



► Interface OCR avec Azure

Reconnaissance optique de caractères

L. Berville ►

GRETA – Simplon ► 07/04/2024

Résumé :

Le projet vise à examiner les exigences d'une entreprise cliente pour une nouvelle application intégrant un service d'intelligence artificielle. Cela impliquait de comprendre les besoins métier, les défis à surmonter et les attentes des utilisateurs. L'objectif était de définir avec précision les objectifs de développement en tenant compte des besoins identifiés et des contraintes techniques, tout en garantissant la faisabilité et la viabilité du projet. Suite à cette analyse, le client exprime le besoin d'améliorer sa procédure de pré-traitement des factures. Il souhaite élargir les fonctionnalités de pré-traitement en y ajoutant la reconnaissance optique de caractères (OCR), permettant ainsi d'automatiser le reporting de sa comptabilité fournisseurs.

Pour répondre à ces exigences, les interfaces de l'application ont été développées en utilisant des outils et langages de programmation appropriés, tout en respectant les spécifications fonctionnelles, techniques et de gestion des données établies. Les détails des spécifications sont inclus dans ce document, décrivant en détail les fonctionnalités requises, les interactions avec les utilisateurs et les résultats attendus.

Points à améliorer :

- Application web – Tableau de bord
- Traitement des données OCR pour les accents
- Permettre aux prix d'être supérieur à 99 €

Objectif du projet

Afin de rester compétitives sur le marché et de répondre aux demandes croissantes en termes d'efficacité et de précision, les entreprises peuvent automatiser davantage leurs processus de gestion de documents en utilisant les techniques basées sur le Machine Learning. Cette automatisation permet non seulement de réduire les coûts associés à la saisie manuelle des données, mais aussi d'améliorer la qualité des informations extraites. En conséquence, les entreprises peuvent réaliser des gains d'efficacité et de productivité significatifs, libérant ainsi des ressources pour se concentrer sur des tâches à plus forte valeur ajoutée et accélérer leur prise de décision.¹

À la lumière de ces constatations, notre client a formulé le besoin d'optimiser sa méthode de pré-traitement des factures fournisseurs. Dans cet objectif, ils aspirent à élargir les fonctionnalités du processus de pré-traitement en y intégrant la reconnaissance optique de caractères, afin de faciliter l'automatisation du reporting de leur comptabilité fournisseurs. Afin de concrétiser cette mutation, ils ont mis à notre disposition les données nécessaires de leur processus de traitement, accessibles via une API dédiée permettant d'accéder aux factures numérisées².

Interface OCR avec Azure

Reconnaissance optique de caractères

Les techniques de « Reconnaissance Optique de Caractères » basées sur l'apprentissage automatique permettent l'extraction de texte imprimé à partir d'images telles que des factures. Le texte est généralement extrait sous forme de mots, de lignes de texte, de paragraphes ou de blocs de texte, facilitant ainsi l'accès à une version numérique du texte numérisé. Ce processus élimine ou diminue significativement le besoin de saisie manuelle des données, ce

qui améliore l'efficacité et la précision des tâches de traitement de données.

Avantages

Grâce à l'utilisation de techniques de machine learning, les systèmes OCR peuvent s'adapter et s'améliorer avec le temps. Cette capacité d'adaptation continue permet aux systèmes OCR de rester pertinents et efficaces face à l'évolution des besoins et des exigences, offrant ainsi des performances améliorées à mesure que de nouvelles données sont traitées.

Qu'est-ce qu'Azure AI Vision ?

Une solution de pointe dans ce secteur est Azure AI Vision, une offre de services fournie par Microsoft Azure, qui met à disposition des outils avancés pour le traitement d'images et la reconnaissance visuelle.

Les deux versions de Read disponibles dans Azure AI Vision prennent en charge plusieurs langues, y compris le français. La compétence OCR est utilisée pour extraire le texte à partir de différents types de fichiers image tels que JPEG, JPG, PNG, BMP et TIFF. L'étape initiale dans le processus de traitement est appelée le craquage de document. Cette étape est importante car elle permet de séparer le texte et le contenu de l'image. En d'autres termes, elle identifie les parties de l'image qui contiennent du texte et les isole pour permettre une reconnaissance plus précise. Cette séparation est essentielle pour garantir que le texte soit correctement interprété, indépendamment de la mise en page ou de la complexité de l'image.

