

# *Rapport de PFA*

## *Création d'un système d'analyse de correspondance entre une description de poste et le CV d'un candidat*

*Présenté et soutenu le ..... par :*

- *Mouhyeddine Youssef*
- *Radi Anas*
- *ElBouamri Hamza*

*Devant le jury composé de :*

*Professeur..... : Président de jury*

*Professeur ..... : Examineur*

*Professeur Said NOUH : Encadrant académique*

*Année Universitaire : 2023-2024*

# *Remerciements*

On tient à exprimer notre profonde gratitude envers toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail. Nos remerciements vont tout particulièrement à Mr. Said Nouh pour son soutien, et ses conseils avisés et son inspiration.

On tient à exprimer aussi notre reconnaissance à l'EMSI pour avoir rendu possible la concrétisation de ce projet.

Nous n'oublions pas de saluer également tous ceux qui, de près ou de loin, ont contribué à faire de ce projet une réalité.

Merci à chacun d'entre vous pour votre précieuse

Enfin, on tient à dédier ce travail à nos chers parents et à toutes nos familles.

## Table des matières

Remerciement.....	i
Résumé en français .....	ii
Résumé en anglais.....	iii
Résumé en arabe.....	iv
Introduction générale .....	9
<b>Chapitre I : Cahier de charge, objectifs et gestion du projet .....</b>	<b>10</b>
Introduction .....	11
Contexte du projet.....	11
Problématique .....	11
Etude de l'existant.....	12
Objectif à atteindre .....	13
Besoin fonctionnels.....	14
Besoin non fonctionnels.....	14
Approche de Travail.....	15
Méthodologie .....	15
Formalismes adoptés .....	15
Conclusion.....	17
<b>Chapitre II : Analyse et conception du projet .....</b>	<b>18</b>
Introduction .....	19
Introduction sur LucidChart.....	19
L'outil utilisé .....	20
Le Workflow du Projet.....	21
Conclusion.....	22

<b>Chapitre III : Réalisation du projet .....</b>	<b>23</b>
Introduction .....	24
Environnement de développemnt .....	23
Langage de programmation .....	24
Bibliothèques Python .....	25
Datasets .....	30
Exemplaire Job Description .....	32
Exemplaire CV .....	33
Machine Learning .....	34
Deep Learning .....	35
CosineSimilarity.....	36
CountVectorizer .....	37
Présentation du projet réalisé (captures d'écran) .....	38
Interfaces .....	42
Conclusion.....	43
Conclusion générale .....	48
Webographie.....	49

## Table des figures

Figure 1:Linkedin Logo .....	12
Figure 2:ReKrute Logo .....	13
Figure 3: Jira Logo.....	15
Figure 4: Kanban Logo .....	16
Figure 5: Jira .....	17
Figure 6: LucidChart Logo .....	19
Figure 7: Workflow .....	21
Figure 8: PyCharm Logo .....	23
Figure 9: Python Logo .....	24
Figure 10: Streamlit Logo.....	25
Figure 11: Scikit-learn Logo.....	26
Figure 12: Matplotlib Logo.....	27
Figure 13: NumPy Logo .....	28
Figure 14:PdfPlumber Logo .....	29
Figure 15: Kaggle Logo .....	30
Figure 16: Public Datasets AWS Logo .....	31
Figure 17: Job Description.....	32
Figure 18: CV .....	33
Figure 19: Machine Learning Logo .....	34
Figure 20: Deep Learning Logo.....	35
Figure 21:Cosine Similarity.....	36
Figure 22:CountVectorizer .....	37
Figure 23: Importation des biblio ScreenShot .....	38
Figure 24: Traitement des fichiers ScreenShot.....	38
Figure 25: AboutUs et ContactUs Screenshots.....	40
Figure 26: Main Function ScreenShot .....	40
Figure 27: Extraction et Affichage ScreenShot .....	41
Figure 28:Home Page Interface .....	42
Figure 29:Upload CV Interface .....	43
Figure 30:Job Description Selection Interface.....	44
Figure 31: CV Selection Interface .....	44
Figure 32:Resultat Upload CV Interface .....	45
Figure 33: About Us Interface .....	46
Figure 34:Contact Us Interface.....	47

# Résumé en Français

*Ce projet de fin d'année a été une expérience captivante axée sur l'analyse de la correspondance entre les descriptions de stage et les curriculum vitae des candidats. En mettant en œuvre des techniques de traitement du langage naturel (NLP) et de comparaison de texte, l'objectif était d'évaluer la pertinence des CV par rapport aux exigences spécifiées dans les descriptions de stage.*

*Le processus impliquait l'extraction de données à partir des descriptions de stage et des CV, suivi de l'utilisation de méthodes telles que la similarité cosinus pour déterminer le pourcentage de correspondance. Les résultats obtenus ont permis d'identifier les candidats les mieux adaptés aux opportunités de stage, facilitant ainsi le processus de sélection.*

# Résumé en Anglais

*This end-of-year project has been a captivating experience focused on analyzing the correspondence between internship descriptions and candidates' resumes. By implementing natural language processing (NLP) techniques and text comparison, the goal was to assess the relevance of CVs to the specified internship requirements. The process involved extracting data from internship descriptions and CVs, followed by using methods such as cosine similarity to determine the match percentage. The results obtained helped identify candidates best suited for internship opportunities, streamlining the selection process*

# Résumé en Arabe

كان هذا مشروع نهاية العام تجربة مثيرة مركزة على تحليل التوافق بين وصف الفرص التدريبية والسير الذاتية للمرشحين. من خلال تنفيذ تقنيات معالجة اللغة ومقارنة النصوص، كان الهدف هو تقييم صلة السير الذاتية بالمتطلبات الطبيعية المحددة للفرص التدريبية. شملت العملية استخراج البيانات من وصف الفرص التدريبية والسير الذاتية، تلاها استخدام أساليب مثل التشابه لتحديد نسبة التوافق. ساعدت النتائج المحصلة في تحديد المرشحين الأكثر تأهيلاً لفرص التدريب، مما يسهل عملية الاختيار



# Introduction générale :

*Ce projet de fin d'année constitue la synthèse collaborative des efforts d'une équipe trinôme, représentant l'apogée de notre parcours académique. À travers cette initiative conjointe, nous avons exploré les contours de l'innovation technologique dans le domaine de la sélection des candidatures en fonction des critères spécifiés dans les descriptions de travaux demandés tout en nous focalisant sur l'analyse de la correspondance entre les descriptions de stage et les curriculums vitae des candidats.*

*Par l'adoption de techniques avancées de traitement du langage naturel (NLP) et de comparaison de texte, notre équipe s'est attelée à perfectionner la concordance entre les attentes des recruteurs et les compétences des candidats, simplifiant ainsi la tâche complexe de sélection.*

# *CHAPITRE* 1

---

## ***Cahier de charge, étude de l'existant, objectifs et gestion du projet***

## **Introduction**

Ce chapitre présente les bases fondamentales de notre projet dont nous aborderons la problématique inhérente au domaine de notre étude, offrant ainsi une vision claire des enjeux que nous cherchons à résoudre.

### **1. Contexte du projet**

Actuellement, le processus de correspondance entre les descriptions de postes et les CV des candidats est principalement manuel, ce qui entraîne des retards, des erreurs potentielles et une inefficacité globale.

