

CHAPITRE 4

IMPLÉMENTATION ET TEST

Dans ce chapitre nous allons présenter notre environnement de travail (Python,HTML et CSS ,PHP,PostgreSQL,XAMPP,Tomcat), ainsi que la bibliothèque Keras pour l'apprentissage et la classification. on va aussi présenter et expliqué brièvement les différentes interfaces GED et OCR ,et la fin une discussion des résultats de traitement de notre OCR.

4.1 ENVIRONNEMENT DE TRAVAIL :

4.1.1 PYTHON :



FIGURE 4.1. Python.

Python est un langage de programmation interprété, multiparadigme est multiplateforme compatible avec de nombreux systèmes d'exploitation, il favorise la programmation impérative structurée, fonctionnelle et orientée objet.

Python est placé sous une licence libre (Python software foundation license). l'utilisation de python est adaptée dans de nombreux contextes grâce à ses nombreuses bibliothèques optimisées, il est cependant particulièrement utilisé comme langage de script pour automatiser des tâches simples mais fastidieuses.

La version utilisée pour notre projet est python 3.6. [Lutz, 2009]

4.1.2 HTML ET CSS



FIGURE 4.2. HTML et CSS.

HTML : est un langage de balisage conçu pour représenter les pages web, il permet d'écrire de l'hypertexte ainsi que la structuration sémantique et logique pour mettre en forme le contenu des pages. [Gillies and Cailliau, 2000]

CSS : les feuilles en cascade forment un langage informatique qui décrit la présentation des documents HTML et HML, il est utilisé dans la conception des sites web. [Lie, 2005]

La combinaison Html et CSS est un véritable standard en informatique, les deux langages se trouvent à la base de tout projet web car ils ont un rôle qui les rend incontournables, le HTML va donc créer la structure des pages tandis que le CSS va nous permettre de modifier l'apparence des contenus de la page. Concernant l'environnement de travail, on a choisi Brackets comme éditeur de texte car il dispose d'une excellente ergonomie. [Gir, 2004]

4.1.3 PHP



FIGURE 4.3. PHP.

PHP : est un langage de programmation libre impératif orienté objet.

Il est considéré comme une des bases de la création de sites web dits dynamiques mais également des applications web, c'est un langage de script qui est particulièrement adapté au développement web. Rapide et flexible il intègre tous les outils nécessaires à la création de sites dynamiques. Lorsqu'une page PHP est exécuté par le serveur, alors celui-ci renvoie généralement au client (aux visiteurs du site) une page web qui peut contenir du HTML, XHTML, CSS, JavaScript ... [[w3s, 2003](#)]

4.1.4 POSTGRESQL



FIGURE 4.4. PostgreSQL.

PostgreSQL : est un système de gestion de base de donnée relationnelle et objet (SGBDRO). c'est un outil libre créé par Michael Stonebraker.

Ce système multiplateformes est largement connu et réputé à travers le monde, surtout pour être respectueux des normes ANSI SQL, ce projet est géré par une communauté de développeurs.

PostgreSQL fonctionne sous plusieurs systèmes d'exploitation dont Linux, MacOS, Windows (depuis la version 8.0) etc. plusieurs interfaces utilisateurs existent : psql, pgAdmin et PHPAdmin, pour notre projet on a utilisé PgAdmin. [Pos, 2005]

4.1.5 XAMPP



FIGURE 4.5. XAMPP.

Xampp : Est un ensemble de logiciels qui permettent de mettre facilement en œuvre un serveur web local, il offre une bonne souplesse d'utilisation car il est simple et rapide à installer, d'ailleurs il est à la portée d'un grand public car il n'exige pas de connaissance particulière, en plus il fonctionne sur plusieurs systèmes d'exploitation.

En d'autres termes XAMPP est un programme simple et gratuit qui permet aux utilisateurs d'héberger des sites web sur leur pc. [des, 2009]

4.1.6 TOMCAT



FIGURE 4.6. Tomcat.

Tomcat : Est un conteneur web libre de servlets il implémente les spécifications des servlets, utilisé en association avec un autre serveur web en général Apache.

Principe de fonctionnement :

- Le serveur web s'occupe des pages web traditionnelles (.html, .php par exemple).
- Il délègue à Tomcat les pages relevant spécifiquement d'une application web Java (Servlet, JSP...). [[Wiki, 2019](#)]

4.1.7 INTELLIJ IDEA

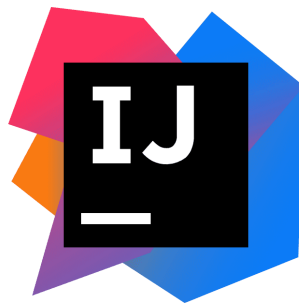


FIGURE 4.7. IntelliJ-IDEA.

IntelliJ IDEA : est un EDI disponible en deux versions, l'une communautaire, open source, sous licence Apache 2 et l'autre propriétaire, protégée par une licence commerciale. En plus de java IntelliJ permet également de supporter plusieurs langages comme HTML, CSS, PHP, Python, JS etc... Il est disponible sous plusieurs Systèmes d'exploitations (Windows, Linux et MacOS) . . [[will, 2015](#)]

4.1.8 KERAS

Keras est une interface de programmation d'application de réseaux de neurones de haut niveau, écrite en Python et capable de s'exécuter sur TensorFlow, CNTK ou Theano. Elle a été développée pour permettre une expérimentation rapide. Pouvoir faire de la recherche de qualité est essentiel pour pouvoir passer de l'idée au résultat le plus rapidement possible. [[Git, 2008](#)]

4.2 INTERFACES

Nous allons présenter les différentes interfaces et fonctionnalités d'Alfresco qui a été personnalisé et modifié afin de correspondre à l'organisme qui va l'utiliser (le centre des archives), ainsi que les

différentes interfaces de notre OCR.