En résumé, les fonctionnalités d'OCR dans Azure AI Vision permettent de traiter efficacement les fichiers image pour extraire le texte, et le craquage de document est la première étape de ce processus, assurant une reconnaissance précise du texte³.

¹ <https://portal.vision.cognitive.azure.com/gallery/featured>

² <https://invoiceocrp3.azurewebsites.net/invoices>

³ <https://learn.microsoft.com/fr-fr/azure/search/cognitive-search-skill-ocr>

Comment utiliser Azure AI Vision ?

Pour exploiter les fonctionnalités d'Azure AI Vision, nous avons la possibilité d'utiliser les API fournies par Microsoft Azure. Ces interfaces de programmation sont spécifiquement conçues pour être facilement intégrées dans des applications, permettant ainsi de bénéficier des capacités de vision par ordinateur d'Azure sans nécessiter le développement de modèles à partir de zéro.

Image Analysis 4.0

Configuration des Variables d'Environnement pour les Clés et les Points de Terminaison.

Pour configurer les variables d'environnement nécessaires à l'accès aux services cognitifs de Vision Azure, il est indispensable de définir la clé d'accès et le point de terminaison. Ces informations permettront à notre application de s'authentifier et de communiquer efficacement avec la ressource Azure associée, facilitant ainsi l'intégration des fonctionnalités de Vision Azure dans notre application.

1. Utilisez la commande⁴ pour définir la variable d'environnement VISION_KEY. Remplacez "your-key" par l'une des clés fournies pour votre ressource Azure.
2. Utilisez la commande⁵ pour définir la variable d'environnement. Remplacez "your-endpoint" par le point de terminaison fourni pour votre ressource Azure. Le point de terminaison indique où se trouve le service que vous souhaitez utiliser.

Analyser une image

Exécutez cette commande pour installer le SDK d'analyse d'images : « pip install azure-ai-vision-imageanalysis ». Puis, nous utiliserons le code python donné à cette adresse : [https://learn.microsoft.com/en-us/azure/ai-](https://learn.microsoft.com/en-us/azure/ai-services/computer-vision/quickstarts-sdk/image-analysis-client-library-40?tabs=visual-studio%2Cwindows&pivots=programming-language-python)

[services/computer-vision/quickstarts-sdk/image-analysis-client-library-40?tabs=visual-studio%2Cwindows&pivots=programming-language-python](https://learn.microsoft.com/en-us/azure/ai-services/computer-vision/quickstarts-sdk/image-analysis-client-library-40?tabs=visual-studio%2Cwindows&pivots=programming-language-python)

Traitement des données

Les données renvoyées par l'OCR d'Azure Vision sont sous la forme de Bounding Boxes (cadres de délimitation). Ces bounding boxes sont des rectangles qui encadrent chaque zone de texte détectée dans l'image. En fournissant ces cadres de délimitation, l'OCR permet une localisation précise des mots ou des lignes de texte dans l'image. Les coordonnées de ces bounding boxes, souvent exprimées en pixels, sont également fournies pour chaque zone de texte détectée. Elles permettent de spécifier la position relative du texte dans l'image, ce qui facilite sa localisation et sa manipulation ultérieure. Les factures ayant toutes la même structure, il a été possible de délimiter les zones d'intérêt.



INVOICE FAC_2019_0004	
Issue date 2019-01-02 07:27:00	
Bill to Thibault Martinez de la Boucher	
Address Stedla Romolo, 8 Appartamento 62	
35018, San Martino Di Lupat (PD)	
Déchirer visite mieux âme.	9 x 61.92 Euro
Creuser refuser digne inquiéter.	4 x 53.72 Euro
Well stock operation hope.	6 x 92.54 Euro
Floor international animal can.	6 x 40.44 Euro
Hors comment arracher renoncer.	1 x 55.50 Euro
Suivre parole banc endroit.	6 x 22.64 Euro
Unde repellendus velit velit.	6 x 93.18 Euro
Unde delectus dolore suscipit.	6 x 28.01 Euro
Ordre nez joindre compter.	8 x 45.67 Euro
TOTAL	2853.88 Euro