Les principales motivations derrière ce projet résident dans la nécessité d'améliorer considérablement la manière dont nous identifions les candidats potentiels. Nous faisons face à une augmentation du volume de candidatures, ce qui rend la tâche de tri et de sélection de plus en plus complexe. De plus, le marché du recrutement évolue rapidement, et il est impératif que nous adoptons des méthodes innovantes pour rester compétitif

### **2. Problématique :**

- Charge de Travail Manuelle Élevée : Les recruteurs passent un temps considérable à examiner manuellement chaque CV par rapport aux exigences spécifiques d'un poste. Cela entraîne non seulement une augmentation de la charge de travail, mais également un risque accru d'omission d'informations cruciales.
- Risques d'Omission et d'Erreur : Le processus manuel expose l'entreprise à des risques d'omission de compétences clés ou de critères spécifiques liés à un poste. Les erreurs humaines peuvent entraîner une sélection inadéquate des candidats et influencer négativement la qualité des nouvelles embauches.
- Gestion Inefficace du Volume de Candidatures : L'augmentation du nombre de candidatures rend la gestion manuelle de la correspondance de plus en plus complexe. Les recruteurs sont confrontés à des défis pour trier rapidement et efficacement les candidatures, entraînant une baisse potentielle de la qualité du processus de sélection.

### 3. Etude de l'existant :

Notre étude met en lumière le besoin d'une solution plus efficace et flexible. Bien que certaines alternatives existent telles que la correspondance manuelle et les plateformes de recrutement comme LinkedIn et ReKruite qui sont deux acteurs majeurs dans le domaine des plateformes de recrutement et se posent en concurrents significatifs dans le contexte de la correspondance entre les CV et les descriptions de postes, aucune n'a pleinement réussi à résoudre les défis liés à la correspondance entre les CV et les descriptions de postes de manière exhaustive et efficiente. Notre projet vise à combler ces lacunes en développant une solution novatrice qui offre précision, automatisation, et adaptabilité aux besoins spécifiques de notre processus de recrutement.

Malgré leur influence, ces plateformes présentent certaines limites que notre projet ambitieux aspire à dépasser.

#### - LinkedIn :



*Figure 1: LinkedIn Logo*

LinkedIn, en tant que réseau professionnel de premier plan, permet aux utilisateurs de créer des profils détaillés exposant leur parcours professionnel, compétences et éducation. Bien qu'il offre une visibilité considérable aux candidats et aux recruteurs, la correspondance sur LinkedIn repose souvent sur des critères de mots-clés, manquant parfois de précision contextuelle. Notre projet se distingue en adoptant des techniques avancées de traitement du langage naturel (NLP) pour une compréhension plus approfondie du contenu des CV et des descriptions de postes, visant ainsi à accroître la précision et l'efficacité de la correspondance.

- **ReKruite :**



*Figure 2:ReKruite Logo*

ReKruite, une plateforme de recrutement automatisée, propose des solutions pour le tri des CV et la correspondance avec les postes vacants. Bien que ces solutions offrent une certaine automatisation, elles peuvent parfois manquer de flexibilité pour s'adapter à des processus de recrutement spécifiques. Notre projet se positionne comme une solution plus souple et adaptable, mettant l'accent sur la précision, l'automatisation et l'adaptabilité pour répondre de manière optimale aux besoins uniques de chaque processus de recrutement.

Ainsi, tout en reconnaissant la contribution notable de LinkedIn et ReKruite dans le secteur, notre projet aspire à surmonter les limitations actuelles en offrant une approche plus complète et innovante. En concentrant nos efforts sur la précision de la correspondance, l'automatisation intelligente et l'adaptabilité, nous visons à redéfinir les normes de l'efficacité et de la personnalisation dans le domaine de la correspondance CV-poste.

#### **4. Objectifs à atteindre :**

Le projet vise à automatiser la correspondance entre les CV et les descriptions de postes, avec des objectifs clés tels que la réduction des délais, l'amélioration de la précision, la personnalisation du processus, une intégration transparente, la formation de l'équipe, la garantie de la sécurité des données, la réduction des coûts, et l'obtention d'un feedback positif des candidats et des recruteurs. Ces objectifs convergent vers une transformation efficiente du processus de recrutement.

## **5. Besoins :**

Les besoins de l'application se déclinent en deux catégories essentielles : les besoins fonctionnels et non fonctionnels.

### **Besoins Fonctionnels :**

- Automatisation de la Correspondance : La solution doit être capable d'automatiser la correspondance entre les compétences des candidats et les critères spécifiques des postes, réduisant ainsi la charge de travail manuelle.
- Personnalisation du Processus : Les recruteurs doivent avoir la possibilité de personnaliser les critères de correspondance en fonction des besoins spécifiques de chaque poste, assurant ainsi une flexibilité adaptée à la diversité des profils recherchés.
- Formation Intuitive : L'interface utilisateur doit être intuitive, facilitant une prise en main rapide pour l'équipe de recrutement avec un minimum d'efforts de formation.

### **Besoins Non Fonctionnels :**

- Sécurité des Données : La solution doit garantir la sécurité des données des candidats, conformément aux normes de protection des données en vigueur.
- Performances Élevées : La solution doit offrir des performances élevées, assurant une correspondance rapide et précise même avec un volume élevé de CV et de postes.
- Évolutivité : L'application doit être évolutive pour s'adapter à la croissance future du volume de candidatures et à l'évolution des besoins de recrutement.
- Conformité Réglementaire : La solution doit être conforme aux réglementations en matière de protection des données et de recrutement, garantissant une utilisation éthique et légale.

## 6. Approche de Travail

### 6.1. Méthodologie :

Notre approche pour le développement de la solution de correspondance automatisée a été guidée par l'agilité, avec une focalisation particulière sur le Framework Kanban, complété par l'utilisation de Jira pour la gestion de projet. Cette approche agile nous a permis d'ajuster dynamiquement notre trajectoire en fonction des retours et des changements de priorités. En adoptant une approche itérative, nous avons favorisé une flexibilité accrue, mettant en œuvre des fonctionnalités par étapes pour une meilleure adaptabilité.

### 6.2. Formalismes adoptés :

Le projet a été soigneusement orchestrée avec l'utilisation du tableau Kanban dans Jira. Les cartes Kanban représentaient chaque tâche et fonctionnalité, de la conception à la mise en œuvre, offrant une vue visuelle complète du flux de travail. Les sprints et les cycles de développement étaient planifiés et gérés dans Jira, offrant une traçabilité complète des progrès et des problèmes.

- **Jira** : développé par Atlassian, est une plateforme de gestion de projet qui prend en charge différentes méthodologies, y compris Scrum et Kanban. Elle offre un ensemble complet d'outils pour la planification, le suivi des tâches, la collaboration d'équipe, et la gestion des problèmes. Jira utilise un tableau Kanban pour visualiser les tâches, mais il propose également des fonctionnalités plus avancées telles que la gestion des sprints, la création de feuilles de route, et la personnalisation des workflows. Jira est apprécié pour sa polyvalence, sa capacité à gérer des projets complexes, et sa traçabilité détaillée des tâches. Il facilite la collaboration d'équipe en centralisant les informations liées au projet, des exigences aux rapports de bugs.



*Figure 3: Jira Logo*

- **Kanban** : méthodologie de gestion de projet visuelle qui se concentre sur l'optimisation du flux de travail. Les principaux éléments de Kanban comprennent un tableau Kanban, des cartes représentant les tâches, et des colonnes décrivant les différents statuts du processus. Le tableau offre une vue d'ensemble des tâches, de leur création à leur achèvement. Les équipes utilisent Kanban pour maximiser l'efficacité en minimisant les goulots d'étranglement et en facilitant la collaboration.

Les avantages de Kanban incluent sa simplicité, sa flexibilité et sa capacité à s'adapter à des flux de travail changeants. Il permet également une gestion visuelle et transparente des tâches, offrant une visibilité en temps réel sur les progrès.



*Figure 4: Kanban Logo*

#### **Quelques avantages de Kanban par rapport à Scrum :**

- **Pas de Sprint Fixe** :

**Kanban** : Il n'y a pas de sprint fixe dans Kanban. Les équipes travaillent sur des tâches en continu sans les contraintes de sprints.

**Scrum** : utilise des sprints de durée fixe, généralement de deux à quatre semaines. Cela peut créer des contraintes de temps pour certaines équipes.