Interface d'authentification :

Au lancement de la GED on aura une interface d'authentification qui exige à l'utilisateur de saisir les informations demandées à savoir le nom d'utilisateur et mot de passe afin de se connecter à la plateforme.[FIGURE 4.8]



FIGURE 4.8. Interface authentification GED.

Interface d'accueil :

Après avoir saisi les bonnes informations, vous aurez accès à la plateforme d'accueil qui se compose de :[FIGURE 4.9]

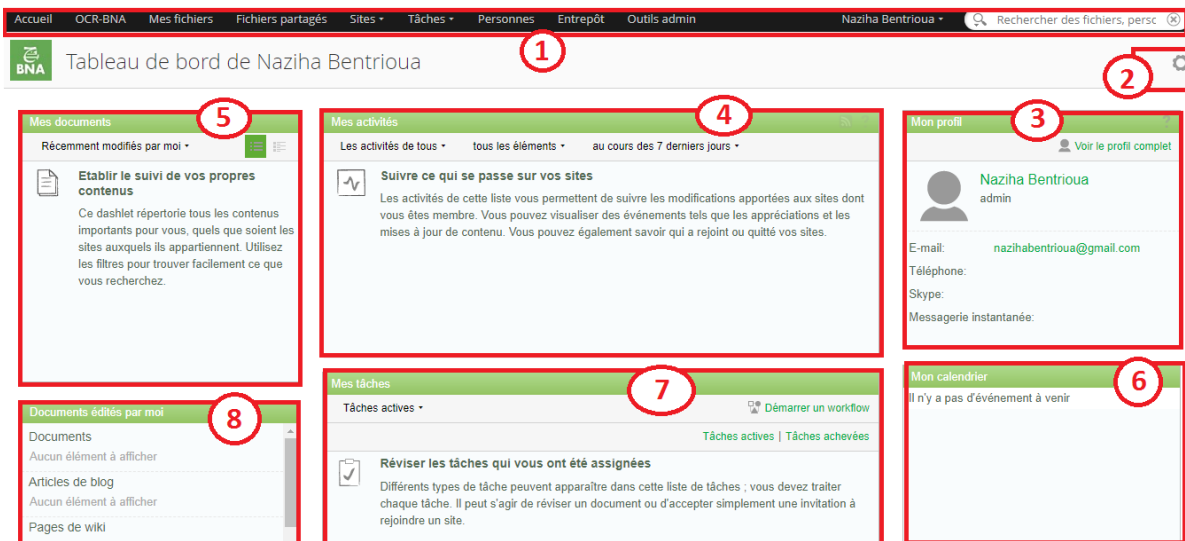


FIGURE 4.9. Interface d'accueil GED.

(1) tableau de bord utilisateur : quelque soit votre emplacement dans la GED vous aurez toujours ce tableau à votre portée pour vous faciliter la navigation.

(2) personnaliser le tableau de bord : vous pouvez modifier votre tableau de bord en changeant la disposition de vos dashlets ou encore les supprimés et ajouter d'autre celons vos besoins.

(3) Mon Profil : permet d'afficher l'identité de l'utilisateur en cours, une possibilité de mettre a jour le profil est mis en place afin d'ajuster des détails ou changer le mot de passe utilisateur.

(4) Mes activités : cette section permet d'afficher les dernières activités (celons une certaine échéance de temps) que l'utilisateur en cours à effectuer.

(5) Mes documents : ici une liste de documents récemment modifiés ou ajoutés par l'utilisateur sera affichée.

(6) Mon calendrier : permet de vous rappeler les prochains événements à venir.

(7) Mes tâches : cette section affichera une liste des tâches qui vous ont été assigné afin de les accomplir, une fois la tâche terminée elle disparaîtra de votre liste.

(8) Documents édités par moi : comme son nom l'indique, cette partie permet d'affiché les documents édités par l'utilisateur à savoir des documents, article de blog etc.

Interface gérer les utilisateurs :

Afin de gérer les utilisateurs qui auront accès à la GED, seul l'administrateur aura la possibilité d'ajouter ou faire des modifications ou suppressions sur les utilisateurs.

Pour accéder à l'interface de la gestion des utilisateurs, on clique sur (1) « outils admin » dans le tableau de bord ensuite sur (2) « utilisateurs », une liste des utilisateurs existents s'affichera [FIGURE 4.10].

The screenshot shows the 'Outils admin' interface. The top navigation bar has 'Outils admin' highlighted with a red circle and the number 1. The left sidebar has 'Utilisateurs' highlighted with a red circle and the number 2. The main area shows a table of users with a 'Nouvel utilisateur' button highlighted with a red circle and the number 3.

	Nom	Nom d'utilisateur	Intitulé du poste	E-mail	Utilisation	Quota
	Alice Beecher	abeecher	Graphic Designer	abeecher@example.com	8 Mo	
	Naziha Bentrioua	admin	admin	nazihabentrioua@gmail.com	0 octets	
	Amina Robaline	Amina		amina.rob@gmail.com	0 octets	
	Guest	guest			0 octets	
	Mike Jackson	mjackson	Web Site Manager	mjackson@example.com	8 Mo	

FIGURE 4.10. Interface gérer user GED.

Pour ajouter un nouvel utilisateur on clique sur (3) et une interface de ce genre [FIGURE 4.11] s'affichera pour remplir les informations nécessaire à la création de l'utilisateur.

Nouvel utilisateur

Informations

Prénom: *

Nom:

E-mail: *

À propos de l'utilisateur

Nom d'utilisateur: *

Mot de passe: *

Vérifier le mot de passe: *

FIGURE 4.11. Ajouter user GED.