Figure 1 : Exemple de facture

Base de données SQL

Le but de la base de données est de stocker et d'organiser de manière structurée les informations de manière à ce qu'elles puissent être facilement accessibles, gérées et mises à jour. La base de données offre un moyen efficace d'administrer les données clients en permettant une organisation selon un modèle logique, tel que le modèle relationnel. Cela permettra d'établir des relations entre différentes données, ce qui facilite la recherche, l'analyse et l'extraction d'informations pertinentes. De plus, la base de

⁴ « export VISION_KEY=your-key »

⁵ « export VISION_ENDPOINT=your-endpoint »

données doit être conçue et gérée de manière à respecter les principes et les exigences du Règlement Général sur la Protection des Données (RGPD) pour assurer la protection des données personnelles et la vie privée des individus concernés. Ainsi, l'accès aux données doit être restreint aux employés essentiels et aux services appropriés pour assurer la conformité au RGPD.

Dans une première requête, les deux premières tables de notre base de données seront implémentées ; « Clientèle » et « Facturation ». Lors d'une seconde requête, les tables « Produits » et « Ventes » vont être incrémentées.

Azure SQL Database

Azure SQL Database est une base de données en tant que service (DBaaS) relationnelle hébergée dans Azure, qui appartient à la catégorie **PaaS** (Plateforme en tant que service).

DBeaver & JDBC

DBeaver est un outil de gestion de bases de données qui nous permet de nous connecter et d'interagir avec la base de données, via JDBC. Java Database Connectivity, est une API (Interface de Programmation Applicative) Java qui permet aux programmes d'interagir avec des bases de données relationnelles. Grâce à JDBC, DBeaver établit des connexions sécurisées, permettant aux utilisateurs d'exécuter des requêtes SQL, de visualiser et de modifier les schémas de base de données, et d'explorer les données de manière intuitive.

Le processus typique d'utilisation de JDBC implique la création d'une connexion à la base de données, l'envoi de requêtes SQL à la base de données à l'aide de cette connexion, la récupération et le traitement des résultats, puis la fermeture de la connexion une fois que les opérations sont terminées.

La base de données Billing

La base de données est représentée (Figure 2), par un schéma visuel utilisé pour représenter les tables et les relations entre elles. Les tables sont représentées par des rectangles, avec le nom de l'entité à l'intérieur. Les relations entre les tables sont représentées par les lignes qui relient les entités concernées.

Les diagrammes ER aident à visualiser la conception de la base de données de manière claire et concise, ce qui facilite la création et la maintenance de bases de données efficaces et bien organisées. Chaque table correspondant à un concept du monde réel, ici « Produit », « Facturation », « Vente » et « Clientèle ». La gestion de l'accès aux tables pourra donc se faire en fonction des services appropriés services et aux employés essentiels pour assurer la conformité au RGPD.

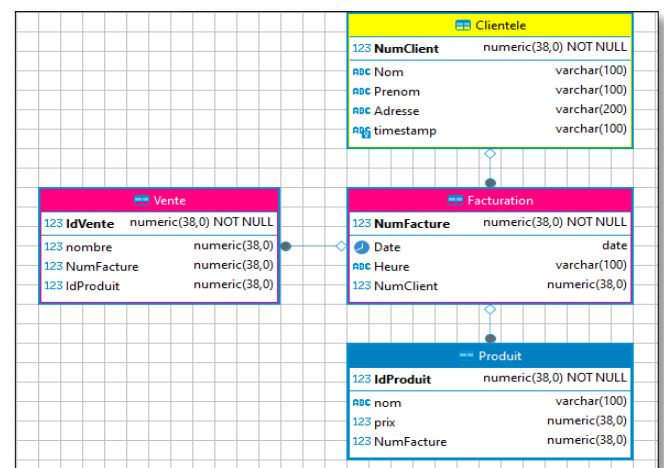


Figure 2 : Diagramme Entité-Relation (ER) de notre base de données.

La déclaration de la clé primaire de chaque table a été effectuée, et indique à la base de données quelles colonnes doivent être utilisées. Ce sont des attributs qui identifient de manière unique chaque enregistrement dans une table. La clé primaire garantira que chaque ligne, d'une table, de la base de données, est distincte et peut être identifiée de manière unique. Ainsi un produit, une facture, un client et une vente ne pourront pas être insérer en doublons dans une table. Les requêtes sont conçues de manière à ce que les données des nouvelles factures soient intégrées à celles déjà existantes dans la base de données, sans nécessiter de recommencer le processus d'OCR pour l'ensemble des factures. La table

"Clientèle" nous permet de définir les origines de nos clients, à savoir les États-Unis, la France et l'Italie. La table "Facturation" nous permet d'établir le nombre de ventes en fonction du temps, ainsi que le chiffre d'affaires annuel.

