

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOHAMED BOUDIAF - M'SILA

**FACULTE : MATHÉMATIQUES ET
D'INFORMATIQUE**
DEPARTEMENT : D'INFORMATIQUE
N° :



DOMAINE : INFORMATIQUE
FILIERE INFORMATIQUE
**OPTION : : RESEAUX ET TECHNOLOGIES
DE L'INFORMATION ET DE LA
COMMUNICATION**

**Mémoire présenté pour l'obtention
Du diplôme de Master Académique**

Par: ZEYNEB BERBIT et AMINA HAMDI

Intitulé

**Conception et Réalisation d'une Application
Web pour la Reconnaissance Intelligente des
Sommes Monétaires**

Soutenu devant le jury composé de :

Benazi Makhlouf	Université de M'sila	Président
Tahar Mehenni	Université de M'sila	Rapporteur
Hamani Mohamed Said	Université de M'sila	Examineur

Année universitaire : 2023/2024



Dédicace

«بسم خالقي و ميسر اموري و عصمت امري،

كل كل الحمد و الامتنان و الرضا "

من كلكه الله بالهيبه و الوقار...الى من أحمل اسمه بكل فخر...وزين أسمي بأجمل الألقاب ،
الى من حصد الأشواك عن دربي ليمهد لي طريق العلم بعد فضل الله ما أنا فيه يعود الى أبي ،
الرجل الذي سعى طوال حياته لكي نكون أفضل منه

أبي الغالي

الى من جعل الله الجنة تحت أقدامها، إلى ملاكي في الحياة التي احتضنتني بقلبها قبل يديها
وسهلت لي الشدائد بدعائها، التي كانت لي الأم والأخت ووجهتي الدائمة

أمي الحبيبة

الى مصدر قوتي، وضلعي الثابت وأمان أيامي الى ملهمي نجاحي من شددت عضدي بهم
فكانوا لي يناييع أرتوي منها

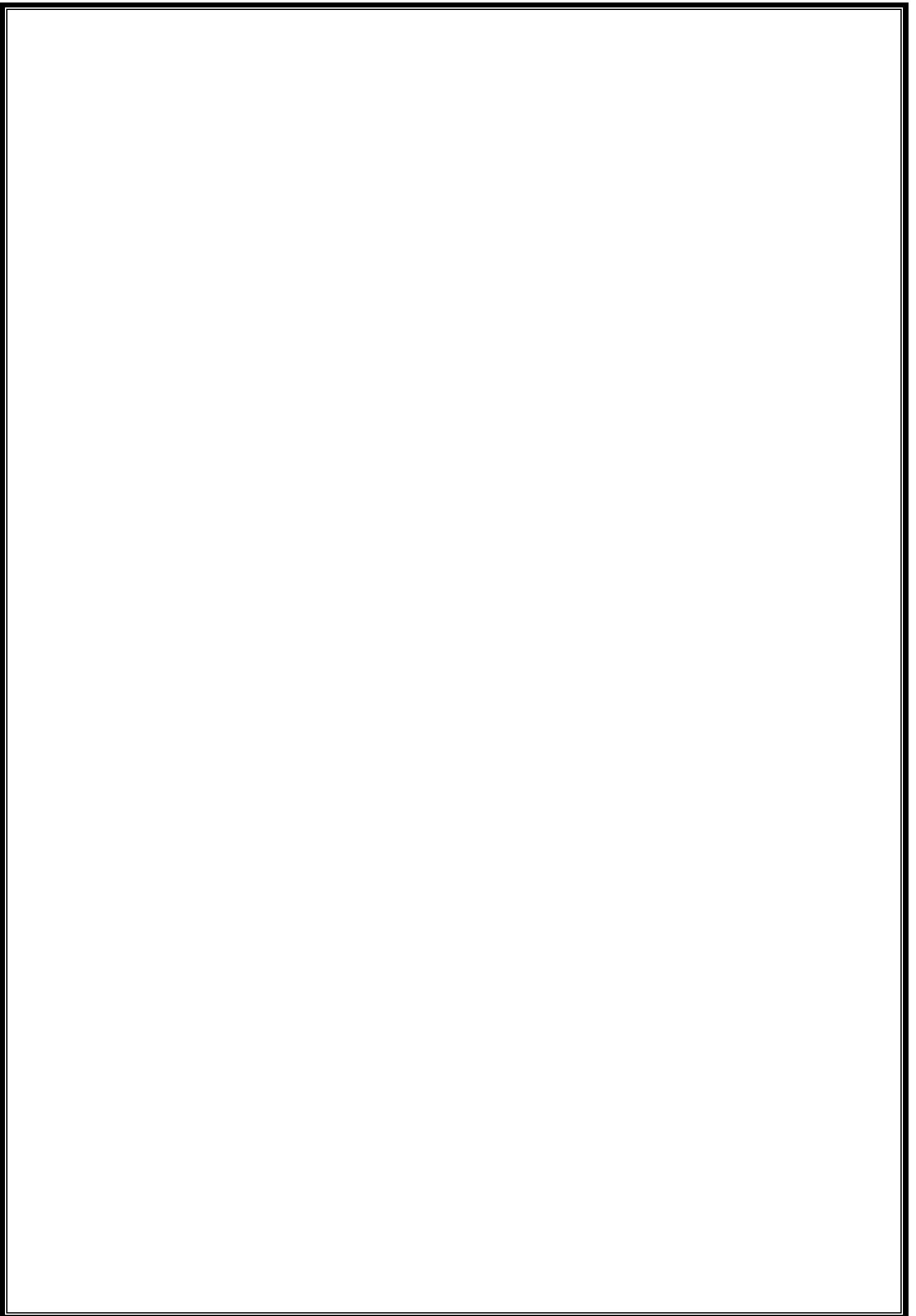
أخواي وأختاي

الى رفيقة الروح التي شاركتني خطوات هذا الطريق، خيرة أيامي و صفوتها .

رفقاء الدرب كنتم عوننا وسندا في هذا الطريق كل باسمها.

الى مصدر قوتي، أرضي الصلبة و جداري المتين، الداعمين في الخفاء.

ها أنا اليوم أتممت أول ثمرة نجاحي الذي هو نجاح والداي الذي لطالما تمنياه، لله الشكر
كله أن وفقني لهذه اللحظة، فالحمد لله رب العالمين.



Dédicace

الحمد لله حبا وشكرا وامتنانا على البدء والختام

لم تكن الرحلة قصيرة ولا ينبغي لها ان تكون لم يكن الحلم قريبا لا الطريق كان محفوظا
بالتسهيلات لكني فعلتها فالحمد لله الذي يسر البدايات وبلغنا النهايات بفضلة وكرمة
اهدي تخرجي إلى من احمل اسمه بكل فخر الذي كان دوما بقلبي رغم مرضه إلا أنه دائما
بجانبي وسندي بلا حدود إلى من علمني أن الدنيا صراع وسلاحها العلم

والمعرفة

والذي حفظه الله

إلى المرأة التي جعلت مني فتاة طموحة وسهلت علي الصعاب بدعائها الخفي إلى إلى القلب
الحنون والشمعة التي كانت لي في الليالي المظلمة

أمي حفظها الله

الى خيرة أيامي وصفوتها كانوا لي سندا وداعمين ومشجعين دائما أزاحوا عن طريقي
المتاعب ممهدين الطريق زارعين الثقة والاصرار بداخلي سندي وكتفي الذي استند عليه
دائما الى الذين غمروني بالحب والتوجيه امدوني بالقوة إلى الذين لطالما كانوا الظل لهذا

النجاح

اخواني واختي حفظهم الله

واخيراً من قال أنا لها "نالها" وأنا لها ان أبت رغما عنها اتيت بها ماكنت لأفعل لولا توفيق
من الله ها هو اليوم العظيم هنا اليوم الذي أجريت وسنوات الدراسة الشاقة حالمة فيها حتى
توالت بمنه وكرمه لفرحة التمام الحمد لله الذي به خيرا واملأ إلا وأغرقنا سرورا وفرحا
ينسيني مشقتي.

أمينة

TABLE DES MATIERES

Introduction generale.....	1
Chapitre 1 : Informations générales sur la poste et le web.....	3
1. Introduction:.....	3
2 .Présentation de la société « algérie poste »	3
2.1 Qu'est-ce que « algérie poste » ?.....	3
2.2 Services d'algérie poste.....	4
2.3 Vision de la poste algérienne.....	5
2.4 Transformation numérique en algérie poste	5
3. Applications web.....	7
4. Technologies web.....	8
4.1 Technologies côté client *front-end*.....	8
4.2 Technologies côté serveur*back-end*.....	9
5. Architecture client/serveur.....	9
6. Conclusion:.....	11
Chapitre 2. La reconnaissance de la parole.....	12
1. Introduction.....	12
2. Principes de base.....	13
3. Historique.....	14
4. Matériel technique	14
5. Les modèles.....	15
6. Classification.....	15
7. Logiciels de reconnaissance de la parole.....	16
8. Comment fonctionne la reconnaissance de la parole ?	16
9. Conclusion.....	16
Chapitre 3. Analyse et Conception.....	17
1. Introduction.....	17
2. Aperçu général sur UML.....	17

2.1 UML.....	17
2.2 Développement du Langage de Modélisation Unifié (UML).....	18
2.3 Diagrammes d'UML.....	18
3. Présentation générale de notre application	19
4. Diagrammes de cas d'utilisation.....	20
5. Diagrammes de séquences.....	20
6. Diagrammes d'activité.....	22
7. Conclusion.....	23
Chapitre 4. Implémentation.....	24
1. Introduction :.....	24
2. Outils utilisés pour le back-end.....	24
3. Outils utilisé pour le front-end.....	25
4. Visual Studio code pour le web.....	26
5. Bibliothèques utilisées:.....	26
1- numpy.....	26
2- levenshtein.....	26
3- re.....	27
4- numbers_to_words.....	27
5- whisper.....	27
6. fonctions utilisées.....	28
7. démarrage de l'application.....	29
8. page d'insertion.....	30
8.1 entrer les chiffre par écrits :.....	30
8.2 entrer les chiffre par voix(message vocal)	33
9.Conclusion.....	35
Conclusion générale.....	36

Liste de figures

Figure 1.1: Logo de la société Algérie Poste	3
Figure 1.2 : Direction d'Algérie Poste.....	4
Figure 1.3 : Représentation d'un site Web statique	6
Figure 1. 4 : Représentation d'un site Web dynamique.....	7
Figure 1. 5 : L'architecture client/serveur.....	10
Figure 1. 6 : Architecture client/serveur à 1 niveau	11
Figure 2.1 Linux kernel and gaming input-output latency	12
Figure 3.1 Le diagramme de cas d'utilisation.....	20
Figure 3.2 : Le diagramme de séquences conversion numerique	21
Figure 3.3 : Le diagramme de séquences conversion audio	21
Figure 3.3 : Le diagramme d'activités	22
Figure 4.1: Le request_response_cycle_1.	25
Figure 4.2: Le request_response_cycle_2.	25
Figure 4.3: Le logo de python.	25
Figure 4.4 : Création d'un compte Azure Open AI Studio.....	27
Figure 4.5 : Inscription au système.....	28
Figure 4.6:leDémarrage de back end 1.	29
Figure 4.7:le Démarrage de back end 2.	29
Figure 4.8: le Démarrage de front end 1.	30
Figure 4.9: Entrer les chiffre par écrits	30
Figure 4.10:les messages d'erreur 1	31
Figure 4.11:les messages d'erreur 2.	31
Figure 4.12:les messages d'erreur 3.	32
Figure 4.13: Choisir la devise.....	32
Figure 4.14:Le résultat par des lettres	33
Figure 4.15: Entrer les chiffre par voix 1	34
Figure 4.16: Entrer les chiffre par voix 2	34

INTRODUCTION GENERALE

Face à l'explosion technologique et au développement rapide, il était nécessaire pour toutes les institutions de tous types, quelle que soit leur activité, d'acquérir diverses techniques et technologies modernes, telles que les moyens de communication parmi lesquels Internet, les technologies de communication par satellite, la téléphonie mobile, et les ordinateurs à grande vitesse, en plus du développement de l'utilisation de médias automatisés et d'autres technologies modernes.

Ces technologies avancées ont été rassemblées, et cette connexion a contribué à l'émergence de l'intelligence artificielle et au grand développement qui l'a accompagnée dans le domaine de l'informatique et de l'efficacité technologique. L'Algérie, comme d'autres pays, s'efforce toujours d'être attentive à ce domaine, alors que le secteur postal est entré dans la dynamique de modernisation du secteur des services et de diversification de ses services, comme les services sont considérés.

Le service postal est aujourd'hui le plus utilisé en Algérie, et cela est dû à son utilisation par diverses catégories de personnes, les travailleurs, les retraités et toutes les couches de la société. Cependant, aux portes des bureaux de poste, les passants rencontrent souvent de longues files de personnes âgées faisant la queue pour retirer leur argent. Ce qui est intéressant dans la scène, c'est que ces gens attendent le jour où leurs salaires arrivent pour les retirer, mais ce qui est douloureux, c'est quand l'un d'eux demande à quelqu'un de remplir le chèque postal à sa place, et qu'on le voit parler à voix basse, et suppliant quelqu'un d'écrire pour lui dans cette case vide le montant qu'il retirera en lettres.