Pour faire des modifications et suppressions sur un utilisateur, on clique sur le nom d'utilisateur affiché ou recherché, et on aura accès à la modification ou suppression de la personne, dans la partie modification on peut assigné l'utilisateur à des groupes de travail.[FIGURE 4.12]

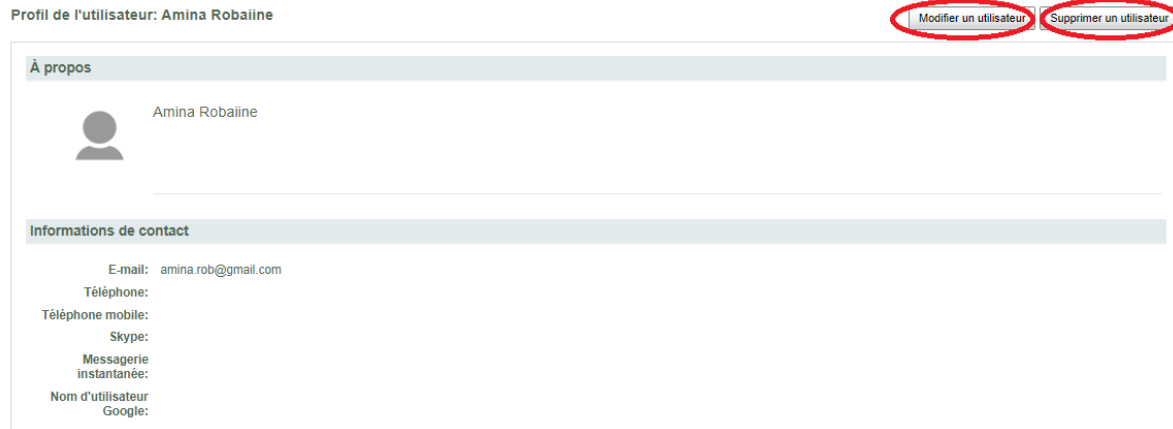


FIGURE 4.12. Modifier supprimer user GED.

Interface gérer les groupes :

Afin de gérer les équipes de travail de notre département (centre des archives), des groupes peuvent être créés afin d'organiser et faciliter la coordination entre groupes.[FIGURE 4.13]

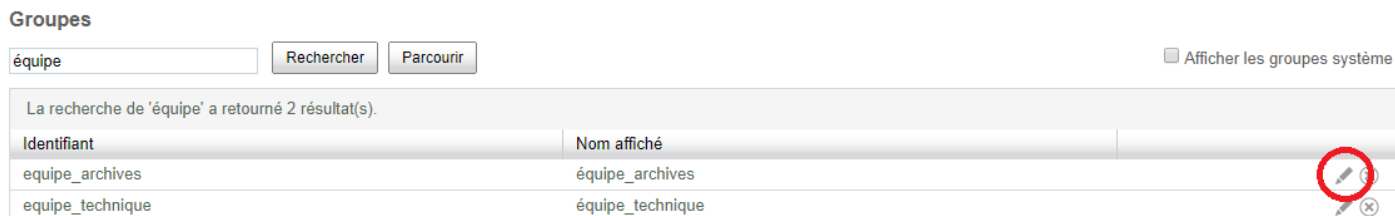


FIGURE 4.13. Gérer groupe GED.

Interface gérer les sites :

Un site est une zone de travail dans laquelle les membres peuvent partager du contenu et travailler ensemble. Le créateur du site devient son gestionnaire et peut confier la gestion du site à d'autres membres.

Créer un site est simple et rapide, dans le tableau de bord on trouve la section « sites » dans laquelle on a la possibilité de créer un site en cliquant sur « créer site ».[FIGURE 4.14]

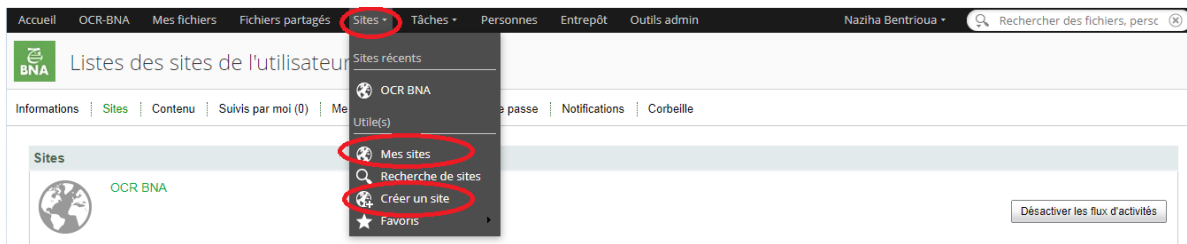


FIGURE 4.14. Section sites GED.

Dans la même section « site », on peut accéder à un site déjà créé on obtient une interface comme ceci [FIGURE 4.15] qui se compose de :

(1) Membres du site. (2) Contenus du site (espace documentaire). (3) Activité du site. (4) Mes discussions : (pour faciliter la communication entre les membres du site).

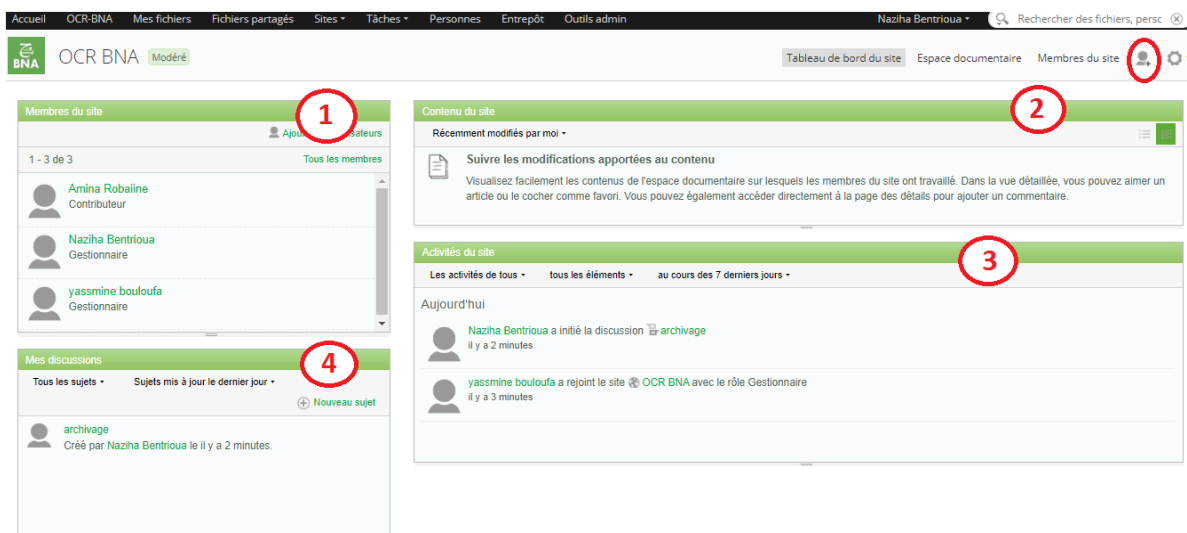


FIGURE 4.15. site GED.