Résultats

L'analyse des factures de 2019 à 2024 (au 07/04/2024) nous a permis d'obtenir les détails de 6 215 factures, avec un chiffre d'affaires stable d'une année à l'autre. Au total, 124 697 produits ont été vendus, avec les clients achetant en moyenne 5 produits différents à chaque transaction. Cette stabilité du chiffre d'affaires et du nombre de produits vendus témoigne d'une certaine régularité dans les activités commerciales de la période analysée. Voici trois exemples de résultats que nous pouvons obtenir à partir des données stockées dans la base de données.

- 1- Le nombre de vente réalisé. Ici affiché par trimestre.

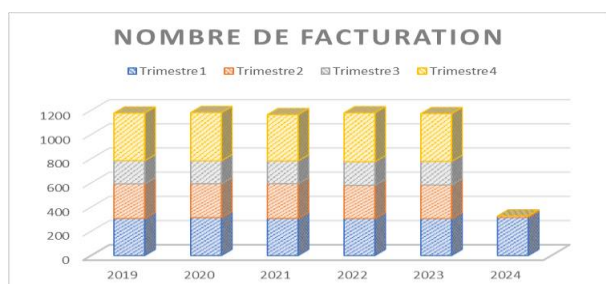


Figure 3 : Nombre de factures émises entre 2019 et 2024, pour chaque trimestre.

Il est à noter que le troisième trimestre semble être moins propice à la vente.

- 2- Le chiffre d'affaires, ici affiché annuellement.

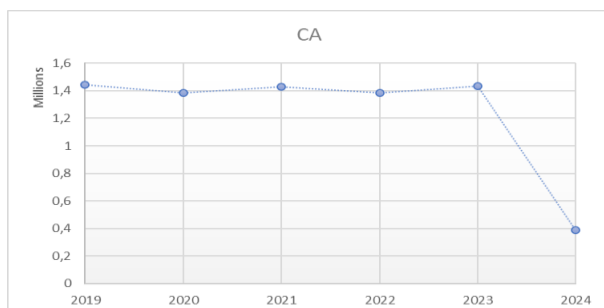


Figure 4 Chiffre d'affaires annuelle, entre 2019 et 2024.

3- L'origine géographique des clients

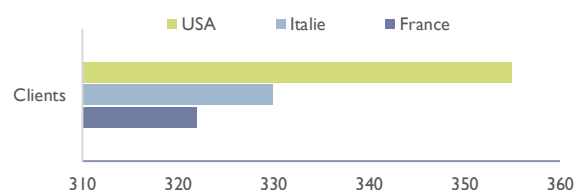


Figure 5 : Origine géographique des clients

Monitoring

En surveillant continuellement la base de données, les anomalies telles que les erreurs de saisie peuvent être détectées rapidement, assurant ainsi la fiabilité des informations utilisées pour prendre des décisions commerciales. De même, un système de monitoring dédié peut identifier et corriger rapidement les erreurs de reconnaissance de l'OCR, minimisant ainsi les risques de mauvaise interprétation des données. Ici, nous avons créé 4 vues dans la base de données pour suivre les données problématiques, telles que les quantités de produits égales à zéro et les noms de produits non reconnus.

Application

La cohérence et la stabilité observées dans les données des factures de 2019 à 2024 soulignent l'importance de maintenir une surveillance étroite de l'activité commerciale. Une application de monitoring dédiée serait donc essentielle pour permettre une gestion proactive et efficace. Cette application pourrait fournir des analyses en temps réel, permettant ainsi aux responsables de l'entreprise de détecter rapidement toute anomalie ou tendance inattendue. De plus, elle faciliterait la gestion des stocks en offrant une visibilité complète sur les ventes et les produits les plus demandés. En investissant dans une telle application, l'entreprise renforcerait sa capacité à prendre des décisions stratégiques éclairées, tout en assurant une continuité dans ses opérations commerciales.