être remarqué par l'utilisateur et affecter le flux. La taille de l'échelle Mistral, bien qu'améliorée, consomme toujours des ressources de traitement relativement élevées, ce qui limite l'efficacité du traitement du langage naturel (NLP) dans les appareils aux capacités limitées.

4.7.2 Limites humaines et comportementales :

Sur le plan humain et psychologique, certains défis sont apparus dans l'acceptation générale de l'assistant intelligent. Un certain nombre d'utilisateurs ont exprimé leur réticence à interagir avec un personnage virtuel plutôt que de parler avec un vrai médecin humain, en particulier sur des questions psychologiques délicates qui nécessitent une empathie humaine directe. Certains médecins ont également noté une réserve à l'égard du recours total au diagnostic automatisé, préférant le diagnostic humain traditionnel, ce qui indique la nécessité de plus de sensibilisation sur la précision et la fiabilité des systèmes d'intelligence artificielle dans le domaine de la santé.

En ce qui concerne la confidentialité des données, bien que le système repose sur un cryptage complet de la communication et des données, certains patients ont exprimé des préoccupations au sujet du stockage de leurs informations psychologiques sur des plateformes en nuage, qui souligne l'importance de fournir des garanties ou des solutions alternatives telles que le stockage local. Ou permettre à l'utilisateur de gérer davantage ses autorisations de données.

4.7.3 Défis lors du développement :

Le projet a fait face à d'importants défis techniques et linguistiques au cours des étapes de développement, notamment la formation du modèle Mistral sur le dialecte algérien, car il n'y avait pas de base de données prête à l'emploi, ce qui a nécessité la collecte manuelle de phrases et la reformulation. Les traduire et les mettre en forme pour couvrir des contextes psychologiques. Et les différences sociales, qui ont nécessité beaucoup d'efforts et de temps pour assurer le réalisme des données.

Quant au niveau de réponse instantanée de l'assistant intelligent, il a fallu ajuster le Wake Word précisément, tout en ajustant le niveau de sensibilité pour réduire les fausses activations d'une part, et éviter les retards dans l'interaction d'autre part, afin de maintenir une expérience utilisateur naturelle.

En ce qui concerne l'intégration du matériel, Il y a eu des difficultés de formatage des données sensorielles provenant de l'ESP32, notamment lors de la intégration des lectures du microphone (audio) avec le capteur de mouvement (MPU6050), en plus de assurer le moment de la transmission simultanée à Firebase d'une manière qui maintenir la séquence des événements et l'exactitude de l'analyse en temps réel dans l'application.

4.7.4 Limites de la portée du projet :

Malgré les résultats positifs obtenus par le système durant la phase de test, il y a quelques limitations qui doivent être prises en compte. Plus particulièrement, le nombre de participants était limité (seulement 11 personnes), ce qui ne permet pas d'étendre les résultats à tous les segments de la société. De plus, le système n'a pas encore été testé sur différents groupes d'âge comme les personnes âgées, ou sur des patients psychiatriques chroniques qui nécessitent un suivi à long terme.

De plus, l'expérience 3D des personnages n'a été réalisée que sur Android et n'a pas encore été portée ou adaptée pour fonctionner sur d'autres plateformes comme iOS ou le Web, ce qui limite actuellement la portée du système.

4.8 Conclusion :

Ce chapitre représente la conclusion de l'évaluation réelle du système intelligent proposé, où une série de tests en conditions réelles ont été effectués pour mesurer l'efficacité de chaque composant technique et fonctionnel, à commencer par les appareils portables, en passant par l'application intelligente, et se terminant par le personnage virtuel soutenu par l'intelligence artificielle. Les résultats ont montré que le modèle Mistral, en intégration avec le microphone INMP441, était capable d'analyser le contenu tonal et textuel avec une grande précision, tandis que le capteur physiologique MPU6050 fournissait des lectures précises de la condition physique et de l'activité motrice.

La réponse fluide et naturelle de l'assistant intelligent a également été mise en évidence, soutenue par le caractère 3D réaliste, qui a amélioré le sens de l'interaction humaine de l'utilisateur. Les utilisateurs et les médecins ont exprimé leur satisfaction à l'égard de cette expérience, soulignant que le système a déjà les capacités nécessaires pour soutenir la santé mentale, en particulier dans des environnements qui souffrent d'une pénurie de spécialistes ou de difficultés d'accès au soutien thérapeutique.