Pour ajouter un utilisateur au site on clique sur l'icône « Ajouter membre », une interface se présentera [FIGURE 4.16] dans la quelle on trouvera 3 étapes à suivre :

(1) sélectionner l'utilisateur à ajouté.

(2) définir le rôle de l'utilisateur sur :

Gestionnaire : dispose des droits d'accès à l'ensemble du contenu du site.

Collaborateur : peut modifier mais ne pas supprimer le contenu du site.

Contributeur : dispose des droits d'accès du contenu dont il est le propriétaire uniquement.

Lecteur : dispose des droits d'accès au site comme lecteur uniquement.

(3) Ajouter l'utilisateur au site.

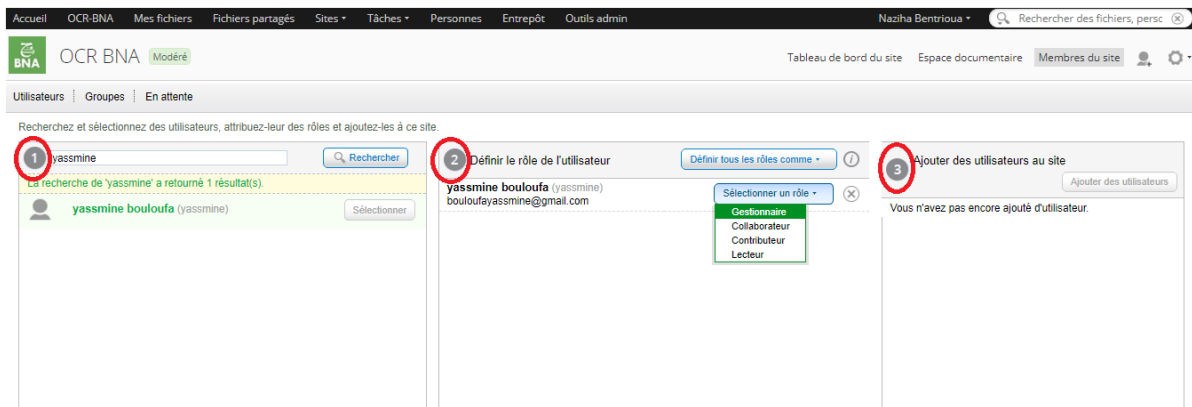


FIGURE 4.16. Ajouter membres au site GED.

Interface gérer les documents :

L'espace documentaire donne une vue global des différents documents ajoutés au site, plusieurs options sont mis à disposition y compris la modification et suppression tel que [FIGURE 4.17] :



FIGURE 4.17. Option espace documentaire GED.

(1) Gérer les droits d'accès : cette fonction prévaut sur les rôles assignés par défaut, un membre peut avoir plus d'accès ou moins sur un document par rapport aux autres contenus documentaires.

(2) Éditer en hors-ligne : Elle permet de télécharger un document sur votre ordinateur et le verrouiller dans l'espace documentaire afin d'empêcher d'autre utilisateur de le modifier en parallèle.

(3) Recherche par filtrage : on peut filtrer notre recherche et cela dépend des besoins recherché ex. (tous les documents, modifié par moi, modifié par d'autre, récemment modifié, récemment ajouté, mes favoris).

(4) démarrer un workflow : aide à organiser les tâches que les utilisateurs doivent accomplir, on peut assigner chaque tâche à une personne avec un document précis sur une échéance de temps à respecter.[FIGURE 4.18]

Accueil OCR-BNA Mes fichiers Fichiers partagés Sites ▼ Tâches ▼ Personnes Entrepôt Outils admin

BNA Démarrer un workflow

Workflow: **Nouvelle tâche ▼**

* Champs requis

Général

Message: ?

Échéance: J J/MM/AAAA

Priorité: Moyenne ▼

Personne assignée

Assigner à: *

Sélectionner

Éléments

Élément(s):

Pas d'élément sélectionné

Ajouter Tout enlever

Autres options

☒ Envoyer des notifications par e-mail

Démarrer un workflow Annuler

FIGURE 4.18. Démarrer Workflow GED.

Interface OCR :

Concernant notre OCR la 1ère interface [FIGURE 4.19] propose de choisir un type de document sachant que notre OCR s'applique sur les chèques bancaires et les bordereaux de versements AADL de la BNA :