- **Aucun Rôle Prédéfini** :

**Kanban** : n'impose pas de rôles prédéfinis en permettant aux membres de l'équipe de collaborer sans nécessiter des rôles spécifiques comme le Product Owner ou le Scrum Master.

**Scrum** : définit des rôles spécifiques, comme le Product Owner, le Scrum Master et l'équipe de développement. Cela peut être contraignant pour des équipes.



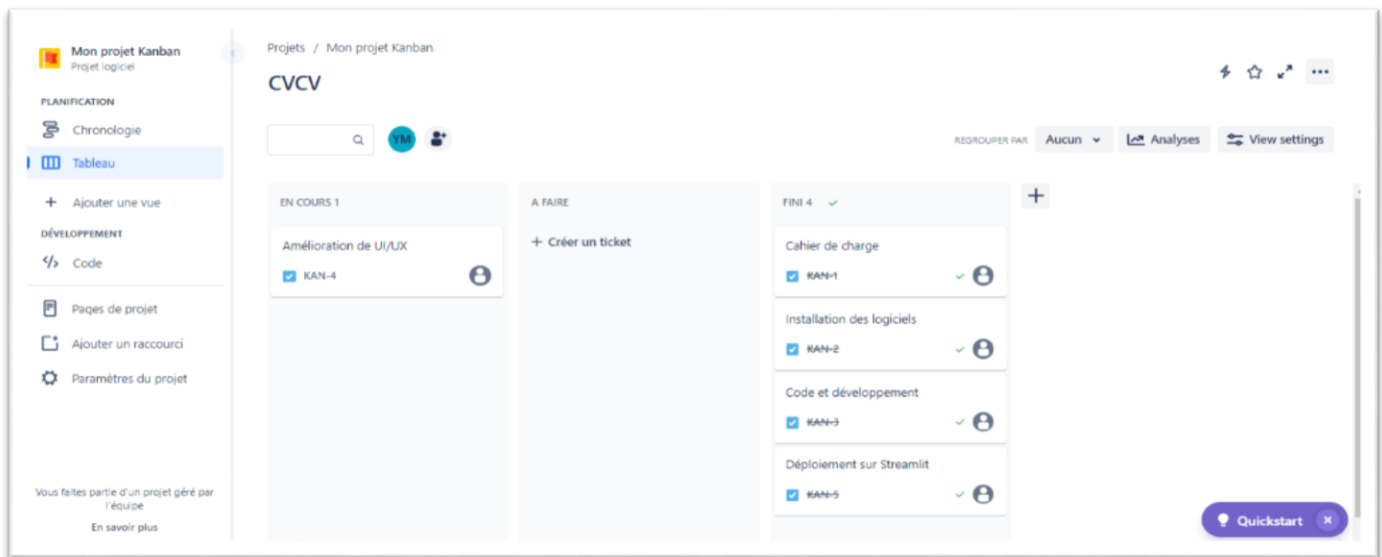


Figure 5: Jira

## Conclusion

L'étude de l'existant a permis de poser la problématique du sujet, d'adopter la méthodologie du travail et de sélectionner les données essentielles pour la conception de l'application.

# CHAPITRE 2

## *Analyse et conception du projet*

## Introduction

Dans ce chapitre, nous plongerons dans la conception et l'analyse détaillées du projet, mettant en lumière les étapes clés du processus. Nous explorerons également l'utilisation de Lucidchart, un outil puissant de création de diagrammes, pour visualiser et comprendre le workflow du projet

## Introduction sur LucidChart

Lucidchart offre une plateforme intuitive permettant de créer des diagrammes de flux, des organigrammes et d'autres représentations visuelles. Son interface conviviale facilite la modélisation des processus, offrant une vision claire du déroulement des activités du projet.



*Figure 6: LucidChart Logo*

- **Interface Collaborative** : Permet la collaboration en temps réel entre plusieurs utilisateurs, facilitant le travail d'équipe et la communication.
- **Bibliothèque de Formes** : Propose une vaste gamme de formes et d'icônes prédéfinies pour faciliter la création de diagrammes professionnels.
- **Intégrations** : S'intègre souvent à d'autres outils de productivité tels que Google Workspace, Microsoft Office, Slack, etc.

Lucidchart est largement utilisé dans les entreprises, les écoles et d'autres contextes où la création de diagrammes et de schémas est nécessaire. Il facilite la visualisation des idées, la planification de projets, la documentation des processus, et bien d'autres activités liées à la communication visuelle.

## **L'outil utilisé :**

Dans le cadre de la conception et de l'analyse du projet, Lucidchart sera l'outil principal pour élaborer le workflow. Nous exploiterons les formes, les connexions et les annotations pour créer un diagramme informatif et esthétiquement plaisant. Cette démarche permettra une communication efficace des idées au sein de l'équipe et avec les parties prenantes.

## **Introduction sur le Workflow :**

Un workflow (ou flux de travail en français) est une séquence d'activités, de tâches ou d'étapes interconnectées qui sont réalisées pour atteindre un objectif spécifique. Il s'agit d'une représentation visuelle ou d'une description séquentielle des différentes étapes d'un processus, depuis le début jusqu'à la réalisation de l'objectif final.

- **Définition des Étapes** : Un workflow définit clairement les étapes ou les actions nécessaires pour atteindre un résultat spécifique. Chaque étape peut impliquer des actions manuelles ou automatisées.
- **Séquence Logique** : Les étapes d'un workflow sont généralement organisées de manière logique et séquentielle, reflétant la façon dont le processus doit être exécuté.
- **Collaboration** : Certains workflows impliquent la collaboration entre différentes parties prenantes. Les responsabilités et les autorisations peuvent être clairement définies pour chaque étape.
- **Automatisation** : Dans de nombreux cas, les workflows peuvent être automatisés à l'aide de systèmes informatiques ou de logiciels de gestion de workflow. Cela permet d'améliorer l'efficacité, de réduire les erreurs et d'accélérer le déroulement du processus.
- **Contrôle et Suivi** : Les workflows peuvent inclure des mécanismes de contrôle et de suivi pour garantir que le processus se déroule correctement. Des alertes, des notifications ou des rapports peuvent être intégrés pour informer les parties prenantes de l'avancement.

## Le WorkFlow du projet

Le workflow du projet est une représentation visuelle du processus de correspondance entre les descriptions de poste et les CV. À travers un workflow diagram conçu sur Lucidchart, nous détaillerons chaque étape, mettant en évidence les points de décision et les résultats attendus.