Malgré les défis techniques et certains obstacles humains, le modèle a prouvé sa capacité à effectuer des tâches efficacement, en faisant un prototype prometteur qui pourra être développé à l'avenir pour constituer une des solutions avancées dans le domaine de la psychothérapie basée sur l'intelligence artificielle et l'Internet des objets.

Conclusion générale

A la conclusion de ce travail, on peut dire que notre projet est venu comme une réponse réaliste et tangible à un besoin sociétal croissant dans le domaine de la santé mentale, notamment dans des environnements qui souffrent du manque de spécialistes, d'une pénurie de services psychologiques, et des barrières culturelles et sociales élevées qui empêchent beaucoup de personnes de demander du soutien. De ce point de vue, nous avons pu concevoir un système intelligent intégré qui combine des appareils portables qui surveillent les indicateurs de l'état physiologique, et une application intelligente qui fonctionne comme un assistant virtuel qui soutient la thérapie comportementale cognitive (TCC), supporté par un modèle linguistique local (Mistral) capable de comprendre et d'analyser les textes et le ton sensoriel dans le dialecte algérien en plus d'une interface 3D qui adopte un caractère interactif (3D Avatar + LipSync), il améliore le sens de l'interaction humaine et réduit l'éloignement numérique dont souffrent la plupart des systèmes psychologiques numériques. En outre, la plateforme Firebase a permis de stocker et de suivre les données et a permis aux médecins de suivre les patients à distance de manière sûre et efficace.

Les tests qui ont été réalisés ont montré que le système fonctionne efficacement dans des conditions réalistes, car il a montré une grande capacité à analyser le langage et les émotions, une réponse naturelle et lisse lors de l'interaction avec le son, avec une grande précision dans l'analyse du mouvement et de l'activité physiologique à l'aide des capteurs MPU6050 et INMP441. Les évaluations que nous avons obtenues des participants à l'essai et des médecins qui ont supervisé le suivi confirment que ce système constitue un outil prometteur pouvant être utilisé en complément du soutien psychologique traditionnel, et peut même combler un réel vide dans des environnements souffrant de marginalisation ou d'éloignement par rapport aux centres spécialisés. L'expérience de l'utilisateur a également montré une grande satisfaction en termes de facilité d'utilisation, de qualité d'interaction et du sentiment de l'utilisateur que l'assistant comprend réellement et répond d'une manière qui est proche d'un humain plutôt que d'une machine rigide.

Malgré ce succès, certains des défis auxquels nous avons été confrontés pendant le développement du système ne peuvent être ignorés, y compris le nombre limité de participants à l'essai (seulement 11 personnes), et l'incapacité de tester le système sur des groupes d'âge différents ou des patients souffrant de troubles chroniques.

De plus, l'expérience des personnages 3D était actuellement limitée à Android et n'a pas encore été portée sur les versions web ou iOS. Cependant, tous ces obstacles sont normaux dans le contexte d'un projet initial, ils peuvent faire l'objets de travaux futurs

Sur la base de ce qui précède, nous proposons un ensemble de recommandations de développement pour assurer l'amélioration du système et étendre son utilisation, à savoir : la nécessité d'élargir la base de données linguistiques pour inclure plus de dialectes locaux et de terminologie, qui améliore la précision de l'analyse linguistique et l'interaction réaliste dans le dialecte algérien; Développer des versions parallèles pour le web et iOS afin d'élargir l'accès; Intégrer un système de recommandation intelligent qui avertit automatiquement le médecin lorsque des signaux de langage dangereux ou sensibles sont détectés; Activation de la fonction d'évaluation automatique hebdomadaire qui envoie des rapports périodiques au patient et au médecin ; utiliser plus de capteurs médicaux certifiés tels que le MAX30102 pour mesurer le pouls et la pression artérielle ,des capteurs visuels pour l'analyse faciale ; Développer un mécanisme d'auto-apprentissage qui permet au modèle d'améliorer ses performances en fonction de l'interaction de l'utilisateur; Fournir plusieurs voix et personnalités dans l'assistant intelligent selon le sexe ou l'état psychologique (voix masculine / féminine, calme ou tonalité stimulante).

Nous recommandons également d'explorer la possibilité d'intégrer le système dans les programmes nationaux, en particulier dans le secteur de la santé scolaire et universitaire, ou même dans des initiatives du ministère de la santé numérique. Un tel système pourrait être utilisé pour la détection précoce, la sensibilisation et l'accompagnement psychologique initial, en particulier dans les régions éloignées.