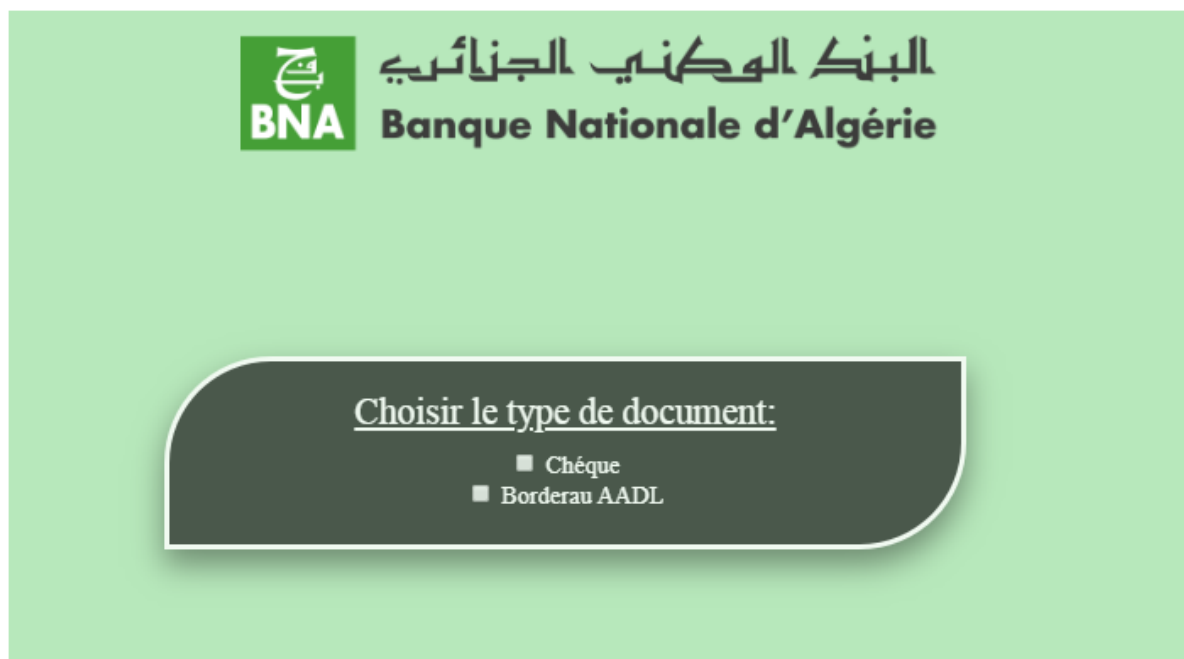


FIGURE 4.19. Interface accueil OCR.

une fois le type "chèque" sélectionné on obtient 3 catégories : [FIGURE 4.20]

(1) Insertion manuelle.

(2) Appliquer OCR.

(3) Recherche.



FIGURE 4.20. Interface chèque.

En cliquant sur L'insertion manuelle on aura une interface comme ceci, elle se compose d'un formulaire dont chaque case correspond à une donnée essentielle à l'archivage du chèque bancaire du client . Cette catégorie nécessite une insertion manuelle des données afin de les ajoutées à la base de données en cliquant sur le bouton « valider ».[FIGURE 4.21]



The image shows a screenshot of a web application interface for the Banque Nationale d'Algérie (BNA). At the top, the BNA logo and name are displayed in Arabic and French. Below this, a dark green rounded rectangle contains the title 'Insertion manuelle'. Inside this rectangle, there is a form with six input fields, each preceded by a small icon: a person icon for 'Nom et prenom client', a person icon for 'N° de compte', a stack of coins icon for 'Montant DA', a group of people icon for 'a l'ordre de Mr/Mme', a document icon for 'N° Serie de cheque', and a calendar icon for 'jj/mm/aaaa'. At the bottom left of the form is a green button labeled 'Valider'.

FIGURE 4.21. Saisie manuelle.

La 2ème catégorie représente l'application de l'OCR, son utilisation est simple il suffit de choisir l'image du chèque scanné, un traitement de quelques secondes sera effectué le temps d'appliquer des filtres, de segmenter et reconnaître les caractères des zones qui nous intéressent afin d'envoyer chaque information dans sa case adéquate du formulaire, la possibilité de modifier le résultat obtenu est possible pour être ajouté à la base de données.[FIGURE 4.22]

البنك الوطني الجزائري
BNA Banque Nationale d'Algérie

Choisir un fichier cheque1555gg.PNG OK

OCR

GUETTAF KHEIR

1020001xxxx

23 000 v0 |

moi meme

968xxx

16/09/2019

Valider

FIGURE 4.22. OCR.

Et finalement la 3ème catégorie représente la recherche, [FIGURE 4.23], l'utilisateur a la possibilité de faire une recherche bien précise, il suffit d'introduire les informations qu'on veut rechercher (par exemple une recherche par N° de compte ou la date). Et là on aura le résultat de la recherche sous forme de tableau.

Resultat de la recherche

Nom et Prenom	N° de compte	Montant	Ordre	N° de cheque	Date
OULD HAMOUDA WASSILA	1020002 XXXX	15800 DA	moi meme	143 XXXX	2019-09-12
GUETTAF KHEIR	1020001 XXXX	1500 DA	moi meme	968 XXXX	2019-09-12

FIGURE 4.23. Recherche par date.


Concernant le 2ème type de document "Bordereau AADL", il suffit de cliquer sur ce type pour avoir l'interface suivante [FIGURE 4.24], une fois le type "Bordereau AADL" sélectionné on obtient 3 catégories :

- (1) Insertion manuelle.
- (2) Appliquer OCR.
- (3) Recherche.



FIGURE 4.24. Interface bordereau AADL.

En cliquant sur L'insertion manuelle on aura une interface comme ceci, elle se compose d'un formulaire dont chaque case correspond à une donnée essentielle à l'archivage du bordereau d'AADL du client . Cette catégorie nécessite une insertion manuelle des données afin de les ajoutées à la base de données en cliquant sur le bouton « valider ».[FIGURE 4.25]



The screenshot displays the 'Insertion manuelle' (Manual Insertion) interface of the Banque Nationale d'Algérie (BNA) application. The interface is set against a light green background with the BNA logo and name in Arabic and French at the top. The main form is a dark grey rounded rectangle containing four input fields, each with a corresponding icon on the left: a person icon for 'Nom et prenom client', a group of people icon for 'N° de compte AADL', a document icon for 'N° de bordereau', and a calendar icon for a date field labeled 'jj/mm/aaaa'. A green 'Valider' button is positioned at the bottom center of the form.

FIGURE 4.25. Saisie manuelle bordereau.