L'objectif de notre projet est donc de créer une application web pour la reconnaissance intelligente des sommes monétaires, interactif, fiable et simple d'utilisation. Cette application a pour objectif principal de faciliter la rédaction correcte des montants dans une monnaie déterminée.

Ce mémoire inclut les chapitres suivants:

Le 1er chapitre: "Informations générales sur la poste et le web", présente un aperçu de la poste et des applications et sites web, de leurs types et des outils utilisés pour les développer.

Le 2ème chapitre : "la reconnaissance vocale", donne en détail le principe de reconnaissance vocale, ses architectures et technologies permettant de la développer.

Le 3ème chapitre : "Analyse et conception de l'application proposée", qui présente les différents diagrammes de conception de notre application, ainsi que la représentation des différentes phases de son élaboration.

Le 4ème chapitre: "Implémentation de l'application", qui présente les outils de programmation utilisés, ainsi que les fonctions principales de l'application.

CHAPITRE 1

INFORMATIONS GENERALES SUR LA POSTE ET LE WEB

1. Introduction:

L'utilisation des technologies de l'information représente une étape importante et décisive vers la transition aux services électroniques et le passage des formes de communication directes entre les citoyens et l'administration à la communication virtuelle, car elle repose sur l'utilisation de moyens et d'appareils et d'informations modernes et avancées, ainsi que des programmes qui apportent des solutions aux problèmes personnelles et bureaucratiques rencontrés. Malgré les efforts déployés au niveau de la Société Algérienne des Postes (Algérie Poste) pour offrir des services qui s'appuient sur les technologies de l'information et de la communication, nous les trouvons insuffisants en raison des obstacles qui entravent son efficacité.

2. Présentation de la Société « Algérie Poste »

2.1 Qu'est-ce que « Algérie Poste » ?

Algérie Poste (Fig. 1.1) est l'institution officielle des services postaux en Algérie, et son activité industrielle et économique s'articule autour des métiers postaux. C'est un établissement public à caractère industriel et commerciale [1]:

- Placé sous la tutelle du Ministère de la Poste, des Télécommunications, des Technologies et du Numérique (MPTTN),
- Administré par un Conseil d'Administration est présidé par le ministre chargé des postes ou son représentant;
- Dirigé par un Directeur Général, nommé par décret présidentiel.



Figure 1.1: Logo de la société Algérie Poste

Définie par un arrêté ministériel du 21 janvier 2010 et encadrée par deux (02) Comités, exécutif et Commercial & marketing, l'organisation de l'E.P.I.C Algérie Poste, est bâtie autour des métiers postaux classiques sur trois niveaux (Central, Régional, et Local)

La Direction Générale d'Algérie Poste et ses services centraux sont installés à Bab ezzouar 16024, au Lot n°01, Parcelle n°04, Zone d'affaires, dans un immeuble de douze (12) étages.



Figure 1.2 : Direction d'Algérie Poste

2.2 Services d'Algérie Poste

Les points suivants expliquent certains des services fournis par Algérie Poste :

- **Services postaux :** Ce sont les services d'envoi du courrier de base et supplémentaires qui nécessitent un traitement particulier, la location de boîtes postales, le stockage du courrier pendant 3 mois, le transfert du courrier vers un autre bureau de poste, le renvoi du courrier à une autre adresse, l'envoi de télégrammes par de nouvelles méthodes techniques et le service de courrier express. [1].
- **Services financiers:** Il est possible de retirer de l'argent dans les bureaux de poste, de signer des instruments et du service d'épargne du Fonds national d'épargne et de réserve, en plus de fournir une application appelée Postal Mob Web (ou Baridy Mob), à travers laquelle le paiement peut être effectué. [1].
- **Services de paiement en espèces :** qui comprend plusieurs services, notamment la demande d'une carte Gold et le retrait sans carte. Grâce à la carte, il est possible de retirer et de transférer de l'argent d'un compte à un autre, de consulter le solde, de payer des factures et de payer dans les centres commerciaux. Ils offrent également le service d'émission d'une carte distinctive, car elle est émise le même jour et peut également être délivrée au client.[1].

- **Services corporatifs :** Algérie Poste fournit également des services financiers postaux et monétaires aux professionnels et aux institutionnels, des services postaux, des services de proximité et un service de paiement à travers l'application destinée notamment à ces établissements.[1].
- **Services fournis en ligne:** La société offrent une large gamme de services en ligne, notamment Mob mail, le paiement de factures en ligne, la demande d'une carte Gold, la vente de timbres en ligne, la vente de livres en ligne, le chargement d'un compte *idoom*, le téléchargement de publications, le suivi d'un envoi de courrier privé, l'ouverture et la fermeture d'un compte privé, le rechargement d'une ligne mobile privée et son paiement. Prix des billets de voyage et des assurances en ligne.

2.3 Vision de la Poste Algérienne

L'institution Algérie Poste aspire à devenir leader dans la distribution de services de proximité et à domicile, et à être l'institution de référence en matière d'inclusion financière des citoyens, et l'élément le plus efficace en matière de courrier et de gros colis. Cela se fait en accélérant sa transformation numérique et la compatibilité entre ses réseaux physiques et numériques.[1].

2.4 Transformation numérique en Algérie Poste

Dans le but d'améliorer la qualité des services fournis aux bénéficiaires et afin de suivre l'évolution technologique, Algérie Poste s'est consacrée à la transformation numérique dans tous les bureaux de poste, en créant un système d'information complet qui permet à chaque utilisateur d'effectuer de nombreuses opérations dont il a besoin partout en Algérie.

3. Sites et Applications Web

Le World Wide Web (WWW), littéralement « toile d'araignée mondiale », communément appelé Web, est un système hypertexte public fonctionnant sur Internet. Le Web permet de consulter, avec un navigateur, des pages accessibles sur des sites. L'image de la toile d'araignée provient des hyperliens qui relient les pages Web entre elles. Le Web n'est qu'une des applications Internet distinctes d'autres applications telles que le courrier électronique, la messagerie instantanée et le partage de fichiers peer-to-peer. [2]

3.1 Historique

La communauté scientifique reconnaît la paternité du World Wide Web à Tim Berners Lie qui a créé le concept en 1990. Cet éminent chercheur est devenu président du Consortium 3W, basé à Genève. Internet est né de la rencontre de deux cultures, d'une part celle de l'ARPA (Advanced Research Project Agency) dont le mandat est de stimuler et encourager l'innovation technique aux Etats-Unis, d'autre part celle des informaticiens, hackers, imprégnés d'une culture de liberté, de valeurs d'autonomie individuelle et pour la vision paradoxale d'une habitude de partage des connaissances et de coopération.

3.2 Sites Webs

Un site Internet est constitué d'un ensemble de documents structurés, appelés pages web, stockés (hébergés) sur un ordinateur (serveur) connecté au réseau mondial (internet). Une page Web contient principalement du texte, et est souvent enrichie d'images, de sons, de vidéos et de liens vers d'autres pages Web.[2].

Il existe deux types de sites Web : les sites Web statiques et les sites Web dynamiques. [3].

- **Sites Web statiques** : Ce sont des « fenêtres ». Les visiteurs peuvent voir le contenu du site mais ne peuvent pas participer. Pour les réaliser, deux langages sont nécessaires : (X) HTML et CSS (en théorie, il est cependant possible d'utiliser uniquement HTML ou XHTML)
 - (X)HTML permet d'écrire le contenu de la page.
 - Le CSS permet de le décorer. .[3].



Figure 1.3 : Représentation d'un site Web statique

- **Web dynamiques** : Une page Web dynamique est une page Web générée à la demande, par opposition à une page Web statique. Ces sites ne sont pas seulement des « vitrines » : les visiteurs peuvent y participer. Pour le réaliser, nous avons également besoin de (X)HTML et CSS, mais nous avons également besoin d'un autre langage qui

créé dynamiquement les pages. Il existe plusieurs langages pour créer ces pages (Java, PHP, Python, etc...). Leur rôle est d'analyser les demandes des visiteurs puis de créer une réponse adaptée. [3].



Figure 1. 4 : Représentation d'un site Web dynamique

3.3 Applications Web

Une application web (aussi appelée web application, de l'anglais et français) est une application manipulable directement en ligne grâce à un navigateur web et qui ne nécessite donc pas d'installation sur les machines clientes, contrairement aux applications mobiles. De la même manière que les sites web, une application web est généralement installée sur un serveur et se manipule en actionnant des *widgets* à l'aide d'un navigateur web, via un réseau informatique (Internet, intranet, réseau local, etc.).

Certaines des principales caractéristiques et fonctionnalités des applications web peuvent comprendre:

- **Architecture Client-Serveur** : Les applications web suivent un modèle client-serveur, où le client (navigateur web) envoie des requêtes au serveur, et le serveur répond avec les données demandées ou effectue l'action demandée.
- **Compatibilité multi-plateformes** : Les applications web sont indépendantes de la plateforme, ce qui signifie qu'elles peuvent fonctionner sur n'importe quel système d'exploitation ou appareil avec un navigateur web compatible.
- **Interface utilisateur** : Les applications web ont une interface utilisateur (UI) qui est rendue dans le navigateur web de l'utilisateur, généralement en utilisant une combinaison de HTML, CSS et JavaScript.
- **Stockage des données** : Les applications web utilisent souvent une base de données ou d'autres mécanismes de stockage côté serveur pour stocker et gérer les données.

- **Accessibilité :** Les applications web peuvent être accessibles de n'importe où avec une connexion Internet, ce qui les rend très accessibles et pratiques pour les utilisateurs.
- **Mises à jour et maintenance :** Les applications web peuvent être mises à jour et maintenues côté serveur sans nécessiter que les utilisateurs installent des mises à jour sur leurs machines locales.

Les applications web font partie de l'évolution des usages et de la technologie du Web appelée Web 2.0. Elles sont construites en utilisant des technologies de développement web telles que HTML, CSS, JavaScript, et des langages de programmation côté serveur comme PHP, Python, Ruby ou Node.js. Elles s'appuient souvent sur des frameworks web et des bibliothèques pour faciliter le développement et fournir des fonctionnalités supplémentaires.

4. Technologies Web

Les technologies modernes du web permettent de créer des applications et services riches et interactifs accessibles depuis n'importe quel appareil connecté à internet. Au cœur de cette révolution se trouve un ensemble de langages, frameworks et principes de conception qui façonnent la façon dont le contenu et les fonctionnalités sont distribués sur le web.

La réalisation d'une application web s'appuie sur un ensemble d'outils et de technologies de développement web. Nous distinguons:

4.1 Technologies côté client *Front-End*

Les technologies web front-end comme HTML, CSS et JavaScript permettent de structurer, mettre en forme et animer les interfaces utilisateur des applications

- **HTML (Hyper Text Markup Language):** Il s'agit de l'ensemble des outils supportés par le navigateur HTML (Hyper Text Markup Language) HTML est un langage permettant de décrire la mise en page et la forme du contenu d'un document Web et incluant des hyperliens. Une page HTML est donc un simple fichier texte portant une extension .htm ou html, et dont l'ensemble des données est entre des balises (appelées marqueurs ou tags). L'hypertexte est un système utilisant le langage HTML, qui permet la création de liens hypertextes. Un document hypertexte est donc un document qui contient des hyperliens. Lorsque les documents sont non seulement textuels, mais aussi audiovisuels, on peut parler de système et de documents hypermédias. [3].

- **CSS (feuille de style en cascade) :** CSS est un langage de formatage qui décrit la présentation d'un document (positionnement des éléments, alignement, polices, couleurs, marges et espacements, bordures, images d'arrière-plan, etc.) écrit en HTML ou XML quelle que soit sa structure. .[3].
- **ReactJS :** React (aussi appelé React.js ou ReactJS) est une bibliothèque JavaScript libre. Elle est maintenue par Meta (anciennement Facebook) ainsi que par une communauté de développeurs individuels et d'entreprises depuis 2013. Le but principal de cette bibliothèque est de faciliter la création d'application web monopage, via la création de composants dépendant d'un état et générant une page (ou portion) HTML à chaque changement d'état. .[3].

4.2 Technologies côté serveur*Back-End*

Le côté serveur, également appelé *back-end*, joue un rôle crucial dans les applications web modernes. Le choix des technologies back-end dépend de plusieurs facteurs, notamment les besoins du projet, les exigences de performances, la disponibilité des ressources et l'expertise de l'équipe. Une architecture back-end solide et évolutive est essentielle pour garantir la fiabilité, la sécurité et l'évolutivité des applications web modernes.

Des langages comme Python, Ruby, PHP ou Node.js, peuvent gérer la logique métier, l'interaction avec les bases de données, la sécurité, les tâches d'arrière-plan et bien plus encore.

5. Architecture client/serveur

L'architecture client/serveur (Fig. 1.5) désigne un mode de communication entre plusieurs ordinateurs d'un réseau qui distingue un ou plusieurs postes clients du serveur ; chaque logiciel client peut envoyer des requêtes à un serveur. Un serveur peut être spécialisé en serveur d'applications, de fichiers, de terminaux ou de messagerie.