Enfin , nous espérons que ce modeste travail constitue une contribution significative dans le domaine des soins psychologiques intelligents.

1. PRESENTATION DE L'ESP32

L'**ESP32** est une puce combinée Wi-Fi et Bluetooth fonctionnant à 2,4 GHz, conçue avec la technologie TSMC 40 nm ultra-basse consommation. Elle offre des performances radio et énergétiques optimales, avec une robustesse et une fiabilité adaptées à de nombreuses applications.

Les modèles de la série incluent : ESP32-D0WDQ6, ESP32-D0WD, ESP32-D2WD et ESP32-S0WD.

1.1 SOLUTIONS INTEGREES

1.1.1 SOLUTION ULTRA-BASSE CONSOMMATION

Conçu pour les appareils mobiles, portables et les applications IoT, l'ESP32 intègre des modes d'économie d'énergie comme la gestion fine de l'horloge, plusieurs modes de veille et l'ajustement dynamique de la puissance. Il peut se réveiller uniquement lorsque c'est nécessaire, minimisant ainsi la consommation. L'amplificateur de puissance a une sortie ajustable, permettant d'optimiser la portée, le débit de données et la consommation énergétique.

1.1.2 SOLUTION TOUT-EN-UN

L'ESP32 intègre un commutateur d'antenne, un balun RF, un amplificateur de puissance, un amplificateur faible bruit, des filtres et des modules de gestion de l'énergie, avec environ 20 composants externes. Il utilise la technologie CMOS avec calibrations intégrées pour simplifier la production sans équipement Wi-Fi coûteux.

1.2 FONCTIONNALITÉS WI-FI

- 802.11 b/g/n
- 802.11n (2,4 GHz), jusqu'à 150 Mbps
- A-MPDU (TX/RX), A-MSDU (RX)
- Blocage ACK immédiat
- Surveillance des balises
- 4 interfaces Wi-Fi virtuelles

1.3 FONCTIONNALITÉS BLUETOOTH

- Compatible Bluetooth v4.2 BR/EDR et BLE
- Transmetteur classe 1, 2, 3 sans amplificateur externe
- Puissance d'émission : +12 dBm
- Récepteur NZIF : sensibilité BLE à -97 dBm
- Saut de fréquence adaptatif (AFH)
- Interface HCI via SDIO/SPI/UART
- Jusqu'à 4 Mbps via UART
- Contrôleur double mode BR/EDR et BLE
- Connexions multiples BT classiques et BLE

1.4 FONCTIONNALITÉS MCU

- Processeur Xtensa® LX6 simple/double cœur, jusqu'à 600 MIPS
- 448 Ko de ROM, 520 Ko de SRAM
- 16 Ko de SRAM dans le RTC
- QSPI pour mémoire Flash/SRAM multiple

2. INMP441 – MICROPHONE MEMS NUMERIQUE

2.1 DESCRIPTION GENERALE

L'**INMP441** est un microphone MEMS numérique omnidirectionnel, basse consommation, avec port inférieur. Il inclut un capteur MEMS, un conditionneur de signal, un convertisseur analogique-numérique 24 bits, des filtres anti-repliement, et une interface I²S.

Il présente une excellente réponse en fréquence large bande, offrant un son naturel et une bonne intelligibilité.

2.2 APPLICATIONS

- Systèmes de téléconférence
- Télécommandes
- Consoles de jeu
- Appareils mobiles
- Ordinateurs portables/tablettes
- Systèmes de sécurité

2.3 CARACTÉRISTIQUES

- Interface I²S numérique 24 bits
- SNR élevé : 61 dBA
- Sensibilité : -26 dBFS
- Réponse en fréquence plate de 60 Hz à 15 kHz
- Faible consommation : 1,4 mA
- Rejet d'alimentation : -75 dBFS
- Boîtier compact 4,72 × 3,76 × 1 mm
- Conforme RoHS/WEEE

FONCTIONNEMENT

- Mode normal : actif après 218 cycles SCK (≈ 85 ms)
- Mode veille : actif après arrêt de SCK + 214 cycles à la reprise
- Mode arrêt : activé par CHIPEN bas
- Interface I²S : 24 bits, complément à deux, MSB first
- La broche SD est tri-état lorsqu'inactive
- 64 cycles SCK par trame stéréo (32 par mot de données)

3. MPU-6000 / MPU-6050 – GYROSCOPE ET ACCELEROMETRE

3.1 DESCRIPTION GENERALE

Le **MPU-60X0** combine un gyroscope 3 axes et un accéléromètre 3 axes dans une seule puce avec un processeur numérique de mouvement (DMP™) intégré. Il permet l'exécution d'algorithmes de fusion de capteurs 9 axes avec très haute précision.