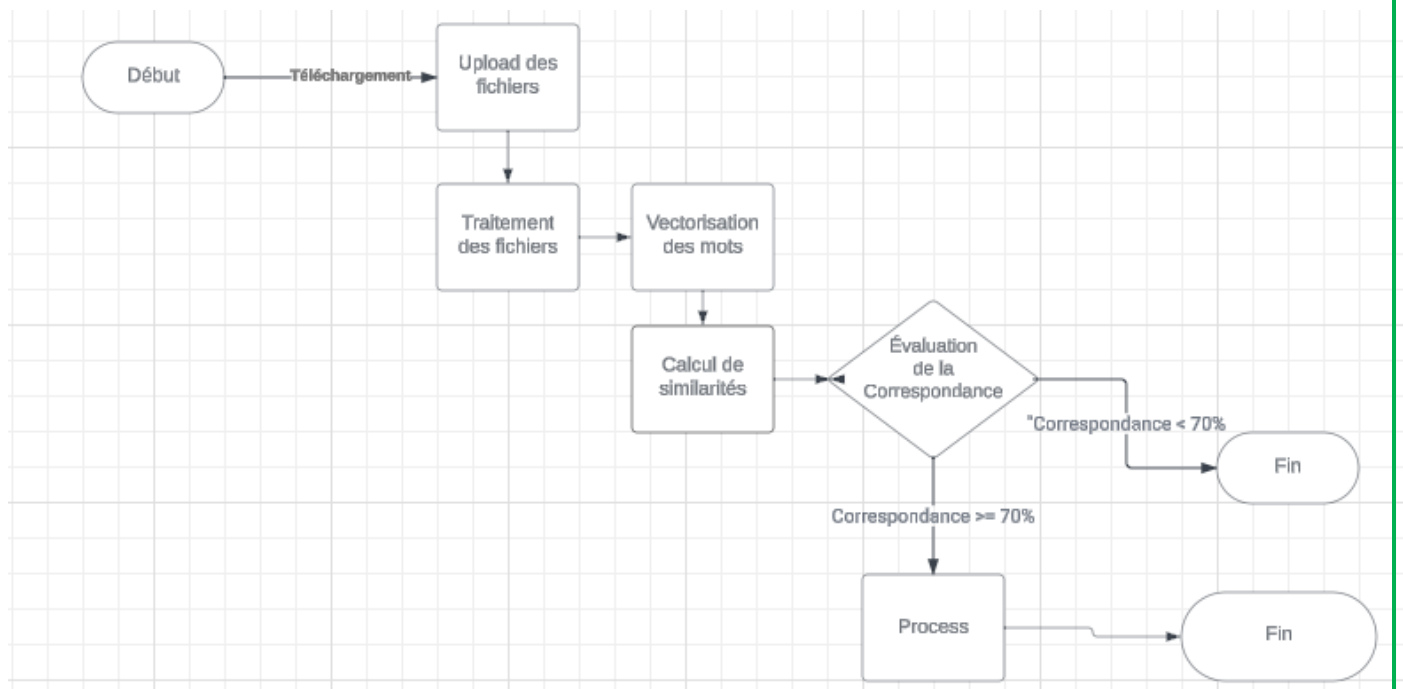


Figure 7: Workflow

## Conclusion:

Ce chapitre a jeté les bases de la conception et de l'analyse du projet en utilisant Lucidchart comme outil visuel essentiel. Le workflow Diagram élaboré servira de référence centrale tout au long du projet, offrant une compréhension visuelle du processus complexe de correspondance entre les descriptions de poste et les CV. Dans les chapitres suivants, nous approfondirons chaque étape du workflow, explorant les aspects techniques et conceptuels du projet.

# *CHAPITRE* 3

---

## *Réalisation du projet*

## Introduction

La réalisation du projet repose sur l'utilisation d'outils de développement spécifiques, du choix de langages de programmation adaptés, et de la création d'un environnement de développement propice à l'efficacité du travail. Dans cette section, nous explorerons ces éléments cruciaux qui ont contribué à la concrétisation du projet

## Environnement de développement :

**PyCharm :** environnement de développement intégré (IDE) spécialement conçu pour les développeurs Python. Il offre un ensemble de fonctionnalités puissantes pour simplifier le processus de développement, le débogage, la gestion de projet, et la collaboration.

- **Interface Utilisateur Intuitive :** PyCharm propose une interface utilisateur conviviale qui facilite la navigation et l'utilisation des fonctionnalités avancées de l'IDE.
- **Analyse de Code :** PyCharm intègre des outils d'analyse statique de code qui aident les développeurs à détecter les erreurs potentielles, à améliorer la qualité du code, et à respecter les bonnes pratiques de codage.
- **Débogage Puissant :** L'IDE prend en charge des fonctionnalités de débogage avancées, notamment le débogage à distance, les points d'arrêt, l'inspection de variables, et la
- **Gestion de Projet :** PyCharm propose des outils de gestion de projet robustes, permettant aux développeurs de créer, d'organiser, et de naviguer dans des projets Python de toutes tailles.
- **Intégration de l'Environnement Virtuel :** PyCharm simplifie la gestion des environnements virtuels Python, facilitant ainsi la gestion des dépendances et des versions pour chaque projet.



Figure 8: PyCharm Logo

## Langage de programmation :

**Python :** Python est un langage de programmation interprété, polyvalent et largement utilisé, réputé pour sa syntaxe claire et sa facilité d'apprentissage

- **Facilité d'Apprentissage :** La syntaxe simple et lisible de Python en font un langage idéal pour les débutants. Il favorise une courbe d'apprentissage douce tout en permettant aux développeurs de créer des applications puissantes.
- **Bibliothèques et Frameworks :** Python dispose d'une vaste bibliothèque standard qui couvre de nombreux aspects du développement logiciel. De plus, il est pris en charge par de nombreux frameworks populaires tels que Django, Flask, TensorFlow et PyTorch.
- **Support Multiplateforme :** Python est compatible avec plusieurs systèmes d'exploitation, ce qui permet aux développeurs de créer des applications qui peuvent être exécutées sur différentes plateformes sans modifications significatives du code source
- **Automatisation et Scripting :** Python excelle dans l'automatisation de tâches et le scripting grâce à sa simplicité et à ses capacités puissantes, ce qui en fait un choix populaire pour les administrateurs système.



Figure 9: Python Logo



## Bibliothèques Python :

**Streamlit :** bibliothèque Python open source conçue pour simplifier le processus de création d'applications web interactives pour la visualisation de données et la démonstration de modèles machine learning. Streamlit se distingue par sa facilité d'utilisation, permettant aux développeurs de créer rapidement des applications sans avoir à maîtriser des compétences avancées en développement web. Voici quelques points clés à propos de Streamlit :

- **Simplicité d'Utilisation :** Streamlit met l'accent sur la simplicité d'utilisation. Les développeurs peuvent créer des applications web en quelques lignes de code, ce qui en fait un outil idéal pour les prototypes rapides et les démonstrations.
- **Support Interactif :** Streamlit prend en charge les éléments interactifs tels que les widgets, les boutons, les curseurs, etc. Ces éléments permettent aux utilisateurs de manipuler les données ou de personnaliser l'affichage de manière interactive.
- **Visualisation de Données :** Streamlit s'intègre facilement avec des bibliothèques de visualisation de données populaires telles que Matplotlib, Plotly et Altair, permettant aux développeurs d'intégrer des graphiques et des visualisations de données dans leurs applications.
- **Documentation Complète :** Streamlit dispose d'une documentation détaillée et conviviale, ainsi que d'une communauté active qui partage des exemples, des astuces et des solutions aux problèmes courants.



*Figure 10: Streamlit Logo*

**Scikit-Learn** : Bibliothèque pour l'apprentissage automatique, utilisée pour la vectorisation des mots et le calcul de la similarité cosinus.

- **Large Éventail d'Algorithmes** : Scikit-learn propose une variété d'algorithmes d'apprentissage machine, notamment pour la classification, la régression, le clustering, la réduction de dimension, et bien d'autres. Ces algorithmes sont implémentés de manière cohérente, facilitant leur utilisation.
- **Interface Cohérente** : L'API de scikit-learn est conçue de manière à être simple, cohérente et facile à comprendre. Cela permet aux utilisateurs de se concentrer sur les concepts fondamentaux de l'apprentissage machine plutôt que sur des détails d'implémentation complexes.
- **Prétraitement de Données** : Scikit-learn propose des outils robustes pour le prétraitement des données, tels que la normalisation, la standardisation, l'imputation des valeurs manquantes, la gestion des catégories, et plus encore.
- **Évaluation de Modèles** : La bibliothèque fournit des métriques d'évaluation de modèles, telles que la précision, le rappel, la courbe ROC, l'erreur quadratique moyenne, etc., facilitant l'évaluation des performances des modèles.
- **Apprentissage Non Supervisé et Supervisé** : Scikit-learn prend en charge à la fois les problèmes d'apprentissage supervisé (classification, régression) et non supervisé (clustering, réduction de dimension), offrant une couverture complète des besoins en apprentissage machine.