- Caractéristiques d'un serveur :

- il est passif (maître) ;
- il est à l'écoute, prêt à répondre aux demandes envoyées par les clients ;
- dès qu'une requête lui parvient, il la traite et envoie une réponse.

- Caractéristiques d'un client :

- il est actif (ou esclave) ;
- il envoie des requêtes au serveur ;

- il attend et reçoit les réponses du serveur.



Figure 1. 5 : L'architecture client/serveur

Le client et le serveur doivent bien entendu utiliser le même protocole de communication. Un serveur est généralement capable de servir plusieurs clients simultanément sur des architectures Web. Voici quelques-uns des principaux types d'architectures client/serveur utilisées dans le développement web :

- **Architecture à 1 niveau** : Une application 1-tier peut facilement être divisée en trois niveaux distincts : les données, le traitement de ces données et la couche de présentation.
 - **Couche de données** : regroupe les mécanismes de stockage et d'accès des données à ce qu'elles peuvent être utilisées par l'application au niveau du traitement.
 - **Couche traitement** : concerne à la fois les tâches à effectuer par l'application sur les données et les traitements nécessaires suite à une action de l'utilisateur : vérification, authentification, calculs divers, etc. [3].
 - **Couche de présentation** : gère l'affichage des données et les interactions de l'application avec l'utilisateur. Cette couche gère la saisie au clavier et à la souris et la présentation des informations à l'écran. Dans une application tierce, les trois couches applicatives sont intimement liées et exécutées sur le même ordinateur

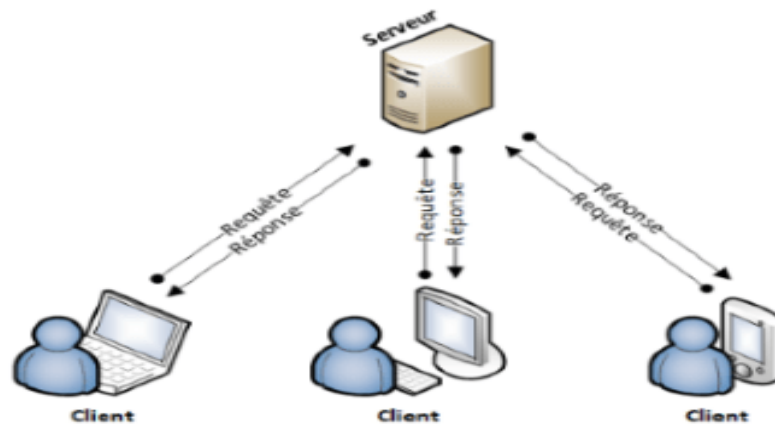


Figure 1. 6 : Architecture client/serveur à 1 niveau

- **Architecture à 2 niveaux :** L'architecture à deux niveaux (appelée aussi architecture 2-tiers) caractérise les systèmes client/serveur pour lesquels le client demande une ressource et le serveur la fournit directement, en utilisant ses propres ressources. Cela signifie que le serveur ne fait pas appel à une autre application pour fournir une partie du service. [3].

6. Conclusion:

Dans ce chapitre, nous avons donné un aperçu général sur Société Algérienne, en l'occurrence Algérie Poste, ainsi qu'un aperçu sur la nature d'Internet, les services qu'il fournit, y compris le Web, et les méthodes de développement du Web en termes de technologies et d'architectures récentes. Dans le chapitre suivant, nous présentons la reconnaissance vocale et ses principes.

CHAPITRE 2

LA RECONNAISSANCE DE LA PAROLE

1. Introduction

La reconnaissance automatique de la parole (Automatique Speech Recognition ASR). C'est une technique informatique qui permet d'analyser un mot ou une phrase captée au moyen d'un microphone pour la transcrire sous la forme d'un texte exploitable par une machine.

Ainsi que la synthèse vocale, l'identification du locuteur ou la vérification du locuteur, font partie des techniques de traitement de la parole. Ces techniques permettent notamment de réaliser des interfaces vocales c'est-à-dire des interfaces homme-machine (IHM) où une partie de l'interaction se fait à la voix. Parmi les nombreuses applications, on peut citer les applications de dictée vocale sur PC où la difficulté tient à la taille du vocabulaire et à la longueur des phrases, mais aussi les applications téléphoniques de type serveur vocal, où la difficulté tient plutôt à la nécessité de reconnaître n'importe quelle voix dans des conditions acoustiques variables et souvent bruyantes (téléphones mobiles dans des lieux publics) [4].

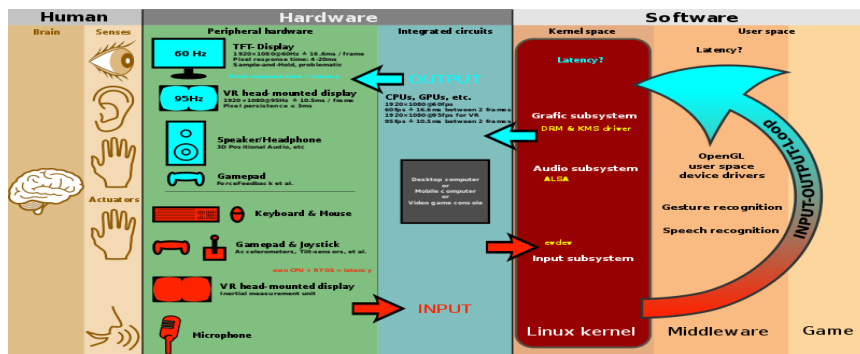


Figure 2.1 Linux kernel and gaming input-output latency

Dans Parole et dialogue homme-machine W. Minker et S. Bennacef expliquent que la reconnaissance automatique de la parole est un domaine complexe, car il existe une différence importante entre le langage formel, qui est compris et utilisé par les machines, et le langage naturel, que les humains utilisent. Le langage formel est structuré par des règles syntaxiques strictes et sans ambiguïté. À l'inverse, dans le langage naturel, des mots ou

des phrases peuvent avoir plusieurs sens selon l'intonation de l'énonciateur ou le contexte par exemple [4].

La reconnaissance vocale peut se rattacher à de nombreux plans de la science : traitement automatique des langues, linguistique, théorie des langages formels, théorie de l'information, traitement du signal, réseaux neuronaux, intelligence artificielle, etc [4].

2. Principes de base

Une phrase enregistrée et numérisée est donnée au programme de *reconnaissance vocale*. Dans le formalisme ASR, le découpage fonctionnel est le suivant :

- Le traitement acoustique (front-end) permet principalement d'extraire du signal de parole une image acoustique la plus significative possible sur des tranches de temps d'environ 30ms. Cette image se présente sous la forme d'un vecteur de caractéristiques (features extraction) de 10 à 15 composantes principales, auxquelles sont ajoutées les différences de premier et second ordre pour obtenir une taille de 30-45 en final.
- La numérisation du signal de la parole sous forme de vecteurs acoustiques qui constituent les données d'observation pour le système de reconnaissance. On utilise pour cela les techniques de traitement du signal : on découpe le signal en tranches de 30 ms en procédant pour chaque tranche à un décalage de 10 ms (technique de fenêtrage de Hamming) afin d'obtenir 10 ms de données significatives pour chaque vecteur. Le signal est alors numérisé et paramétré par une technique d'analyse fréquentielle utilisant les transformées de Fourier (par exemple MFCC, Mel-Frequency Cepstral Coefficients).
- L'apprentissage automatique qui réalise une association entre les segments élémentaires de paroles et les éléments lexicaux. Cette association fait appel à une modélisation statistique entre autres par modèles de Markov cachés (HMM, Hidden Markov Models) et/ou par réseaux de neurones artificiels (ANN, Artificial Neural Networks).
- La reconnaissance (back-end) qui, en concaténant les segments élémentaires de paroles précédemment appris, reconstitue le discours le plus probable. Il s'agit donc d'une correspondance de motif (pattern matching) temporelle, réalisée souvent par l'algorithme de déformation temporelle dynamique (en anglais DTW, Dynamic Time Warping). [4]

3. Historique

Les travaux sur la reconnaissance vocale datent du début du XX^e siècle. Le premier système pouvant être considéré comme faisant de la reconnaissance vocale date de 1952. Elle est évoquée dans Le Premier cercle de l'écrivain dissident soviétique Alexandre Soljenitsyne, comme un outil de répression au service de Staline.

Ce système électronique développé par Davis, Biddulph, and Balashek aux laboratoires Bell Labs était essentiellement composé de relais et ses performances se limitaient à reconnaître des chiffres isolés (voir référence). La recherche s'est ensuite considérablement accrue durant les années 1970 avec les travaux de Jelinek chez IBM (1972-1993). La société Threshold Technologies fut la première à commercialiser en 1972 un système de reconnaissance d'une capacité de 32 mots, le VIP100. Aujourd'hui, la reconnaissance vocale est un domaine à forte croissance grâce à la déferlante des systèmes embarqués [4].

4. Matériel technique

La reconnaissance vocale repose sur un ensemble de composants matériels travaillant de concert pour capturer, traiter et interpréter les signaux vocaux. Voici les principaux éléments matériels impliqués [4]:

- **Microphone** : C'est l'élément clé qui convertit les ondes sonores en signaux électriques analogiques. Il existe différents types de microphones avec des caractéristiques variées en termes de sensibilité, directivité et réponse en fréquence.
- **Convertisseur analogique-numérique (CAN)** : Le signal analogique provenant du microphone doit être converti en données numériques. Le CAN échantillonne le signal à une fréquence donnée et le quantifie en valeurs numériques.
- **Processeur de signal numérique (DSP)** : Les DSP sont des processeurs spécialisés optimisés pour le traitement numérique des signaux. Ils effectuent des opérations comme le filtrage, l'amplification et la mise en forme du signal vocal numérisé pour le préparer à la reconnaissance.
- **Unité de traitement vocale** : De nombreux systèmes de reconnaissance vocale intègrent des puces ou processeurs dédiés au traitement du signal vocal, offrant des performances accrues par rapport au traitement sur CPU généraliste.
- **Accélérateurs matériels** : Pour accélérer les lourds calculs d'apprentissage machine nécessaires à la reconnaissance vocale, on utilise de plus en plus des accélérateurs comme les GPU, TPU ou autres puces dédiées à l'IA.

- **Mémoire** : Une quantité suffisante de mémoire vive et de stockage est nécessaire pour charger et exécuter les modèles d'apprentissage machine volumineux utilisés pour la reconnaissance vocale.

5. Les modèles

Un tel système s'appuie sur trois modèles principaux :

- **Modèle de langage** : ce modèle donne la probabilité $P(W)$ de chaque suite de mots W dans le langage cible ;
- **Modèle de prononciation** : ce modèle donne pour chaque suite de mots W , la ou les prononciations possibles H avec leurs probabilités $P(H|W)$;
- **Modèle acoustico-phonétique** : ce modèle estime la probabilité $P(X|H)$ de la séquence observée de vecteurs acoustiques X étant donné une prononciation possible d'une séquence de mots donnée.

La combinaison de ces trois modèles permet de calculer la probabilité de toute suite de mots étant donné un signal vocal observé. La reconnaissance de la parole consiste à trouver la suite de mots qui a la probabilité la plus élevée. Formellement la solution au problème est la suite de mots W qui maximise l'expression mathématique suivante :

$$P(W) = \sum_H P(H|W)P(X|H)$$

Pour caler ces modèles en vue d'une application, il faut utiliser une grande quantité de corpus annoté. Le corpus doit correspondre aux conditions d'utilisation du système visé [4]

6. Classification

Les systèmes de reconnaissance de la parole peuvent être classifiés selon plusieurs axes :

- **Le type de signal** : signal bruité ou signal non bruité (ex. : microphone casque avec réduction de bruit), signal téléphonique (téléphone fixe ou mobile) ou large bande, signal compressé ou non...
- **Le type de modèle acoustique** : modèle monolocuteur (ex. : dictée vocale), modèle multilocuteur (*speaker-independent* en anglais)
- **La nature des enregistrements** : dictée de texte, commande vocale, dialogue homme-machine, message téléphonique, radio, TV, podcast, etc. ;
- **La langue** : La taille du vocabulaire et la complexité du modèle de langage sont directement liées à la langue et à la nature des données à traiter, de quelques dizaines de mots pour des

commandes vocales à quelques centaines de milliers de mots pour couvrir une langue comme le français ou l'allemand [4].

7. Logiciels de reconnaissance de la parole

Les systèmes de reconnaissance parole modernes utilisent des modèles du langage qui peuvent nécessiter des giga-octets de mémoire ce qui les rend impraticables, en particulier sur les équipements mobiles. Pour cette raison, la plupart des systèmes de reconnaissance vocale modernes sont en fait hébergés par des serveurs distants et nécessitent une connexion internet et l'envoi à travers le réseau du contenu parole.