La 2ème catégorie représente l'application de l'OCR, son utilisation est simple il suffit de choisir l'image du bordereau scanné, un traitement de quelques secondes sera effectué (le temps d'appliquer des filtres, de segmenter et reconnaître les caractères des zones qui nous intéressent afin d'envoyer chaque information dans sa case adéquate du formulaire), la possibilité de modifier le résultat obtenu est possible pour être ajouté à la base de données. [FIGURE 4.26]



FIGURE 4.26. OCR bordereau.

Et finalement la 3ème catégorie représente la recherche, [FIGURE 4.27], l'utilisateur a la possibilité de faire une recherche bien précise il suffit d'introduire les informations qu'on veut rechercher (par exemple une recherche par date). Et là on aura le résultat de la recherche sous forme de tableau.

Recherche

Nom et prenom client

N° de compte AADL

N° de bordereau

08/10/2019

Valider

Resultat de la recherche

Nom et Prenom	N° de compte AADL	N° de bordereau	Date
mane. Hertor	0300053076-16	318391	2019-10-08
ABDELLE AICHA	0300053076-16	1257942	2019-10-08

FIGURE 4.27. Recherche bordereau.

L'ensemble des informations traitées précédemment sont stockées dans une base de données afin de garantir la sauvegarde des données pour pouvoir les restaurées dans le besoin d'une recherche par exemple.[FIGURE 4.28] Ils sont stockés sous PostgreSQL.

Data Output	Explain	Messages	Notifications
id integer	nom character varying (300)	montant integer	ordre character varying (300)
7	GUETTAF KHEIR	23000	Moi même
8	I HAMAD ATEF	156200	Moi même
9	MERAH AHME	6000	Moi même
10	ROBAINE AMINA	8546000	Moi même
11	BOULANDJAS LIAS	50000	moi meme
12	BOULANDJAS LIAS	5000	moi meme
13	GUETTAF KHEIR	23000	moi meme
14	OULD HAMOUDA WASSILA	15400	moi meme

FIGURE 4.28. Base de donnée.

4.3 IMPLÉMENTATION DE L'OCR

Avant de procéder à l'implémentation de notre OCR on devait implémenter d'abord CNN, et pour cela on a utilisé Keras, qui est une bibliothèque pour l'apprentissage en profondeur, elle prend en charge les réseaux convolutionnels et les réseaux récurrents, ainsi que la combinaison des deux.

Le bout de code suivant représente le type du modèle ainsi que les couches qu'on a utilisé pour la construction de notre CNN :[FIGURE4.29]

```
def define_model():
    model = Sequential()
    model.add(Conv2D(28, (3, 3), activation='relu', input_shape=(28, 28, 1)))
    model.add(MaxPooling2D((2, 2)))
    model.add(Conv2D(56, (3, 3), activation='relu'))
    model.add(MaxPooling2D((2, 2)))
    model.add(Flatten())
    model.add(Dense(56, activation='relu', kernel_initializer='he_uniform'))
    model.add(Dense(123, activation='softmax'))
    # compile model
    model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['acc', f1_m, precision_m, recall_m])
    return model
```

FIGURE 4.29. Création d'un modèle.

Dans Keras on assemble des couches pour construire un modèle séquentiel afin de construire un réseau simple et entièrement connecté. [FIGURE3.16] Tel que 784 représente le nombre de pixels d'une

image en entrée (28*28) et 62 représente nos classes de sortie (1,2,3...a,A,b,B...).

Les couches utilisées dans le modèle sont définies comme ceci :

keras.layers.Conv2D :

cette couche crée un noyau de convolution qui est convoluté avec l'entrée de la couche pour produire un tenseur de sorties. Cette couche consiste à filtrer l'image avec un filtre de pixel plus fin et réduire la taille sans perdre la relation entre les pixels, `kernel_size()` permet de paramétrer la hauteur et la longueur de la fenêtre de convolution 2D. Lorsqu'on l'utilise en 1ère couche dans un modèle on doit fournir le `input_shape` qui est paramétré à (28,28,1).

keras.layers.MaxPooling2D :

c'est la 2ème couche de notre modèle elle permet de réduire la taille spatiale et le nombre de paramètres. nous avons sélectionné une taille de regroupement `pool_size()` qui permet de réduire verticalement et horizontalement les dimensions spatiales en sélectionnant les valeurs maximales.

Keras.Layers.Flatten :

Les couches de convolution et de regroupement ont des tenseurs multidimensionnels comme sortie. Afin d'utiliser la couche Dense, on doit transformer ces tenseurs en un tenseur 1D en utilisant `Flatten()`.

keras.layers.Dense :

Dense implémente l'opération, 128 représente l'unité de dimension pour l'espace de sortie. L'activation est une fonction qui est appliquée à la sortie d'une couche de réseau neuronal, qui sera transmise comme entrée à la couche suivante, les fonctions d'activation sont une partie essentielles des réseaux de neurone car sans elles le réseau neuronal se réduit à un simple modèle de régression logistique.

Relu :

Un élément important dans l'ensemble du processus est l'Unité linéaire rectifiée ou ReLU. Les mathématiques derrière ce concept sont assez simples, chaque fois qu'il y a une valeur négative dans un pixel, on la remplace par un 0. Ainsi, on permet au CNN de rester en bonne santé (mathématiquement parlant) en empêchant les valeurs apprises de rester coincées autour de 0 ou d'exploser vers l'infini. [Charle, 2017]

Le résultat d'une couche ReLU est de la même taille que ce qui lui est passé en entrée, avec simplement toutes les valeurs négatives éliminées.

Softmax :

Softmax est bien pour la classification, elle calcule la distribution des probabilités de chaque classe cible parmi toutes les classes possibles.