Figure 11: Scikit-learn Logo

**Matplotlib** : est une bibliothèque de visualisation de données en Python. Elle offre une grande flexibilité pour créer une variété de graphiques, de diagrammes et de visualisations, ce qui en fait l'une des bibliothèques les plus populaires dans le domaine de la visualisation de données. Voici quelques points clés à propos de Matplotlib :

- **Large Gamme de Graphiques** : Matplotlib prend en charge la création d'une large gamme de graphiques, notamment des graphiques linéaires, des histogrammes, des diagrammes en barres, des nuages de points, des graphiques en boîte, des diagrammes de violon, des cartes thermiques, et bien plus encore.
- **Personnalisation** : Matplotlib offre un haut degré de personnalisation pour les graphiques. Les utilisateurs peuvent ajuster presque tous les aspects d'un graphique, y compris les couleurs, les styles de ligne, les légendes, les étiquettes d'axe, et les titres.
- **Interface Orientée Objet** : Matplotlib fonctionne avec une interface orientée objet, ce qui signifie que les graphiques peuvent être créés et modifiés en manipulant des objets Python. Cela permet un contrôle précis sur l'apparence des graphiques.
- **Support Multiplateforme** : Matplotlib fonctionne sur différentes plates-formes (Windows, macOS, Linux) et peut être utilisé avec différentes interfaces utilisateur, notamment Tkinter, PyQt, WxPython, et plus encore.
- **Plusieurs Interfaces d'Utilisation** : En plus de l'interface orientée objet, Matplotlib offre également une interface de style MATLAB via Pyplot, qui est plus adaptée aux utilisateurs familiers avec MATLAB.



Figure 12: Matplotlib Logo

**Numpy :** est une bibliothèque fondamentale en Python pour le calcul numérique. Elle fournit des structures de données, des algorithmes et des outils pour travailler efficacement avec des tableaux multidimensionnels et effectuer des opérations mathématiques avancées.

- **Tableaux Multidimensionnels :** NumPy introduit un type de données appelé "array" qui représente un tableau multidimensionnel. Ces tableaux sont efficaces pour stocker et manipuler des données, en particulier dans le contexte du calcul numérique.
- **Calcul Vectorisé :** NumPy favorise le calcul vectorisé, permettant d'effectuer des opérations sur des tableaux entiers sans avoir besoin de boucles explicites. Cela conduit à des performances améliorées et à un code plus concis.
- **Opérations Mathématiques :** NumPy offre un ensemble étendu de fonctions mathématiques pour effectuer des opérations telles que les fonctions trigonométriques, exponentielles, logarithmiques, ainsi que des opérations sur les matrices et les polynômes.
- **Intégration avec d'Autres Bibliothèques :** NumPy s'intègre bien avec d'autres bibliothèques populaires en Python, notamment pandas (pour la manipulation de données), Matplotlib (pour la visualisation) et SciPy (pour des fonctionnalités avancées).
- **Traitement des Données :** NumPy facilite le traitement des données, en fournissant des fonctionnalités telles que le tri, la recherche, la manipulation, et la statistique descriptive.

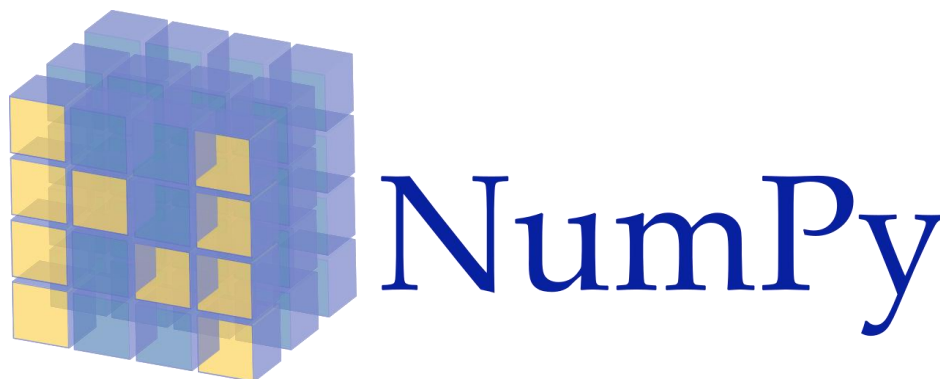


Figure 13: NumPy Logo

**PdfPlumber :** bibliothèque Python utilisée pour extraire du texte, des images et des métadonnées à partir de documents PDF. Elle est construite en utilisant la bibliothèque pdfminer.six et offre une interface de niveau supérieur pour travailler avec des fichiers PDF.

- **Extraction de Texte :** pdfplumber permet d'extraire le contenu textuel des documents PDF. Cela est particulièrement utile lorsque vous avez besoin de convertir le texte d'un PDF dans un format plus accessible pour l'analyse, la recherche ou la manipulation.
- **Extraction de Tableaux :** Les tableaux dans les PDF peuvent être difficiles à manipuler. pdfplumber facilite l'extraction des données tabulaires des PDF, rendant plus simple la conversion des tableaux dans des formats comme CSV ou Excel.
- **Extraction d'Images :** Les PDF peuvent contenir des images, et pdfplumber permet l'extraction de ces images. Cela est utile lorsque vous souhaitez travailler avec des graphiques ou des éléments visuels intégrés dans le PDF.
- **Récupération des Métadonnées :** Les métadonnées comprennent des informations sur le PDF, telles que le titre, l'auteur, la date de création, etc. pdfplumber permet d'accéder à ces métadonnées, ce qui peut être précieux pour cataloguer ou organiser des documents.
- **Manipulation de Pages :** pdfplumber offre la possibilité de manipuler les pages. Par exemple, vous pouvez faire pivoter une page, ce qui peut être utile si l'orientation du contenu doit être ajustée.



Figure 14: PdfPlumber Logo

## Datasets :

Nous avons conservé un très grand ensemble de données (plus de 2 800 CV) dans 24 domaines différents en utilisant Kaggle et AWS Public Datasets

**Kaggle** : une plateforme en ligne dédiée à la science des données, à l'apprentissage machine et à la communauté des analystes de données. Lancée en 2010, Kaggle offre un environnement où les professionnels et les passionnés de la data peuvent collaborer, partager leurs travaux, participer à des compétitions et accéder à des ensembles de données variés. Voici quelques points clés à propos de Kaggle :

- **Compétitions de Science des Données** : Kaggle propose régulièrement des compétitions de science des données où les participants peuvent résoudre des problèmes réels en utilisant des techniques d'apprentissage machine. Ces compétitions couvrent divers domaines, tels que la classification, la régression, la segmentation d'images, et plus encore.
- **Ensembles de Données** : La plateforme met à disposition une vaste collection d'ensembles de données provenant de divers domaines. Les utilisateurs peuvent explorer, télécharger et utiliser ces données pour leurs propres projets de recherche ou d'analyse.

The Kaggle logo is displayed in a large, light blue, lowercase, sans-serif font. The letters are thick and rounded, with a friendly, approachable feel.

*Figure 15: Kaggle Logo*

**AWS Public Datasets :** Les jeux de données publics d'AWS (Amazon Web Services) constituent une initiative visant à rendre accessible au public une variété de jeux de données de grande valeur. Ces jeux de données sont hébergés sur l'infrastructure cloud d'AWS, facilitant ainsi la découverte, l'accès et l'analyse de données sans nécessiter une infrastructure de stockage et de gestion de données à grande échelle.

- **Gamme Diversifiée de Jeux de Données :** Les jeux de données publics d'AWS couvrent un large éventail de domaines et de sujets, incluant la biologie, le climat, l'astronomie, la génomique, les soins de santé, l'économie, et bien d'autres. Cette diversité permet aux chercheurs et aux développeurs de trouver des jeux de données pertinents pour leurs domaines d'intérêt spécifiques.
- **Accès Gratuit et Ouvert :** La plupart des jeux de données sur AWS Public Datasets sont librement accessibles au public. Cette ouverture encourage la collaboration, la recherche et l'innovation dans divers secteurs industriels et disciplines académiques.
- **Programme de Jeux de Données Publics :** AWS soutient le Programme de Jeux de Données Publics, qui permet aux chercheurs et aux organisations de contribuer à leurs jeux de données pour qu'ils soient hébergés sur AWS. Ce programme facilite le partage de jeux de données de valeur avec la communauté mondiale.