Mozilla a lancé un projet communautaire, Common Voice, visant à recueillir des échantillons de la parole dans une base de données libres, pour entraîner des moteurs de reconnaissance parole non-propriétaires. [4]

8. Comment fonctionne la reconnaissance de la parole ?

La reconnaissance de la parole se décline en trois étapes :

1. L'analyse acoustique qui permet de découper le message vocal en vecteurs acoustiques ingérables par la machine,
2. Le machine learning qui associe ensuite les fréquences sonores à des mots.
3. L'analyse de la parole qui combine trois modèles (un modèle de langage, un modèle de prononciation, un modèle acoustico-phonétique) en vue d'identifier les suites de mots les plus probablement prononcés par le locuteur.

9. Conclusion

Dans ce chapitre, nous avons présenté un aperçu sur la reconnaissance parole qui avise à décrypter une expression vocale énoncée par un homme captée par un microphone pour la transcrire sous forme d'un texte exploitable par une machine. La démarche nécessite l'intégration de plusieurs axes de recherches scientifiques pointus, débutant par le traitement automatique acoustique, la numérisation du signal, l'apprentissage automatique via la modélisation du statique, et la reconnaissance. Dans le chapitre suivant nous présentons la conception de l'application web proposée.

CHAPITRE 3

ANALYSE ET CONCEPTION

1. Introduction

La phase de conception constitue la première étape cruciale pour la réussite d'un projet. Elle a pour objectif de formaliser de manière claire et non ambiguë le fonctionnement futur du système afin de faciliter sa réalisation. Pour ce faire, diverses méthodes existent pour formaliser les étapes préliminaires du développement.

Dans ce chapitre, nous allons procéder à un premier découpage de notre application, définir les différentes étapes de conception, identifier les relations entre les cas d'utilisation, et élaborer certains diagrammes UML (Unified Modeling Language) tels que le diagramme de cas d'utilisation pour spécifier les besoins de notre système, le diagramme de séquence pour décrire les interactions entre les acteurs et notre système, ainsi que le diagramme de classe pour définir l'architecture statique de notre système.

2. Aperçu générale sur UML

UML, ou Langage de Modélisation Unifié, est un langage de modélisation graphique standard largement utilisé dans le développement logiciel pour spécifier, visualiser, construire et documenter les éléments d'un système orienté objet.

2.1 UML

L'Unified Modeling Language (UML) est un ensemble de diagrammes intégrés utilisés par les développeurs informatiques pour visualiser les objets, les états et les processus dans un logiciel ou un système.

Ce langage de modélisation permet de créer des modèles qui servent de base pour un projet, garantissant ainsi une architecture d'informations structurée. Il aide également les développeurs à communiquer de manière compréhensible avec des experts externes en présentant une description claire d'un système.

UML est principalement utilisé dans le développement de logiciels orientés objet, mais avec les améliorations introduites dans sa version 2.0, il peut également être utilisé pour représenter des processus de gestion [5].

2.2 Développement du langage de modélisation unifié (UML)

Avant l'introduction d'UML la programmation orientée objet (POO) était déjà largement adoptée dans le développement logiciel. La POO repose sur le concept que tout est un objet, où les composants d'un programme interagissent entre eux en envoyant des messages. Chaque objet est une instance d'une classe supérieure qui définit son comportement. Les objets sont composés de données et de code, avec les données organisées en attributs et le code déterminant les méthodes.

Dans les années 1980 et 1990, de nombreuses méthodes et langages ont été développés pour représenter la POO, créant ainsi une diversité de méthodes différentes. Afin d'unifier ces langages, James Rumbaugh, Grady Booch et Ivar Jacobson ont décidé de fusionner plusieurs langages existants en un langage commun et standardisé, (ce qui a abouti à la création d'UML). [5]

UML est un puissant moyen de communication offrant un cadre structuré pour l'analyse des systèmes. Il permet d'obtenir une précision accrue dans la représentation des concepts, favorise l'utilisation d'outils spécifiques qui simplifient la création et la manipulation des modèles, facilite la compréhension des représentations abstraites complexes, et est gage de stabilité, garantissant la cohérence et la fiabilité des modèles. Sa polyvalence et sa flexibilité en font un langage universel adaptable à différents domaines d'application.

Cependant, l'adoption et la pratique de l'UML nécessitent un apprentissage et une période d'adaptation. Il est important également de noter que l'UML ne couvre pas tous les processus essentiels à la réussite d'un projet.

2.3 Diagrammes d'UML

Un diagramme UML est une représentation graphique qui traite une partie spécifique du modèle. Chaque classe de diagramme UML a une structure et une sémantique spécifique. UML propose plusieurs types de diagrammes pour modéliser différents aspects d'un système :

1. **Diagrammes structurels** : Décrivent les éléments statiques du système
 - Diagramme de classes
 - Diagramme d'objets
 - Diagramme de composants
 - Diagramme de paquetages
2. **Diagrammes comportementaux** : Représentent les aspects dynamiques et fonctionnels du système
 - Diagramme de cas d'utilisation

- Diagramme d'activités
 - Diagramme d'états-transitions
 - Diagramme de séquence
3. **Diagrammes d'interaction** : Modélisent le flux de contrôle et les échanges de messages
- Diagramme de séquence
 - Diagramme de communication
 - Diagramme d'interaction d'ensemble
 - Diagramme de temps
3. **Diagrammes d'implémentation** : Décrivent l'architecture et la configuration physique du système
- Diagramme de composants
 - Diagramme de déploiement

Pour notre projet, nous présenterons trois diagrammes d'UML : Diagramme de cas d'utilisation, Diagrammes de séquence et Diagrammes d'activités.

3. Présentation générale de notre application

De nos jours, les applications mobiles et web facilitent de nombreuses tâches du quotidien. Une fonctionnalité pratique serait de pouvoir convertir rapidement des montants monétaires en chiffres, que ce soit pour vérifier une facture, un relevé bancaire ou libeller un chèque.

C'est dans cette optique que cette application web de conversion de montants a été développée. Son interface simple et intuitive permet à l'utilisateur de saisir un montant en Centimes de Dinars (ou toute autre monnaie), par exemple 150000 centimes, et d'obtenir instantanément sa conversion numérique en Dinars, en chiffres, soit 1500 DA, et en lettres, soit "mille cinq cents dinars". L'application peut effectivement avoir une interface multilingue.

Mais l'application ne s'arrête pas là. Elle intègre également une fonctionnalité de reconnaissance vocale dernier cri. L'utilisateur peut prononcer le montant en Centimes à haute voix, et l'application le transcrit en temps réel, avant d'afficher la conversion en Dinars chiffrée et en lettres du montant correspondant. Plus besoin d'avoir recours à une autre personne pour écrire le chiffre correctement, ni de faire des erreurs pendant le remplissage des factures et des chèques.

Que vous soyez un particulier ou un professionnel amené à manipuler fréquemment des montants d'argent, cette application deviendra vite un outil incontournable. Son interface

épurée, ses capacités de reconnaissance vocale et son algorithme de conversion optimisé offrent une expérience fluide et économiseront un temps précieux.

Dans les sections suivantes, nous explorerons les différents diagrammes de conception de notre application en utilisant le langage UML.

4. Diagrammes de cas d'utilisation

Les diagrammes de cas d'utilisation permettent de représenter le comportement anticipé d'un système à l'avenir. Ces modèles sont adaptés non seulement aux systèmes logiciels, mais aussi pour prédire les procédures au sein d'une entreprise [5].

Le diagramme de cas d'utilisation ci-dessous (Figure 3.1) présente de manière visuelle les différentes interactions entre les acteurs et le système dans le cadre de notre application :

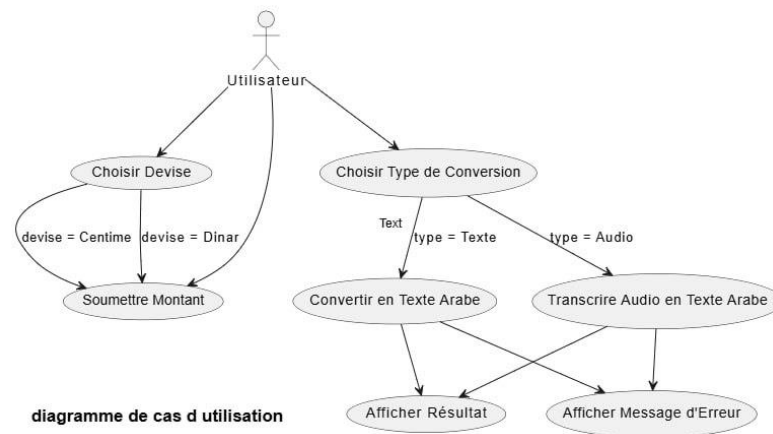


Figure 3.1 Le diagramme de cas d'utilisation

Après avoir recensé l'ensemble des besoins, nous avons distingué les différents acteurs suivants :

- **Utilisateur** : Cet utilisateur a la possibilité de parcourir la page d'accueil en général et saisir la somme en centimes (saisie par des chiffres, ou avec un audio (via un microphone)).

5. Diagrammes de séquences

Le diagramme de séquences est un type de diagramme d'interaction qui représente l'échange de messages entre les objets. Dans ce diagramme, les objets sont modélisés sous forme de lignes de vie, ce qui le rend similaire à d'autres diagrammes comportementaux tels que le diagramme d'activité. Cependant, contrairement à ces derniers, le diagramme de séquence ne vise pas à fournir une vue globale du comportement d'un système, mais plutôt à

présenter en détail un scénario spécifique parmi plusieurs scénarios possibles. Il établit une chronologie et utilise une ligne en pointillés pour représenter le passage dans le temps.

Dans UML 2.0, les messages sont représentés de manière synchrone (flèche avec une tête pleine) ou asynchrone (flèche avec une tête ouverte). Les messages synchrones bloquent le canal de communication jusqu'à ce qu'ils reçoivent une réponse de l'objet cible. Ils définissent le comportement sous forme d'opérations synchrones. Les messages asynchrones contrôlent l'objet source appelant et peuvent être des opérations asynchrones ou des signaux (des paquets de données envoyés entre les actions). [5]

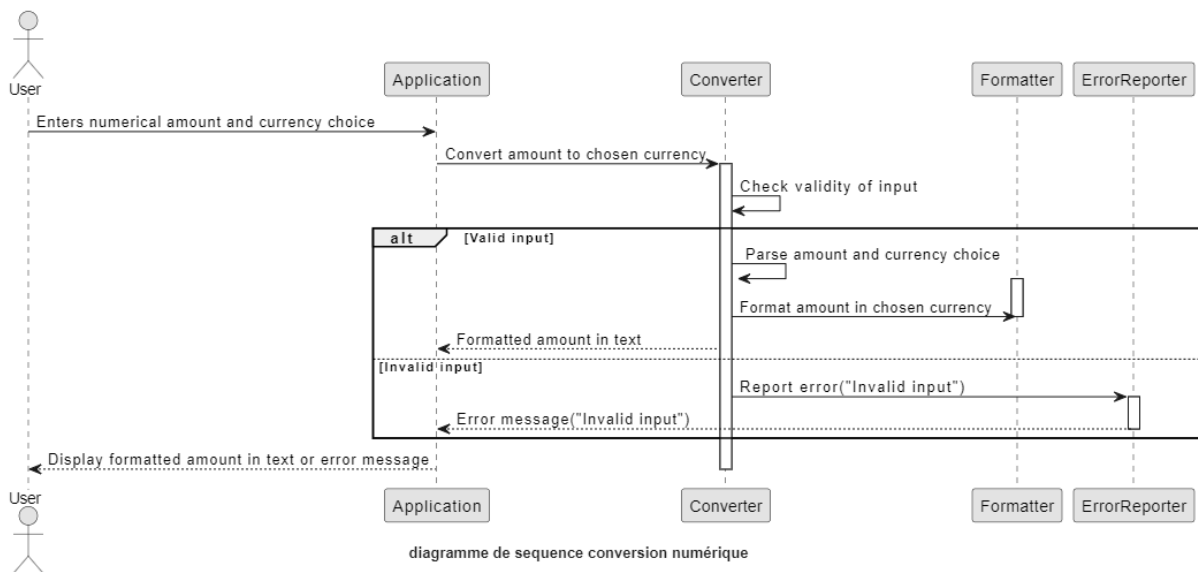


Figure 3.2 :Le diagramme de séquence conversion numérique

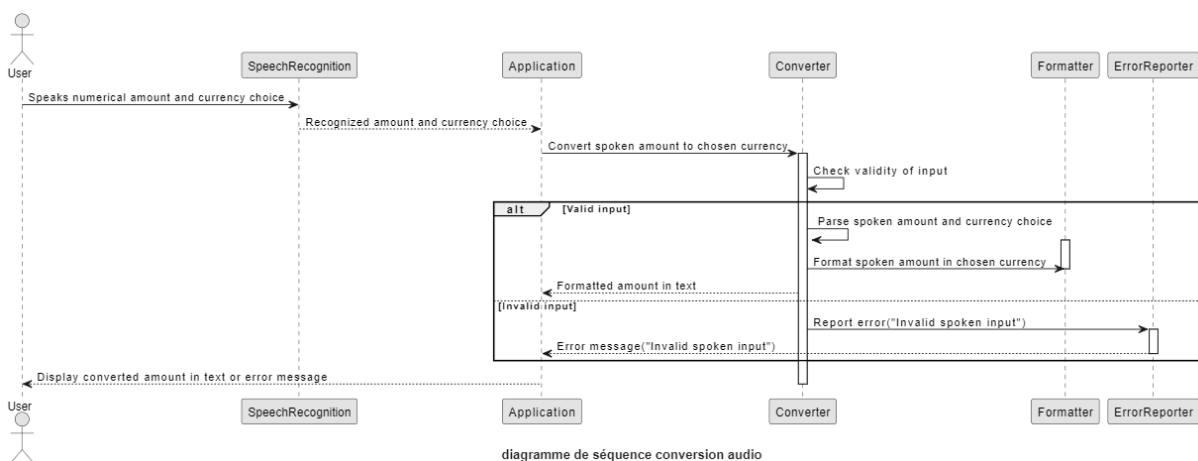


Figure 3.3 :Le diagramme de séquence conversion audio

Dans notre application, le diagramme de séquence (Figure 3.2) et (Figure 3.3) illustre le scénario d'utilisation de l'application. L'utilisateur choisit le mode de saisie (saisie en centimes/dinars), et il acquiesce ses informations via le système pour la conversion, si le

système constate des erreurs de rédaction et de forme, il enverra un message d'erreur, sinon ils sont validé, et la conversion sera réussie. De la même façon, pour le scénario de la reconnaissance vocale, l'utilisateur enregistre le montant en audio, qui sera traité via le système pour le traduire en chiffres, après avoir effectué la conversion. Si le système constate des erreurs de rédaction et de forme, il enverra un message d'erreur.