3.2 AVANTAGES

- Précision accrue
- Élimine les erreurs d'alignement
- Plage gyroscope : ± 250 , ± 500 , ± 1000 , ± 2000 °/s
- Plage accéléromètre : $\pm 2g$, $\pm 4g$, $\pm 8g$, $\pm 16g$
- Compatible I2C (6050) ou SPI (6000)

3.3 AUTO-TEST GYROSCOPE ET ACCELEROMETRE

Le test matériel active un signal interne simulant une force (Coriolis pour gyroscope, force linéaire pour accéléromètre), permettant de comparer la sortie mesurée à une valeur de référence en usine. Si la différence dépasse un seuil, le test échoue.

4– ÉCRAN OLED. SSD1306

4.1 DESCRIPTION GENERALE

Le **SSD1306** est un pilote pour écran OLED/PLED graphique matriciel 128x64. Il intègre une RAM d’affichage, un contrôle de contraste, un oscillateur interne et une interface compatible I2C/SPI/parallèle.

4.2 CARACTÉRISTIQUES PRINCIPALES

- Résolution : 128 x 64
- Alimentation logique : 1,65V à 3,3V
- Tension de pilotage OLED : jusqu’à 15V
- Contrôle de luminosité sur 256 niveaux
- Mémoire embarquée : 128 x 64 bits
- Interfaces : I2C, SPI 3/4 fils, parallèle
- Fonction de défilement vertical/horizontal
- Température de fonctionnement : -40°C à 85°C

Bibliographie

- [1] World Health Organization.. *Mental Health*,2022.
- [2] American Psychiatric Association.. *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*,2013.
- [3] Beck, J. S. *Cognitive Behavior Therapy: Basics and Beyond (2nd ed.)* New York: Guilford Press, 2011.
- [4] Bensaid, M.. *Mental health care in Algeria: Challenges and perspectives*. International Journal of Psychiatry, Vol. 5, No. 2, 2019.
- [5] Topol, E. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books,2019.
- [6] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M.. *Internet of Things (IoT): A vision, architectural elements, and future directions*. Future Generation Computer Systems, 29(7), 1645–1660,2013.
- [7] Deng, L., & Li, X. *Machine learning paradigms for speech recognition: An overview*. IEEE Transactions on Audio, Speech, and Language Processing, 21(5), 1060–1089,2013.
- [8] Taylor, P. *Text-to-Speech Synthesis*. Cambridge University Press,2009.
- [9] ElevenLabs. *Text To Speech API Documentation*,2024.
<https://www.elevenlabs.io>
- [10] TensorFlow Team *TensorFlow Lite Documentation*,2021.
<https://www.tensorflow.org/lite>
- [11] Firebase Google. *Firebase Documentation*,2023.
<https://firebase.google.com/docs>
- [12] Google Android Developers. *Android Studio User Guide* ,2023.
<https://developer.android.com/studio>
- [13] Eckel, B. *Thinking in Java (4th ed.)*. Prentice Hall,2006.

- [14] Android Developers. *Layouts and UI components in XML*,2023.
<https://developer.android.com/guide/topics/ui>
- [15] Cambria, E., & White, B. *Jumping NLP Curves*. IEEE Computational Intelligence Magazine,2014.
- [16] Mistral AI. *Mistral 7B Model Overview*,2023.
<https://mistral.ai>
- [17] Firebase Docs. *Realtime Database and Analytics*,2023.
- [18] HuggingFace. *Fine-Tuning LLMs*. <https://huggingface.co/docs>,2024.
- [19] Navarro, G. *A guided tour to approximate string matching*. ACM Computing Surveys (CSUR), 33(1), 31–88, 2001.
- [20] Picovoice.. *Porcupine Wake Word Engine*,2023.
<https://picovoice.ai/platform/porcupine/>
- [21] Google. *Google Colaboratory*, 2023.
<https://colab.research.google.com>,
- [22] Lutz, M. *Learning Python* (5th ed.). O'Reilly Media ,2013.
- [23] Metahuman by Epic Games. *Real-time 3D Human Avatars* ,2023.
<https://www.unrealengine.com/metahuman>
- [24] Sarikaya, R., et al. *LipSync for AI Avatars*. ACM SIGGRAPH,2021.