# Open Data on AWS Data Exchange

---

Figure 16: Public Datasets AWS Logo

## Exemple Job Description

### HR Manager Job Description

Our innovative and growing company is hiring for a HR manager. We appreciate you taking the time to review the list of qualifications and to apply for the position. If you don't fill all of the qualifications, you may still be considered depending on your level of experience.

#### Responsibilities for HR manager

- Assists management with setting and reviewing employee goals
- Guides and supports management through the process of addressing under-performing employees
- Interprets policies and procedures to ensure management is consistent with disciplinary action and overall decision making
- Act as a business partner to both our Head Office and Retail functions, across a wide spectrum of people matters
- Act as a true generalist, turning your hand to various projects across areas such as
- Participates in disciplinary actions and conflict resolution situations
- Oversees hiring and takes part in the selection, interviewing, and/or hiring of leadership and other key positions in assigned area
- Establish strong partnerships with Site leadership and employees in support of developing the optimal culture to drive high performance, retention and engagement
- Develop and cultivate active, two-way communications with all levels of employees in all functions on all shifts
- Support the China business strategy


#### Qualifications for HR manager

- 3 years or more experience in an HR Generalist or Manager position
- Good IT and HR management system/tools skills
- CIPD or equivalent qualifications would be an advantage
- 7-10 years of HR experience in GCC region
- Experience with multinational organization would be a strong advantage
- Minimum 5 years Human Resources experience managing a significant HR function

*Figure 17: Job Description*



## Exemple CV :



### CONTACT

✉ info@resumekraft.com  
☎ 202-555-0120  
📍 Chicago, Illinois, US  
in linkedin.com/resumekraft

### SKILLS

SAP - Success Factor  
SAP - Field Glass  
eCATS - Candidate Application Tracking System  
Workday  
Canva  
MS Excel, word  
Power Point Presentation  
MS Visio

### LANGUAGES

English ●●●●●  
French ●●●●●  
Arabic ●●●●●  
German ●●●●●

## RICHARD TIM

HUMAN RESOURCE PROFESSIONAL

### SUMMARY

A Human Resource Professional with 6+years of experience in sectors namely - Generic Pharmaceuticals, Critical Care Injectables, Medical Devices and FMCG. A speaker, counsellor and consultant.

### EXPERIENCE

#### HR Manager

Sep 2019 - Present

#### Zydus Wellness Ltd

- Partnering with Sales Managers of 4 businesses namely CPD, NL, Key Accounts & Modern Trade of North India.
- Counselling and consulting with suggestions and support to field & branch employees during organizational integration.
- Handling operating strength of north team and branch.
- Handling Talent acquisition of positions starting from SO, ASM and RSM levels.
- Planning and monitoring joining and onboarding of the new joiner.
- Drive the Induction process in the region for new joiners and ensuring thorough readiness of new joiners for role delivery.
- Ensuring timely execution of changed policies or procedures at branch level.
- Ensuring exit formalities to carry out seamlessly at branch.
- Taking exit feedback and maintaining the records.
- Handling grievances & concerns related to PF, mediclaim, leaves, expense claiming and others, of field employees.
- Ensuring employee engagement activities like birthday celebrations, festivals celebrations.
- Process owner of SHIKHAR program - PGDBA program for sales professional in collaboration with Symbiosis Distance Learning Pvt. Ltd.

#### Deputy Manager

Apr 2018 - Sep 2019

#### Zydus Wellness Ltd.

- Handled Talent acquisition of domestic business of CPD and NL at Pan India level.
- Worked on Talent Retention thoroughly with the help of counselling and consultation.
- Coordinated in employee life cycle mainly confirmation and appraisal.
- Planned field induction, onboarding and took feedbacks.
- Revamped organizational introduction induction PPT for New Joiner.
- Took exit interviews and feedbacks.
- Prepared and doing exit analysis monthly.
- Handled Talent Management "LEAP" program for field staff.
- Involved in Strategic hiring for head office.
- Prepared monthly Hiring, induction, operating strengths report of zones.

Figure 18: CV

## Apprentissage Machine (Machine Learning) :

L'apprentissage machine est une branche de l'intelligence artificielle qui se concentre sur le développement de modèles et d'algorithmes permettant aux systèmes informatiques d'apprendre à partir de données. Plutôt que de programmer explicitement des règles, l'apprentissage machine permet aux machines d'identifier des modèles et de prendre des décisions autonomes en se basant sur l'expérience accumulée. Il existe plusieurs approches en apprentissage machine, notamment l'apprentissage supervisé et l'apprentissage non supervisé

- **Apprentissage Supervisé (Supervised Learning)** : le modèle est entraîné sur un ensemble de données étiquetées, où les entrées sont associées à des sorties connues. Le but est d'apprendre la relation entre les entrées et les sorties afin de faire des prédictions sur de nouvelles données non étiquetées. Exemples : Classification (prédiction d'une étiquette ou d'une classe), Régression (prédiction d'une valeur continue).
- **Apprentissage Non Supervisé (Unsupervised Learning)** : utilise des ensembles de données non étiquetés. Le modèle cherche à découvrir des structures, des relations ou des motifs intrinsèques aux données sans avoir de sortie cible spécifique à prédire. Exemples : Clustering (regroupement de données similaires), Réduction de dimension (compression des données tout en préservant les informations importantes).

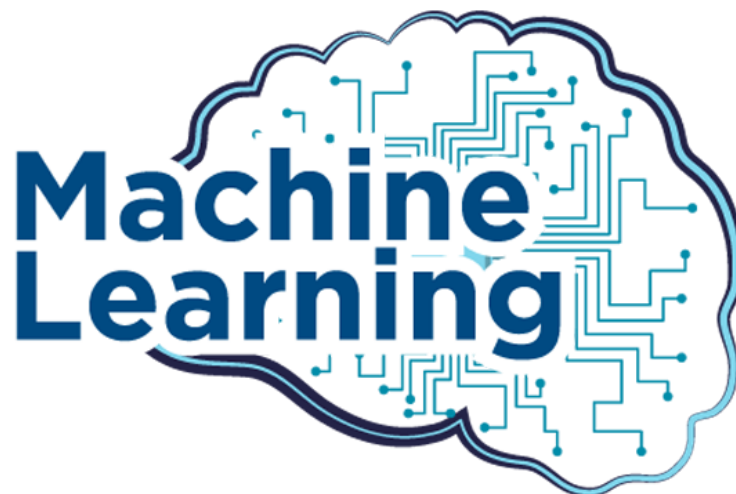


Figure 19: Machine Learning Logo

## Deep Learning (Apprentissage profond) :

Le Deep Learning est une sous-discipline de l'apprentissage machine qui utilise des réseaux de neurones artificiels pour modéliser et résoudre des problèmes complexes. Inspiré par le fonctionnement du cerveau humain, il implique des architectures de réseau neuronal profondes composées de multiples couches (d'où le terme "profond"). Ces réseaux sont capables d'apprendre des représentations hiérarchiques complexes des données, ce qui les rend particulièrement efficaces pour des tâches telles que la reconnaissance d'images, la traduction automatique, la classification de texte, et bien d'autres

- **Réseaux de Neurones Convolutifs (Convolutional Neural Networks - CNN) :** spécialement conçus pour le traitement des images et la reconnaissance visuelle. Ils sont capables d'apprendre des caractéristiques hiérarchiques à partir d'images grâce à l'utilisation de filtres de convolution.
- **Réseaux de Neurones Récurrents (Recurrent Neural Networks - RNN) :** sont utilisés pour traiter des séquences de données, telles que des séquences temporelles, du langage naturel ou des données de texte. Ils possèdent une mémoire interne qui leur permet de conserver des informations sur les états précédents.
- **Réseaux de Neurones Longue Courte Mémoire (Long Short-Term Memory - LSTM) :** sont un type spécialisé de RNN conçu pour surmonter le problème de disparition du gradient. Ils sont particulièrement adaptés pour capturer et mémoriser des dépendances à long terme dans des séquences de données.