6. Diagrammes d'activités

Le diagramme d'activité est un diagramme comportemental d'UML, permettant de représenter le déclenchement en fonction des états du système et de modéliser des comportements (multithreads ou multiprocessus). Le diagramme d'activité est également utilisé pour décrire un flux de travail (workflow). [6]

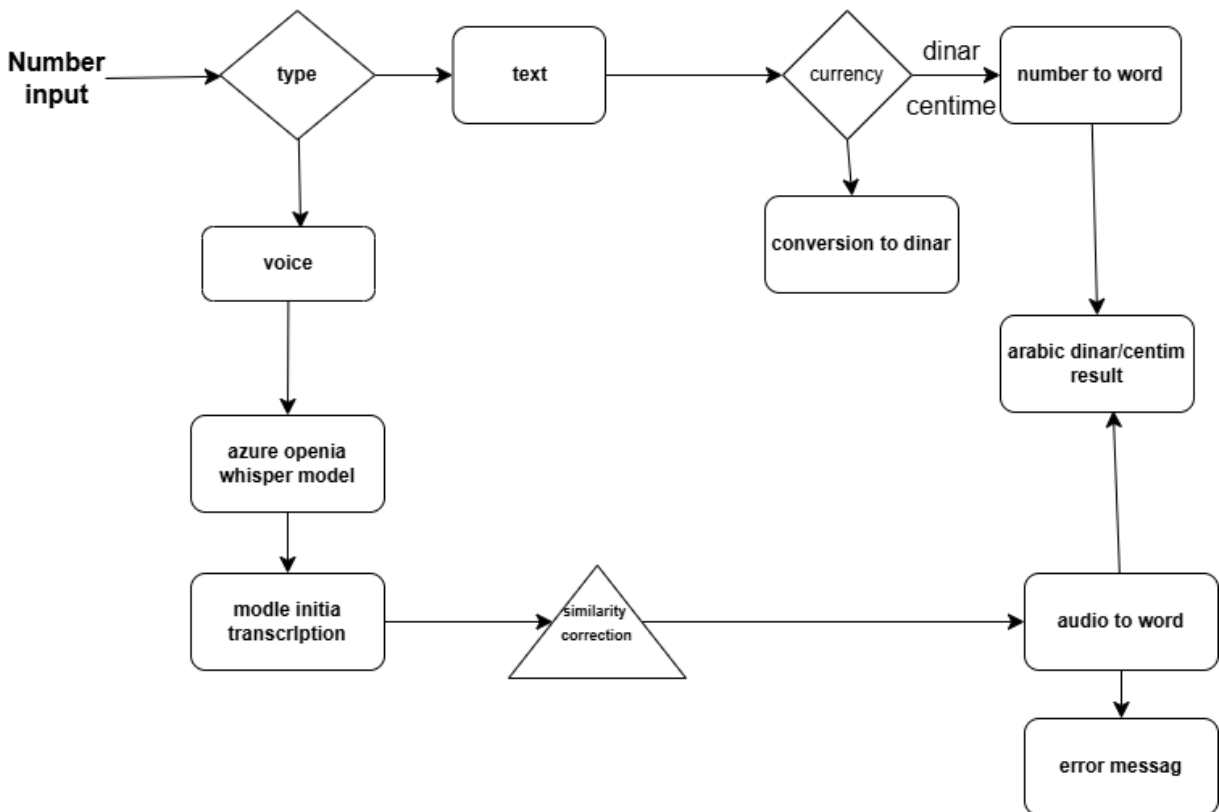


diagramme d activity BACKEND logique

Figure 3.4 : Le diagramme d'activités

Pour notre application, le diagramme d'activités est présenté sur la figure 3.4.

1. L'application démarre et affiche l'interface de conversion de montants.
2. L'utilisateur a deux options :
 - Saisir le montant en toutes lettres via une entrée textuelle

- Activer la reconnaissance vocale pour prononcer le montant
3. Une fois le montant saisi (par texte ou vocalement), l'application convertit ce montant en sa représentation chiffrée.
 4. Le montant converti en chiffres est affiché à l'utilisateur.

Ce diagramme d'activité UML illustre le flux principal de l'application, capturant les deux modes d'entrée (textuel et vocal) et le processus de conversion du montant. Il peut être complété avec des chemins alternatifs, des gardes de condition, etc. pour représenter d'autres cas d'utilisation et comportements.

7. Conclusion

Dans ce chapitre, nous avons exposé le langage de modélisation UML, et la réalisation de la modélisation des différents diagrammes tels que les diagrammes de cas d'utilisation, les diagrammes des séquences et les diagrammes d'activités. Cette étape de modélisation nous a permis d'obtenir une vision globale du comportement théorique des fonctionnalités opérées par notre application. Cette base théorique nous guidera lors du développement de l'application, qui sera abordé dans le chapitre suivant.

CHAPITRE 4

IMPLEMENTATION

1. Introduction :

Ce chapitre a pour objectif de décrire les étapes de développement de notre application web et de fournir une explication détaillée des outils utilisés pour son implémentation. Nous aborderons ce sujet avec un ton formel et professionnel.

2. Outils utilisés pour le Back-End

Les outils utilisés dans ce projet pour la programmation côté back-end : comprennent le Framework web et le langage de programmation. Ces outils sont essentiels pour assurer le bon fonctionnement du système. Pour cela, nous avons choisi le Framework Flask et le langage de programmation Python. Chacun de ces outils a été choisi en fonction de ses fonctionnalités et de sa compatibilité avec les autres outils utilisés dans le projet. Une description détaillée de ces outils sera donnée dans les sections suivantes.

2.1 Pourquoi on utilise Flask ?

Il a été conçu pour faciliter le démarrage rapide de projets web, tout en offrant les outils nécessaires pour développer des applications complexes. Flask est souvent considéré comme un "micro-Framework" car il ne fournit que les composants de base pour le routage des URL, les requêtes et les réponses. [7]

2.2 Flask requête/réponses processus

Cycle de demande/réponse lors du développement d'une application Web avec Flask, nous nous concentrons sur le traitement des demandes reçues et sur la fourniture de réponses (Figure 4.1).

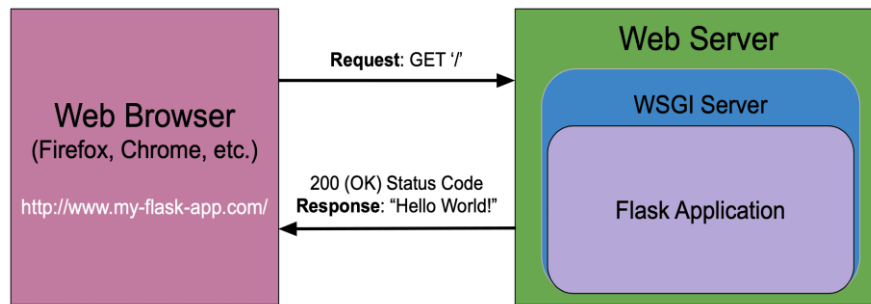


Figure 4.1: Le request_response_cycle_1.

Lorsque vous saisissez une URL dans votre navigateur Web, la requête est envoyée au serveur Web (comme Nginx ou Apache), qui la transmet ensuite à un serveur WSGI (comme Gunicorn). Le serveur WSGI est l'interface standard entre un serveur web et une application Python. Dans notre cas ici, une application Flask s'exécute du côté Python de l'interface WSGI. L'application Flask traite la demande et génère une réponse, qui est renvoyée au navigateur Web via le serveur WSGI et le serveur Web. [8].

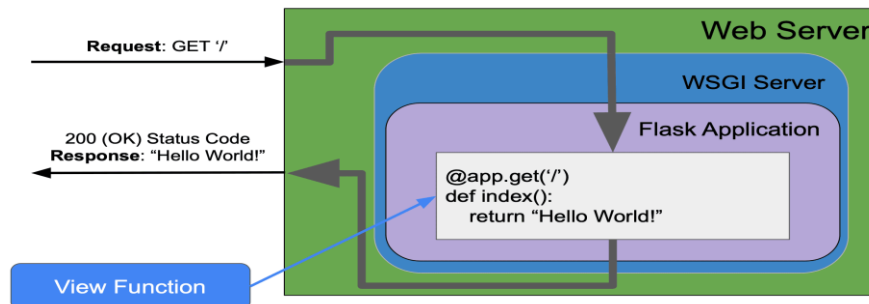


Figure 4.2: Le request_response_cycle_2.

Python, créé en 1991 par Guido van Rossum, est un langage de programmation interprété, haut niveau et orienté objet. Sa popularité croissante s'explique par sa simplicité, sa lisibilité et sa facilité d'apprentissage. Python trouve une large application dans divers domaines tels que le développement web, la science des données, l'intelligence artificielle et l'automatisation des tâches. [9]



Figure 4.3: Le logo de python.

3. Outils utilisé pour le Front-End

Les outils suivants ont été utilisés dans le projet pour la programmation côté front-end :

- **JavaScript (JS) :** est un langage de programmation léger, interprété ou compilé juste à temps avec des fonctions de première classe. Bien qu'il soit surtout connu comme langage de script pour les pages Web, de nombreux environnements autres que les navigateurs l'utilisent également, tels que Node.js, Apache CouchDB et Adobe Acrobat. JavaScript est un langage dynamique, multi-paradigmes, monothread, basé sur un prototype, prenant en charge les styles orientés objet, impératifs et déclaratifs (par
- **ReactJS :** React est une bibliothèque qui ne gère que l'interface de l'application, cette interface étant considérée comme la vue dans le modèle MVC. Elle peut ainsi être utilisée avec une autre bibliothèque ou un Framework MVC comme AngularJS. La bibliothèque se démarque de ses concurrents par sa flexibilité et ses performances, en travaillant avec un DOM virtuel et en ne mettant à jour le compte rendu dans le navigateur qu'en cas de nécessité. [11] .

4. Visuel studio code pour le web

Code pour le web l'IDE Visual Studio est un panneau de lancement créatif que vous pouvez utiliser pour modifier, déboguer et générer du code, puis publier une application.

En plus de l'éditeur et du débogueur standard fournis par la plupart des IDE, Visual Studio inclut des compilateurs, des outils de complétion de code, des concepteurs graphiques et bien d'autres fonctionnalités pour améliorer le processus du développement de logiciels. [12]

5. Les Bibliothèques utilisées:

1-NumPy : NumPy est une bibliothèque Python très populaire utilisée pour le calcul numérique. Elle fournit des structures de données et des fonctionnalités pour travailler avec des tableaux multidimensionnels, ainsi que des fonctions mathématiques pour effectuer des opérations sur ces tableaux. NumPy est largement utilisée dans le domaine de la science des données, de l'apprentissage automatique, de la modélisation mathématique, et dans d'autres domaines où des calculs numériques efficaces sont nécessaires.

2-Levenshtein : Levenshtein est une bibliothèque Python qui implémente l'algorithme de distance de Levenshtein. Cet algorithme calcule la distance entre deux chaînes de caractères en mesurant le nombre minimum d'opérations nécessaires pour les transformer en une autre, où une opération peut être une insertion, une suppression ou une substitution d'un caractère. Cette bibliothèque est souvent utilisée dans le traitement automatique du langage naturel

(TALN) et dans d'autres applications où la similarité ou la distance entre des chaînes de caractères est importante.

3-re : "re" est un module Python intégré qui fournit des fonctionnalités pour travailler avec des expressions régulières. Les expressions régulières sont des motifs de recherche flexibles qui permettent de rechercher et de manipuler des chaînes de caractères en fonction de motifs spécifiques. Le module "re" est largement utilisé pour la validation de données, l'analyse de texte, le nettoyage de données et d'autres tâches qui impliquent la manipulation de chaînes de caractères complexes.