Par exemple, si on donne en entrée la couleur des pixels d'une image de chat, on la multiplie par une matrice de poids qui lui correspond, afin de la transformer en un vecteur de T élément(logits) et ainsi chaque logit sera le score d'un animal. Si le score du chat est le plus important, alors la probabilité donnée par la fonction softmax que l'image est un chat sera la plus importante, d'après l'étude de la couleur des pixels. Mais on peut travailler sur d'autres caractéristiques, et ainsi obtenir d'autres probabilités, afin de déterminer l'animal sur la photo. Au fur et à mesure que l'intelligence artificielle aura d'exemples, plus la matrice de poids s'affinera, et plus le système sera performant : on parle d'apprentissage automatique. [[SAIMADHU, 2017](#)]

Après avoir ajouté les couches précédentes à notre modèle séquentiel, vient la phase de compilation ou on a utilisé :

`optimizer=tf.keras.optimizers.Adam()` :

l'optimisation Adam est une méthode de descente de gradient stochastique basée sur une estimation adaptative des moments du 1er et second ordre.

La méthode est efficace en termes de calcul, requiert peu de mémoire et convient bien aux problèmes de donnée/paramètres.

`loss=tf.keras.losses.categorical_crossentropy()` :

cette fonction est utilisée lorsqu'il existe au moins 2 classes d'étiquettes, les étiquettes sont fournis sous formes d'entrée.

`metrics=['accuracy']` :

précision de classe elle calcule la fréquence à laquelle les prédictions correspondent aux étiquettes.

Lors de l'implémentation de notre OCR on a passé par les phases suivantes :

4.3.1 PRÉ-TRAITEMENT :

Cette étape est principalement effectuée pour garantir à notre système la faciliter à identifier les caractères d'une image, une large gamme de fonctions de pré-traitement est mis en œuvre en fonction de nos besoin. Parmi ces fonctions on a utilisé :

```
ref = cv2.cvtColor(ref, cv2.COLOR_BGR2GRAY)
```

FIGURE 4.30. Conversion en niveau de gris.

Cv2.cvtColor [FIGURE4.30] cette fonction permet de convertir l'image en couleur à une image au niveau de gris.

```
ref = cv2.threshold(ref, 0, 255, cv2.THRESH_BINARY_INV | cv2.THRESH_OTSU)[1]
```

FIGURE 4.31. Conversion en binaire.

Cv2.threshold [FIGURE4.31]cette fonction permet de convertir notre image en niveau de gris en une image binaire(les valeurs des pixels prennent seulement la valeur noir ou blanc).

4.3.2 SEGMENTATION

Concernant la segmentation on a travaillé avec une méthode basée sur les ROI.

Dans un premier temps on a effectué un découpage en zone contenant les informations à extraire de notre document, car il pourra y avoir des régions dans l'image qui ne contiennent aucune information.

Pour extraire les contours de la zone nous utilisant Cv2.findcontours qui produira une liste de contours désordonnée, avant de bouclé sur cette liste de contours on doit d'abord les trié de gauche à droite. Dans cette boucle on va calculer le rectangle de délimitation de chaque contour de cette liste avec la fonction Cv2.boundingrect , si la hauteur et la largeur du caractère sont supérieures ou égales à la hauteur et à la largeur minimales, on va effectuer une extraction du ROI de l'image à l'aide des coordonnée (x,y) et de la largeur/hauteur de notre rectangle englobant, puis on le redimensionne selon la taille des images de notre dataset pour le faire passer par notre modèle entraîné afin prédire sa classe

d'appartenance.

A la fin on effectue une concaténation entre les caractères prédits pour reconstituer le mot.

4.3.3 POST-TRAITEMENT

Cette phase sert à évoluer les taux de reconnaissance des caractères en effectuant une correction manuelle sur les caractères mal prédits, puis ajouter l'image de ces caractères à notre dataset pour l'entraîner sur ces images pour une meilleure reconnaissance.

4.4 TEST ET RÉSULTAT

MNIST est une base de données qui contient des chiffres manuscrits (0 à 9), elle est considérée comme un sous-ensemble d'un ensemble de données plus vaste (EMNIST). [Orhan, 2018]

Elle est le "Hello World" de l'apprentissage automatique sur qui on peut former un algorithme pour tester une nouvelle technologie ou un nouveau modèle afin de s'assurer de leurs bons fonctionnements.

MNIST est divisée en deux jeux de données :

Le jeu d'apprentissage (trainX) qui contient 60000 exemples de chiffres écrits à la main.

Le jeu de test (testX) qui contient 10000 images de chiffres sur lesquelles on effectue les tests de bon fonctionnement de nos modèles.

Ces images ont la même taille et à l'intérieure les chiffres sont centrés et leur taille est normalisée.

Elle peut être utilisée dans l'apprentissage supervisé pour la classification car elle associe à chaque image de chiffre la classe qui lui appartient.

Pour notre cas, on a opté pour MNIST munie d'une extension pour les caractères alphabétiques afin d'enrichir nos données.

Interprétation des mesures de performance :

Une fois que nous avons construit notre modèle, la question la plus importante qui se pose est de savoir quelle est la qualité de ce modèle. L'évaluation de notre modèle est donc la tâche la plus

importante du projet, qui définit la qualité de nos prédictions.[FIGURE 4.32]

Actual Class	Predicted class		
		Class = Yes	Class = No
	Class = Yes	True Positive	False Negative
	Class = No	False Positive	True Negative

FIGURE 4.32. Mesures de performance. [Renuka, 2017]

Les vrais positifs et les vrais négatifs sont les observations correctement prédites et donc affichées en vert. Nous voulons minimiser les faux positifs et les faux négatifs afin qu'ils apparaissent en rouge. Ces termes sont un peu déroutants. Alors prenons chaque terme un par un et comprenons-le bien. [Renuka, 2017]

True Positives (TP) -

Ce sont les valeurs positives correctement prédites, ce qui signifie que la valeur de la classe réelle est égale à la valeur de la classe prédite .

True Negatives (TN) -

Ce sont les valeurs négatives correctement prédites, ce qui signifie que la valeur de la classe réelle est égale à la valeur de la classe prédite.

Faux positifs et faux négatifs, ces valeurs se produisent lorsque votre classe actuelle est en contradiction avec la classe prédite.

Faux positifs (FP) -

Quand la classe réelle est non et la classe prédite est oui.