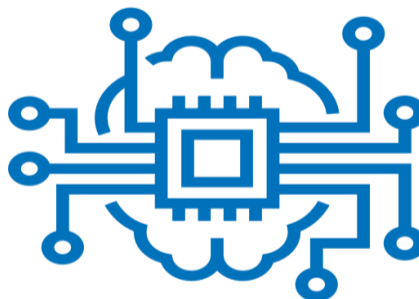
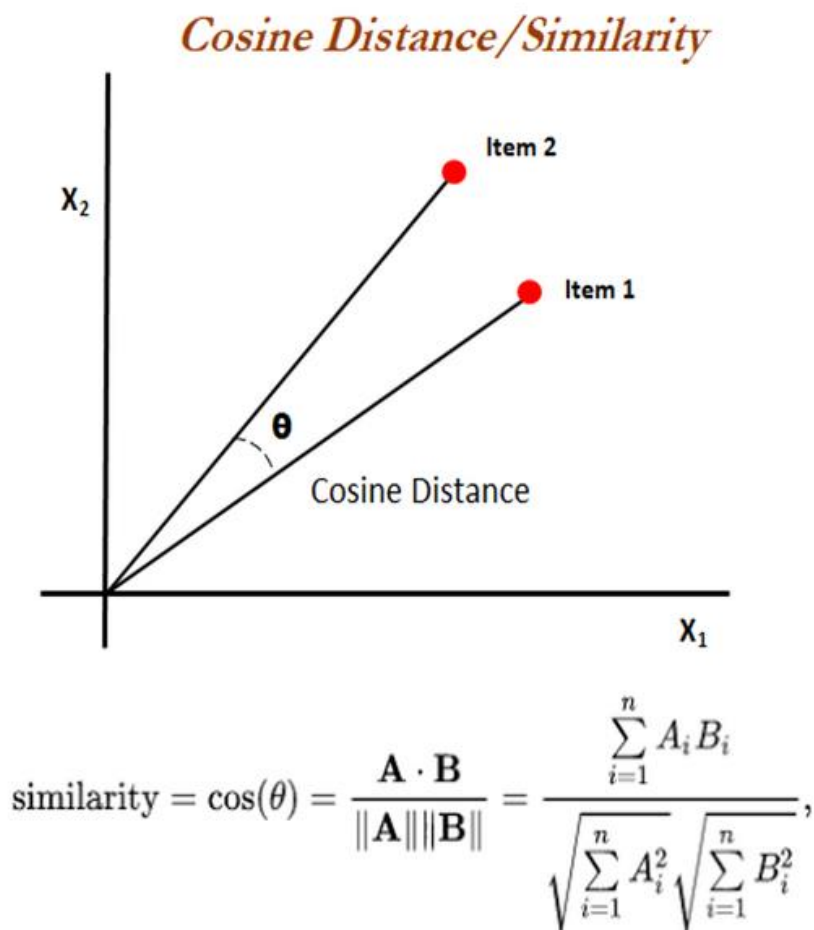


Figure 20: Deep Learning Logo

## *CosineSimilarity :*

La similarité cosinus est une mesure de similarité entre deux vecteurs dans un espace vectoriel, utilisée fréquemment en traitement du langage naturel et en analyse de texte. Elle évalue la proximité directionnelle entre deux vecteurs dans un espace multidimensionnel.

La formule de la similarité cosinus est définie comme le cosinus de l'angle  $\theta$  entre deux vecteurs A et B :



*Figure 21: Cosine Similarity*

La similarité cosinus produit une valeur comprise entre -1 et 1. Une valeur de 1 indique une similarité parfaite, 0 indique une absence de similarité, et -1 indique une dissimilarité parfaite. Cette mesure est largement utilisée dans des applications telles que la recherche d'information, la classification de documents, et la recommandation de contenu. Elle permet de quantifier la proximité sémantique entre des représentations vectorielles de documents, de mots, ou d'autres entités textuelles.

## CountVectorizer :

CountVectorizer est un outil de traitement du langage naturel, notamment une classe de la bibliothèque Scikit-learn en Python, qui permet de convertir une collection de documents textuels en une représentation numérique. Plus précisément, il transforme ces documents en une matrice de comptage, où chaque ligne correspond à un document et chaque colonne à la fréquence d'occurrence des mots.

Imaginons un ensemble de documents tels que des articles, des commentaires ou des critiques. CountVectorizer prend ces documents et crée une matrice où chaque cellule indique le nombre d'occurrences d'un mot particulier dans un document spécifique.

Par exemple, si le mot "chat" apparaît trois fois dans le premier document, la première cellule de la première ligne de la matrice sera égale à 3. Cela permet de représenter chaque document sous forme de vecteur numérique, ce qui est utile pour l'analyse de texte et l'apprentissage automatique.

	about	all	cent	cents	money	new	old	one	two
doc	1	1	3	1	1	1	1	1	1

**In theory (a)**

↓

Index	0	1	2	3	4	5	6	7	8
doc	1	1	3	1	1	1	1	1	1

**In practice (b)**

Figure 22:CountVectorizer

## Présentation du projet réalisé (captures d'écran):

### 1. Importation des bibliothèques :

```
import streamlit as st
import pdfplumber
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics.pairwise import cosine_similarity
```

Figure 23: Importation des biblio ScreenShot

En important les bibliothèques streamlit, pdfPlumber et Scikit-learn, le code peut traiter les fichiers PDF, extraire le texte, et utiliser de méthodes de traitement du langage naturel (NLP) telles que la vectorisation de texte (CountVectorizer) et la mesure de similarité cosinus afin de créer l'application

### 2. Traitement des fichiers :

```
def process_files(job_description, resume):
    with pdfplumber.open(job_description) as pdf:
        job_description_text = " ".join([page.extract_text() for page in pdf.pages])

    with pdfplumber.open(resume) as pdf:
        resume_text = " ".join([page.extract_text() for page in pdf.pages])

    content = [job_description_text, resume_text]

    cv = CountVectorizer()
    matrix = cv.fit_transform(content)
    similarity_matrix = cosine_similarity(matrix)
    match = similarity_matrix[0][1] * 100

    return match, job_description_text, resume_text
```

Figure 24: Traitement des fichiers ScreenShot

La fonction `process_files` prend deux chemins de fichiers en entrée, `job_description` et `resume`, qui représentent respectivement une description de travail (job description) et un curriculum vitae (resume). La fonction effectue les étapes suivantes :

- **Extraction de texte** à partir du fichier de description de travail (`job_description`) : utilise la bibliothèque `pdfplumber` pour ouvrir le fichier PDF de la description de travail et le curriculum vitae (`resume`), et pour chaque page du PDF et le concatène pour former une chaîne de texte
- **Vectorisation de texte et Calcul de Similarité Cosinus** : Crée une liste contenant les textes de la description de travail et du curriculum vitae et utilise la classe `CountVectorizer` de `scikit-learn` pour convertir ces textes en une matrice de comptage de mots, et par la suite, utilise la fonction `cosine_similarity` de `scikit-learn` pour calculer la similarité cosinus entre les deux textes de manière à indiquer à quel point ils sont similaires. Dans ce contexte, il est utilisé pour mesurer la similarité entre le texte de la description de travail et le texte du curriculum vitae.
- **Le résultat de la similarité** cosinus est multiplié par 100 pour obtenir un pourcentage de correspondance. Ce pourcentage (`match`) indique à quel point le CV correspond à la description de travail, et ainsi retourne le pourcentage, le `job_description_text`, et le `resume_text`.