4-numbers_to_words : Ce type de bibliothèque est souvent utilisé dans les applications où les nombres doivent être présentés sous forme de mots, tels que les systèmes de facturation, les générateurs de rapports ou les applications d'apprentissage des langues.

5-Whisper : Whisper est un modèle d'apprentissage automatique pour la reconnaissance et la transcription vocales, type de Transcription software; et est une approche simple de bout en bout, implémentée sous la forme d'un transformateur codeur-décodeur, l'audio d'entrée est divisée en morceaux de 30 secondes, convertir en un spectrogramme log- Mel ; Fondation model; Acoustique model créé par Open AI et publié pour la première fois en tant que logiciel open source en septembre 2022. [13]

Pour le modèle Whisper, il faut passer par des étapes dont la première étapes est la création d'un compte azure Open AI Studio (Voir Figures 4.4 et 4.5).

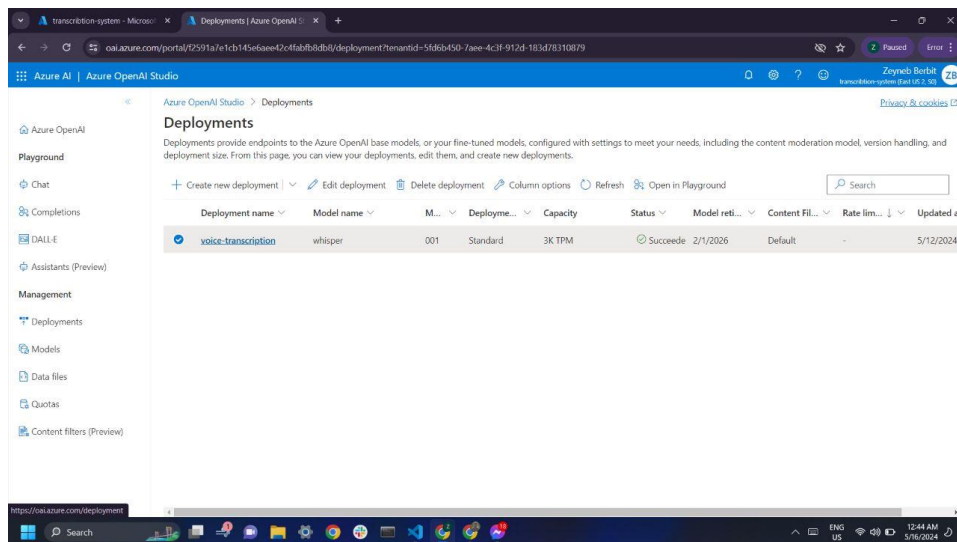


Figure 4.4 : Création d'un compte Azure Open AI Studio

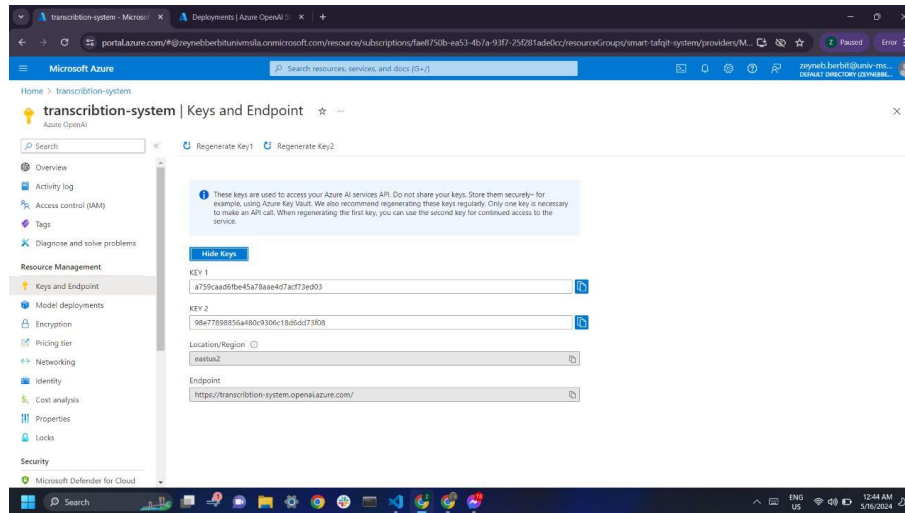


Figure 4.5 : Inscription au système

6. Fonctions utilisées

Ce code fournit des fonctions pour différentes conversions liées à l'unité monétaire et à la voix en utilisant la langue arabe.

1. `centime_to_dinar (centime_input)`: Cette fonction prend un montant en centimes et le convertit en dinar en le divisant par 100. Elle retourne ensuite la valeur en dinar.
2. `numbers_to_words_conversion (dinar_number)` : Cette fonction prend un montant en dinar et le convertit en mots en utilisant la fonction `dinar_number_to_arabic_words` importée depuis le module `numbers_to_words`. Elle retourne une chaîne contenant le montant en mots suivi de "دينار جزائري".
3. `arabic_words_to_voice (arabic_words)`: Cette fonction prend des mots en arabe et devrait effectuer une conversion ultérieure en voix. Actuellement, elle renvoie simplement les mots en arabe.
4. `voice_to_arabic_words (voice)`: Cette fonction devrait convertir la voix en mots en arabe. Le traitement réel de la voix n'est pas implémenté dans ce code, mais des étapes potentielles sont décrites en commentaire. Cela inclut la transcription de la voix en texte en utilisant Whisper AI, la correction automatique des mots arabes, l'extraction des nombres du texte transcrit et leur conversion en mots, et enfin la vérification de chaque mot pour assurer son existence dans un dictionnaire.
5. `Similarity_correction` : Dans cette fonction, chaque mot dans la transcription est parcouru. Pour chaque mot, le code recherche le mot le plus similaire dans la liste de mots de référence en utilisant la distance de Levenshtein pour calculer la similarité.
6. `Arabic_audio_transcript` : est définie. Cette fonction prend un nom de fichier audio comme argument. À l'intérieur de cette fonction, un modèle de transcription audio est

chargé à l'aide de la bibliothèque `whisper`. Le fichier audio est transcrit en texte en utilisant le modèle chargé. Ensuite, les nombres présents dans la transcription sont extraits à l'aide d'une expression régulière.

7. `Dinar_number_to_arabic_words` : Chaque nombre est converti en mots arabes à l'aide de la fonction.

7. Démarrage de l'application

Pour démarrer notre application web, nous avons suivi quelques étapes simples :

- Tout d'abord, il faut démarrer le back end, nous avons utilisé `cd backend`, après l'exécution on ajoute aussi la commande `Python -mflask --run --debug`, et on exécute :

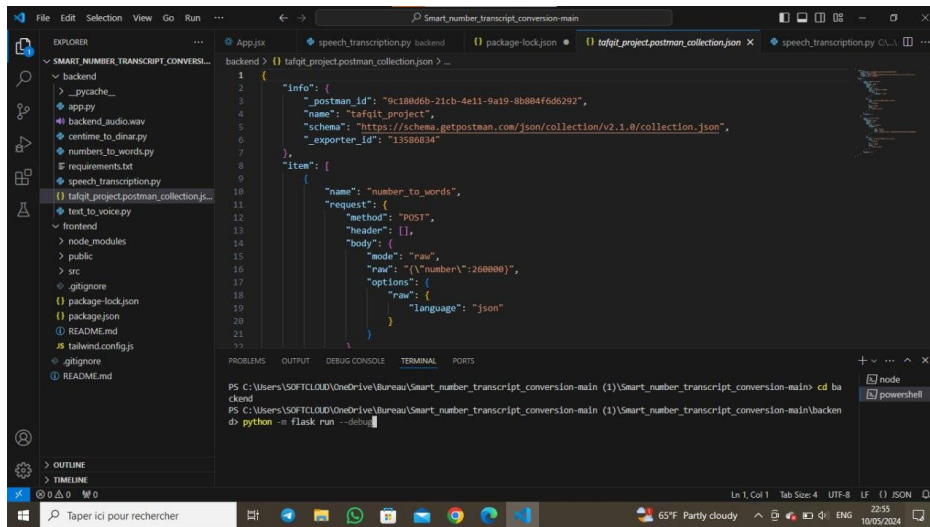


Figure 4.6: le Démarrage de back end 1.

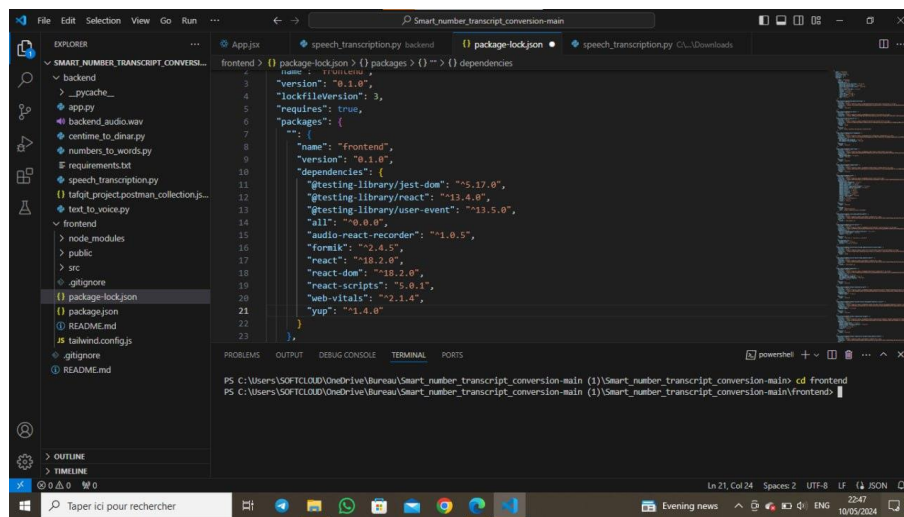


Figure 4.7: le Démarrage de back end 2.

- Ensuite, nous avons lancé le front end , premièrement en ajoute new terminal puis en exécutant la commande cd front_end. Ensuite on ajoute la commande npm start, et on exécute :

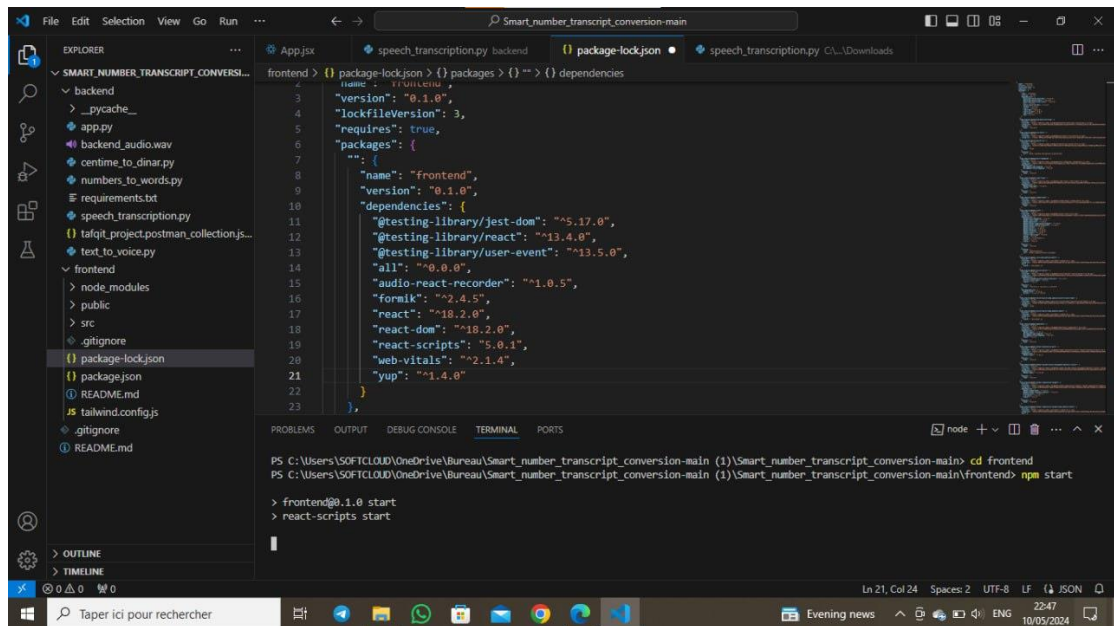


Figure 4.8: le Démarrage de front end 1.