Faux Négatifs (FN) -

Lorsque la classe réelle est oui mais que la classe prédite est non.

Une fois ces quatre paramètres effectués, on peut calculer les scores de précision, Accuracy, Recall et F1 score.

Accuracy-

est la mesure de performance la plus intuitive. On peut penser que, si nous avons une grande précision, notre modèle est meilleur. Oui, la précision est une excellente mesure, mais uniquement lorsque nous disposons de jeux de données symétriques dans lesquels les valeurs de faux positif et de faux négatif sont presque identiques. Pour notre modèle, nous avons 0,752 , ce qui signifie que notre modèle est d'environ . 75% précis.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

Recall -

est le rapport entre les observations positives correctement prédites et toutes les observations de la classe réelle. Nous avons un Recall de 0,728, ce qui est bon pour ce modèle car il est supérieur à 0,5.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Score F1 -

Le score F1 est la moyenne pondérée de la précision et du recall. Par conséquent, ce score prend en compte à la fois les faux positifs et les faux négatifs. Dans notre cas, le score F1 est 0.7303.

$$\text{F1 Score} = \frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$$

1- Entraînement du modèle sur MNIST : Dans un premier temps, nous avons entraîné notre modèle sur la base de données MNIST simple (caractères numériques) pour voir les performances de notre architecture neuronale. On a effectué une étude paramétrique par rapport au nombre d'époques afin de voir au bout de quelle époque notre modèle entraîné sera stable et ainsi pour suivre l'évolution des valeurs des mesures par rapport aux nombres d'époques.

On a évalué notre modèle en utilisant l'algorithme de Adam comme optimiseur.
Les résultats obtenus sont résumées dans le tableau [TABLE 4.1]

Epoques	Accuarcy	Recall	F1-score
1000	75.6	1.035	0.9575
2000	80.54	1.030	0.9586
4000	89.681	1.015	0.96043
6000	98.69	1.006	0.96152
8000	98.698	1.006	0.96159
10000	98.7	1.006	0.9616

TABLE 4.1 – Résultat sur MNIST.

2- Entraînement du modèle sur Extended MNIST :

De la même manière ,on a entraîné notre modèle sur MNIST+caractères alphabétiques, on a effectué une étude paramétrique par rapport au nombre d'époques et on a obtenu les résultats présentés dans le tableau [TABLE 4.2]

Epoques	Accuarcy	Recall	F1-score
1000	69.95	0.7503	0.7445
2000	70.23	0.7452	0.7401
4000	71.98	0.7412	0.7390
6000	73.63	0.7308	0.7310
8000	75.21	0.7289	0.7304
10000	75.22	0.7288	0.7303

TABLE 4.2 – Résultat sur Extended MNIST.

On constate ici que l'accuracy se stagne aux alentours de 75 % Par contre , on voit ici qu'il y a une légère baisse des performances du modèle. Ceci est dû à la taille du nouveau dataset . La proportion des caractères alphabétique est inférieure à celle des caractères numériques. Un dataset plus complet pourra remédier à cela.

4.4.1 CROSS VALIDATION :

La validation croisée est une méthode statistique utilisée pour estimer l'habileté des modèles d'apprentissage automatique.

La procédure a un paramètre unique appelé k qui fait référence au nombre de groupes dans lesquels un échantillon de données donné doit être divisé. En tant que telle, la procédure est souvent appelée validation croisée des k -fold. Lorsqu'une valeur spécifique pour k est choisie, elle peut être utilisée à la place de k dans la référence au modèle, telle que $k = 10$ devenant une validation croisée de 10 fois. [Jason, 2018]

La validation croisée est principalement utilisée dans l'apprentissage automatique appliqué pour estimer les compétences d'un modèle d'apprentissage automatique sur des données invisibles. C'est-à-

dire d'utiliser un échantillon limité afin d'estimer comment le modèle devrait fonctionner en général lorsqu'il est utilisé pour faire des prédictions sur des données non utilisées pendant la formation du modèle.

La procédure générale est la suivante :

- 1-Mélangez le jeu de données de manière aléatoire.
- 2-Diviser le jeu de données en k groupes.
- 3-Pour chaque groupe unique :
 - Prendre le groupe comme un ensemble de données à retenir ou à tester.
 - Prendre les groupes restants comme un ensemble de données d'entraînement.
 - Ajuster un modèle sur le kit d'apprentissage et l'évaluer sur le kit de test.
 - Conserver le score d'évaluation et jeter le modèle.
- 4-Résumer les compétences du modèle à l'aide de l'échantillon de scores d'évaluation du modèle.

Il est important de noter que chaque observation de l'échantillon de données est affectée à un groupe individuel et reste dans ce groupe pendant la durée de la procédure. Cela signifie que chaque échantillon a la possibilité d'être utilisé dans le délai de conservation 1 fois et d'entraîner le modèle k-1 fois. [Jason, 2018]

En fixant la valeur de l'époque à 8000 (c'est l'époque où l'accuracy s'est stabilisée dans l'entraînement) et en variant la valeur de k de 5 à 10, on a obtenu les résultats suivants : [TABLE 4.3]

Valeur de K	Précision
5	74.521
6	75.21
7	74.845
8	75.112
9	75.157
10	75.013

TABLE 4.3 – Résultat du Cross Validation.

Discussions du résultat :

Après une évaluation répétée de notre modèle sur des valeurs différentes de K, on a aboutit à la valeur K=6 dans la [TABLE 4.3] qui était la plus optimal et la plus efficace pour notre cas, car elle nous a offert une meilleure précision. Cette valeur n'est jamais fixe, car elle peut changé selon le modèle de réseau de neurones utilisé.

4.5 CONCLUSION :

Nous avons présenté dans ce chapitre les outils de travail qui nous ont permis de réaliser notre application.

En premier nous avons présentés les outils et les langage utilisés, ensuite nous avons fait la présentation et explication de nos interfaces de travail concernant la GED et l'OCR, et pour finir on a parler des test, et résultats obtenus par notre OCR.