8. Page d'insertion :

8.1 Entrer les chiffre par écrits :

Sur la page d'insertion ,l'utilisateur trouve les informations nécessaire pour saisissse les chiffre par écrits et en respectant le formulaire requis ,comme la figure suivante le montre :

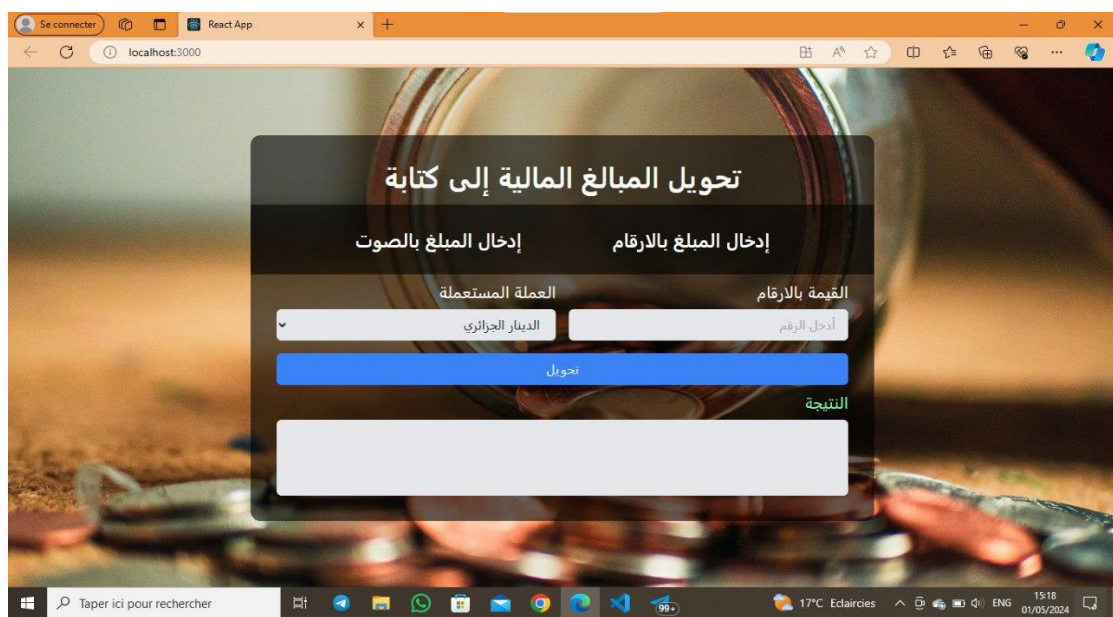


Figure 4.9: Entrer les chiffre par écrits

en cas d'erreur dans les chiffres apparaissant un message d'erreur , comme les figures suivante le montre :

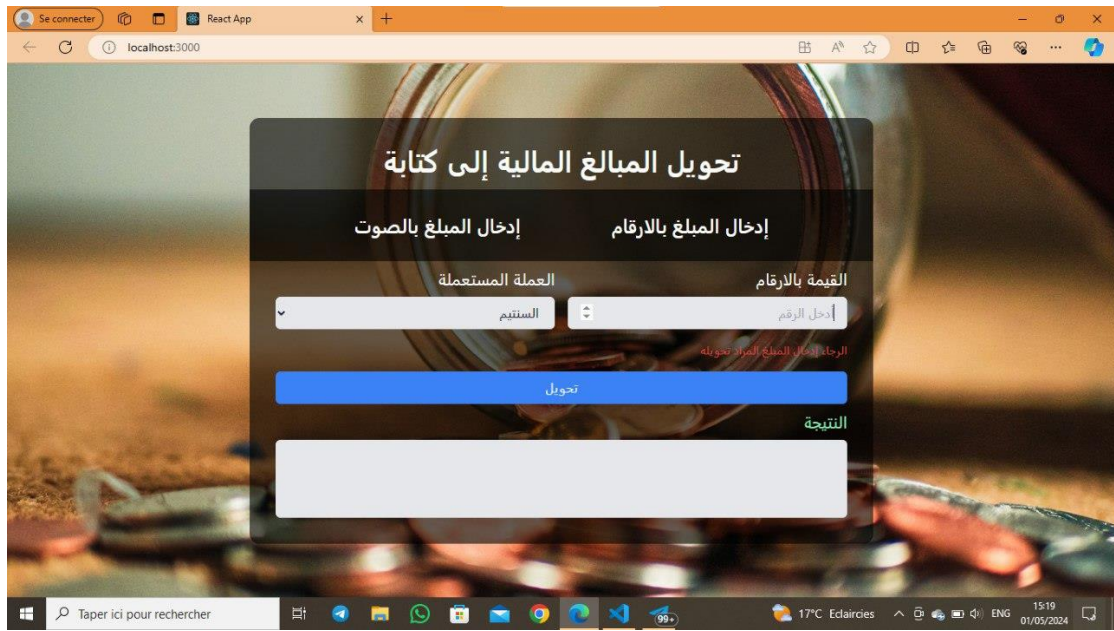


Figure 4.10:avertissement il a une erreur 1

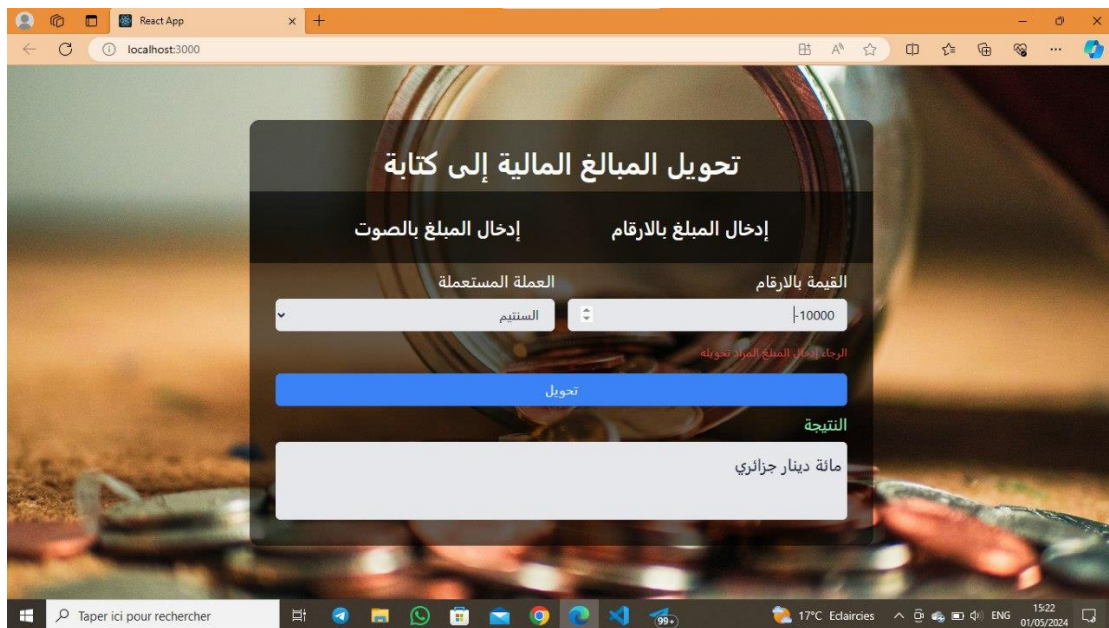


Figure 4.11: avertissement il a une erreur2.

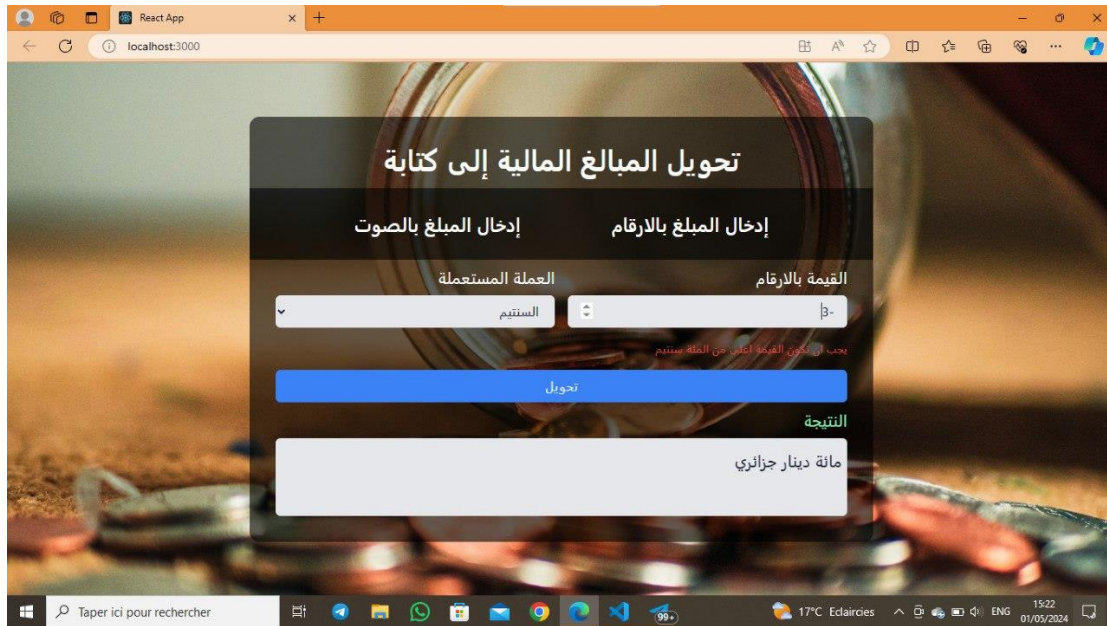


Figure 4.12: avertissement il a une erreur3.

Choisir la devise:

L'utilisateur choisit la devise correspondante, ensuite la saisie de la somme, comme le montre la figure suivante :

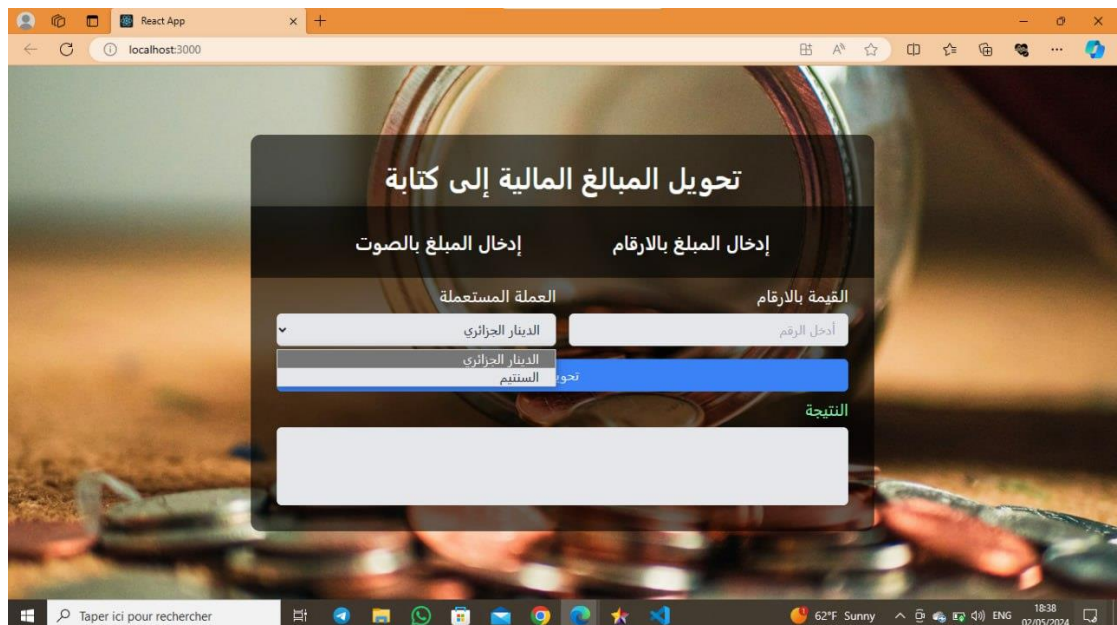


Figure 4.13: Choisir la devise.

Puis, quand on fait la saisie du chiffre, on constatera le résultat dans le bar de (النتيجة) :

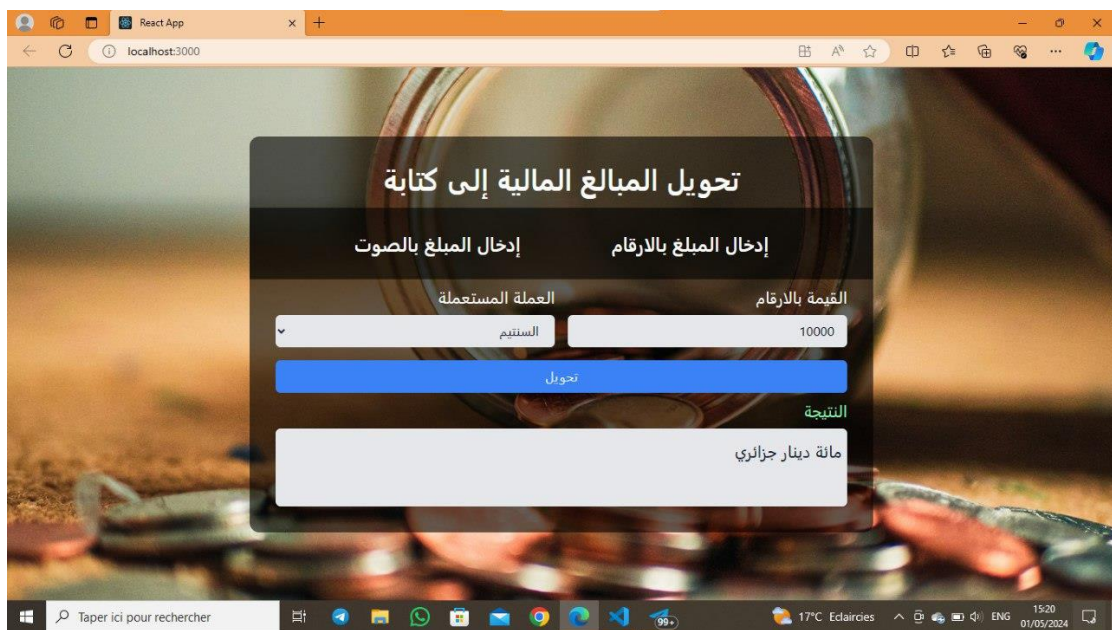


Figure 4.14:Le résultat par des lettres

8.2 Entrer les chiffre par voix(message vocal) :

Sur la page d'insertion, l'utilisateur trouve les informations nécessaires pour la saisie des chiffres par la voix et en respectant le formulaire requis, comme la figure suivante le montre :

Figure 4.15: Entrer les chiffre par voix 1

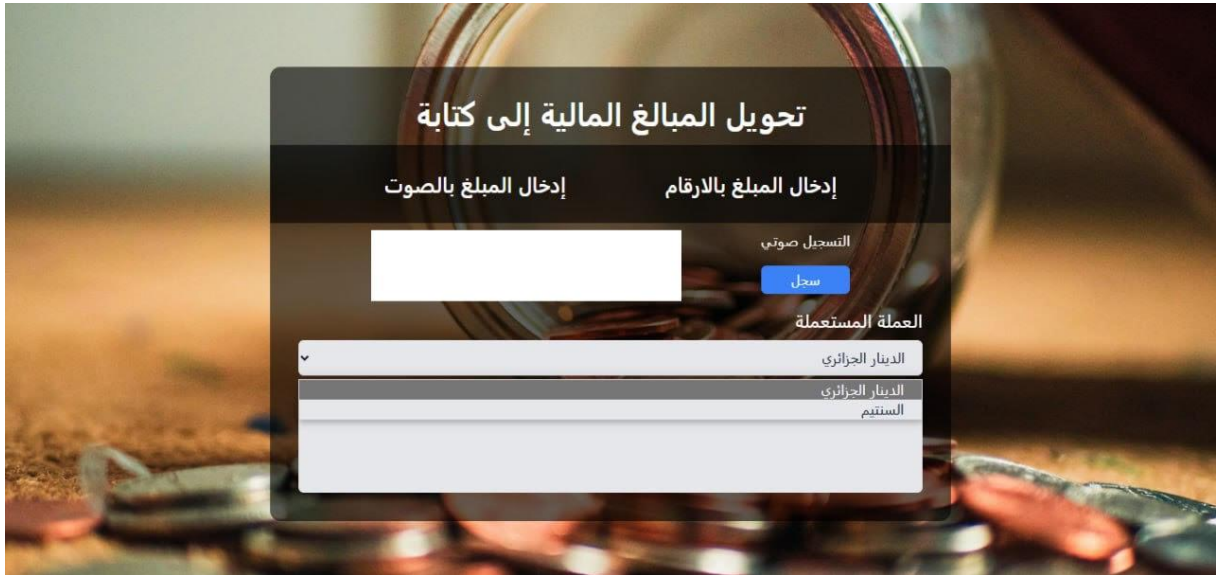


Figure 4.16: choisir la devise

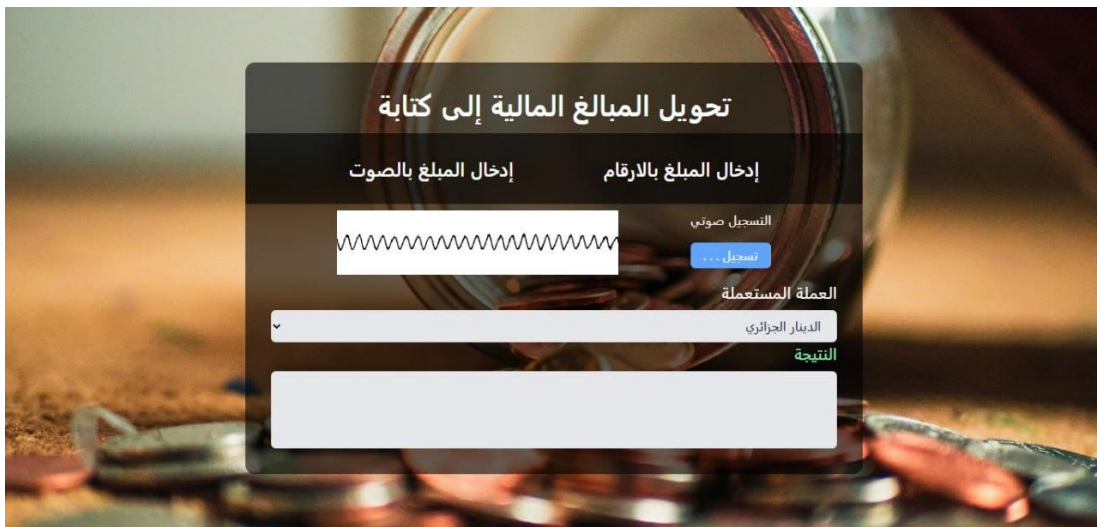


Figure 4.17: Entrer les chiffre par voix 1

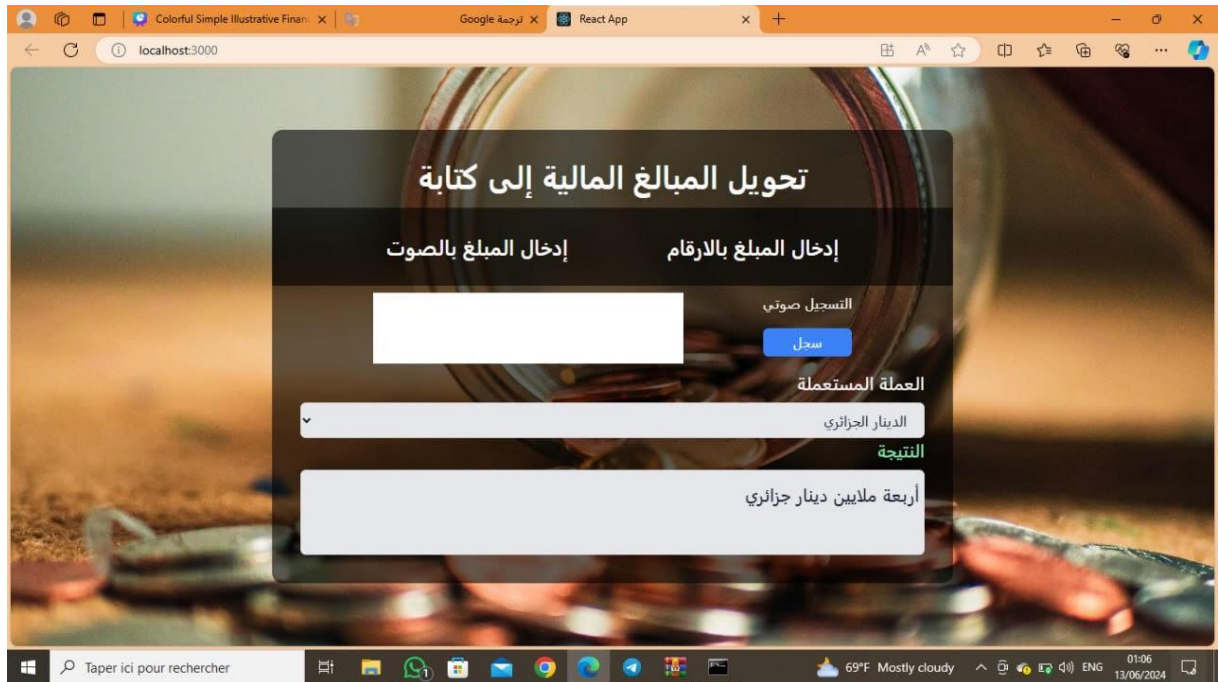


Figure 4.18:Affichage des résultats

9. Conclusion

Cette section a permis de présenter les différents outils utilisés pour la création de notre site web. Nous avons fourni des instructions détaillées sur la façon de configurer un environnement de développement. De plus, nous avons présenté les diverses bibliothèques et les fonctions que nous avons utilisées pour la création de formulaires. Pour terminer ce chapitre, nous avons présenté les principales interfaces graphiques de notre site web pour bien illustrer le travail qui a été réalisé. Nous espérons que cette présentation va contribuer à mieux comprendre les défis du développement d'applications web ,et a donné les clés pour mener à bien les futurs projets.

Conclusion Générale

En conclusion, notre projet de création d'une application web pour la reconnaissance intelligente des sommes monétaires constitue une solution pratique et centralisée pour aider les personnes âgées, les étudiants à écrire leur somme d'une façon correcte et facile. En regroupant toutes les fonctionnalités nécessaires sur une seule application web conviviale, nous avons offert aux utilisateurs un accès simplifié et saisie des chiffres par écrit (lettres/chiffres) ou saisie par la Voix.

Au cours de notre recherche, nous avons abordé les concepts fondamentaux liés au Web et un aperçu sur la poste dans le premier chapitre, ce qui nous a permis de comprendre les bases nécessaires à la création d'une application web solide. Dans le deuxième chapitre nous avons présenté la reconnaissance vocale ; Dans le troisième chapitre, nous avons présenté la conception de notre application en utilisant le langage de modélisation UML, assurant ainsi une planification et une structure efficaces. Enfin, dans le quatrième chapitre, nous avons présenté l'application elle-même, en mettant en évidence les outils de programmation utilisés et en offrant un aperçu des interfaces graphiques développées.

Bibliographie:

- [1] À propos de la Société postale algérienne . <https://www.poste.dz>.
- [2] <https://www.noiise.com/definition/site-web>.
- [3] le mémoire « Design and production of a web site for the management of a kindergarten. » 2021-2022 .
- [4]<https://www.techno-science.net/glossaire-definition/Reconnaissance-vocale.html>.
- [5] IONOS by 11. UML - A Modeling Language for Object-Oriented Programming, n.d.
- [6] diagramme-d 'activité /<https://www.lucidchart.com/pages/fr/diagramme-dactivite-uml>.
- [7] FLASK/ <https://easypartner.fr/blog/flask-vs-django/>.
- [8] [https://testdriven.io/blog/how-are-requests-processed-in- Flask?](https://testdriven.io/blog/how-are-requests-processed-in-Flask?),
- [9]Python Software Foundation. Python. <https://www.python.org/>, N/A. Accessed on : June 13, 2023.
- [10] Mozilla Developer Network. Javascript, n.d .
- [11] <https://react.dev>.
- [12] Microsoft. Visual studio, n.d.
- [13] <https://openai.com> › index › whisper...

الملخص

في إطار العمل الذي قمنا به والذي يتمثل في تطوير نظام التعرف الآلي على المبالغ المكتوبة بحروف اللغة المتداولة (السنتم) وتحويلها الى اللغة العربية مقروءة ومكتوبة وتطبيق هذا النظام على الشيكات البريدية الجزائرية، وذلك باستخدام عدة مصنفات ، بالإضافة إلى إنشاء قاعدة بيانات جديدة لالارقام ، قمنا أولاً بالخطوة الأولى، والتي تتمثل في ادخال الارقام بالسنتم مكتوبة ثم تحويلها الى الدينار الجزائري بصيغة الأرقام والاحرف، اما بالنسبة للخطوة الثانية فتتمثل في ادخال الارقام بالصوت ويكون ذلك بظهور النتيجة على الصيغة الكاملة للمبلغ المالي المراد كتابته في الصك البريدي.

الكلمات المفتاحية: الصك البريدي ، التعرف الآلي ،المبالغ الحرفية المكتوبة، ، ادخال الارقام بالصوت

Le résumé

Dans le cadre des travaux que nous avons entrepris, qui consistent à développer un système de reconnaissance automatique des montants littéraux écrits dans les lettres de la langue commune (centim), à les convertir en arabe lisible et écrit, et à appliquer ce système aux chèques postaux algériens. , en utilisant classificateurs, en plus de créer une nouvelle base de données de nombres. Nous avons d'abord fait la première étape, qui consistait à saisir les nombres en centimes écrits puis à les convertir en dinar algérien sous forme de chiffres et de lettres. Quant à la deuxième étape, elle consiste à saisir les chiffres vocalement, et le résultat apparaîtra sous la forme complète du montant financier à inscrire sur le chèque .

Mots clés : Chèque postal, reconnaissance automatique, montants littéraux écrits, , saisie vocale des chiffres

Abstract

Within the framework of the work that we have undertaken, which consists of developing an automatic recognition system for literal amounts written in the letters of the common language (centim), converting them into readable and written Arabic, and applying this system to Algerian postal checks, using classifiers, in addition to creating a new database for numbers. We first took the first step, which consisted of entering the numbers in written centimes and then converting them to the Algerian dinar in the form of numbers and letters. As for the second step, it is to enter the numbers by voice, and the result will appear in the full form of the financial amount to be written on the check.

Key word: Postal check, automatic recognition, written literal amounts, entering numbers by voice