### 3. About Us/Contact Us :

Les fonctions `about_us` et `contact_us` ajoutent des fonctionnalités importantes à votre application. `about_us` crée une page présentant les créateurs, tandis que `contact_us` permet aux utilisateurs de vous envoyer des messages en utilisant un formulaire interactif. Ces éléments enrichissent l'expérience utilisateur en fournissant des informations sur les créateurs et en facilitant la communication directe.



```
def about_us():
    st.title("About Us")
    st.subheader("Welcome to the About Us page!")
    st.write("Meet the creators of this application.")
    creators_info = {
        "Youssef Mouhyeddine": "Co-founder and developer",
        "Anas Radi": "Co-founder and designer",
        "Hamza El Bouamri": "Co-founder and data scientist"
    }

    for creator, description in creators_info.items():
        st.write(f"**{creator}**: {description}")

def contact_us():
    st.title("Contact Us")
    st.subheader("Use the form below to get in touch with us.")
    name = st.text_input("Your Name:")
    email = st.text_input("Your Email:")
    message = st.text_area("Your Message:")

    if st.button("Submit"):
        st.success("Form submitted successfully! We'll get back to you soon.")
```

Figure 25: AboutUs et ContactUs Screenshots

#### 4. Fonction Principale :

La fonction main définit la logique principale de l'application. Elle utilise la bibliothèque Streamlit pour créer une barre de navigation latérale avec des options telles que "Home", "Upload CV", "About Us", et "Contact Us". La variable nav\_choice enregistre le choix de l'utilisateur parmi ces options, ce qui guide l'affichage de différentes parties de votre application en fonction de la sélection effectuée. Cela permet une navigation fluide entre les différentes sections de l'application.

```
def main():
    nav_choice = st.sidebar.radio("Navigation", ["Home", "Upload CV", "About Us", "Contact Us"])
```

Figure 26: Main Function ScreenShot



## 5. Extraction et affichage des résultats :

Cette partie du code s'exécute lorsque l'utilisateur clique sur le bouton "Process" après avoir téléchargé un fichier de description de travail (job description) au format PDF. Si un fichier a été téléchargé (uploadedJD est évalué à True), la fonction process\_files est appelée avec les fichiers de description de travail et de CV téléchargés

En fonction du pourcentage de correspondance calculé, un message de succès ou un avertissement est affiché. Si le pourcentage est supérieur ou égal à 70%, un message de félicitations est affiché, indiquant que le candidat est bien adapté au poste. Sinon, un avertissement s'affiche, suggérant à l'utilisateur de revoir les critères de correspondance.

```
if click and uploadedJD:
    job_description_match, job_description_text, resume_text = process_files(uploadedJD, uploadedResume)

    job_description_match = round(job_description_match, 2)

    st.write(f"Match Percentage: {job_description_match}%")

    if job_description_match >= 70:
        st.success("Congratulations! The candidate is a good match for the job.")
    else:
        st.warning("The candidate may not be an ideal match for the job. Consider reviewing the criteria.")

    st.subheader("Extracted Text:")
    st.write("Resume:")
    st.text(resume_text)
```

Figure 27: Extraction et Affichage ScreenShot

## Interfaces Réalisées :

### 1. Home Page :



*Figure 28:Home Page Interface*

CVCV, notre application web novatrice qui révolutionne l'analyse de CV et la mise en correspondance des offres d'emploi.

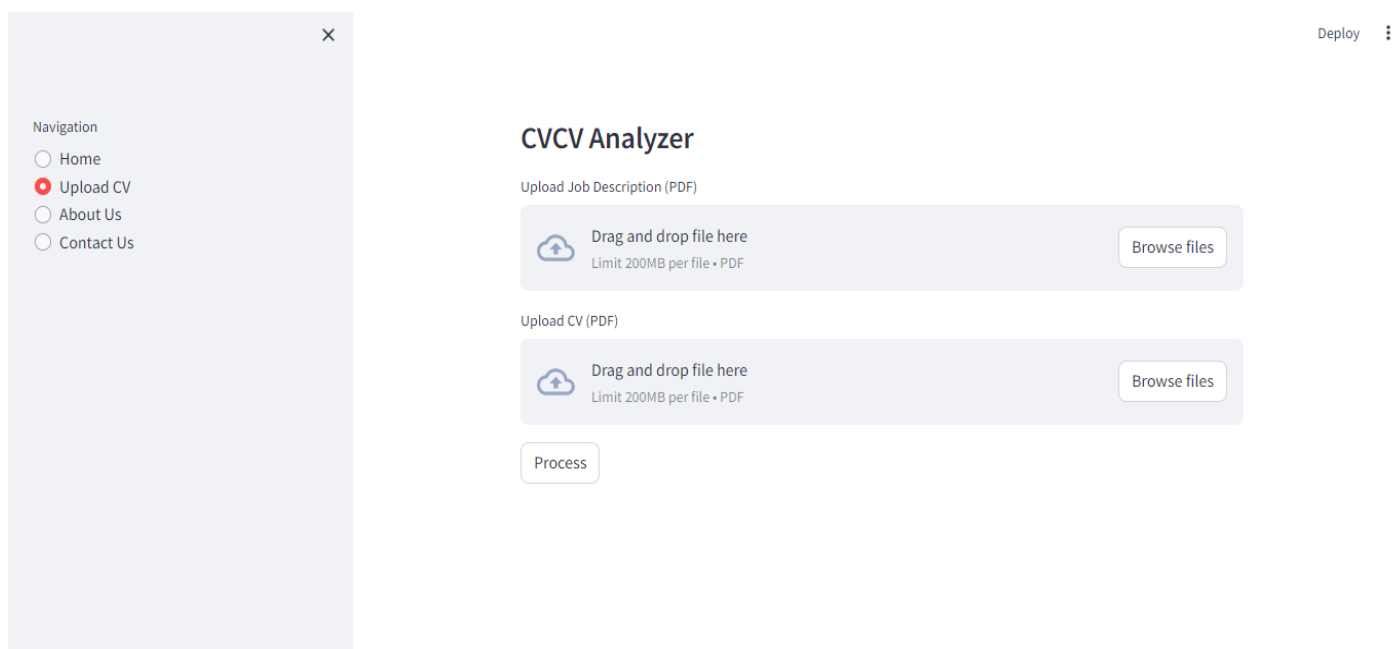
Le Sidebar intuitive permet une exploration aisée des différentes sections telles que "Home" (sur lequel on est déjà positionné), "Upload CV", "About Us" et "Contact Us".

La section "Home" sert de porte d'entrée captivante aux capacités de CVCV, positionnée comme un projet éducatif recommandé par notre école, elle intègre une image captivante, renforçant l'attrait visuel de la plateforme pour une expérience moderne et esthétiquement plaisante.

CVCV incarne la fusion entre la technologie et l'éducation, guidant les utilisateurs dans leur exploration des opportunités de carrière avec une plateforme sophistiquée et accessible pour l'analyse de CV et la mise en correspondance d'emplois.

Embarquez dans un voyage transformateur avec CVCV - là où les CV rencontrent leur correspondance idéale et où les chemins de carrière se dévoilent avec précision.

## 2. Upload CVs :



*Figure 29: Upload CV Interface*

La fonctionnalité "Upload CV" propose une expérience utilisateur fluide et intuitive pour simplifier le processus de soumission de candidature.

Avec un design convivial, les utilisateurs ont la possibilité de télécharger leur CV en quelques clics ou de le faire glisser directement sur la page grâce à la méthode "Drag and Drop" qui simplifie l'expérience utilisateur en permettant de déplacer les PDF à travers l'interface simplement en les saisissant avec la souris, en les déplaçant vers un nouvel emplacement, et en les relâchant.

Grâce à cette interface interactive, le processus de soumission de CV devient rapide et accessible, permettant aux utilisateurs de se concentrer sur l'essentiel de leur candidature. Que ce soit sur ordinateur ou sur appareil mobile, cette fonction de téléchargement intuitif assure une expérience utilisateur optimale, facilitant ainsi le premier pas vers de nouvelles opportunités professionnelles.

## 2.1. Sélection de Job Description :

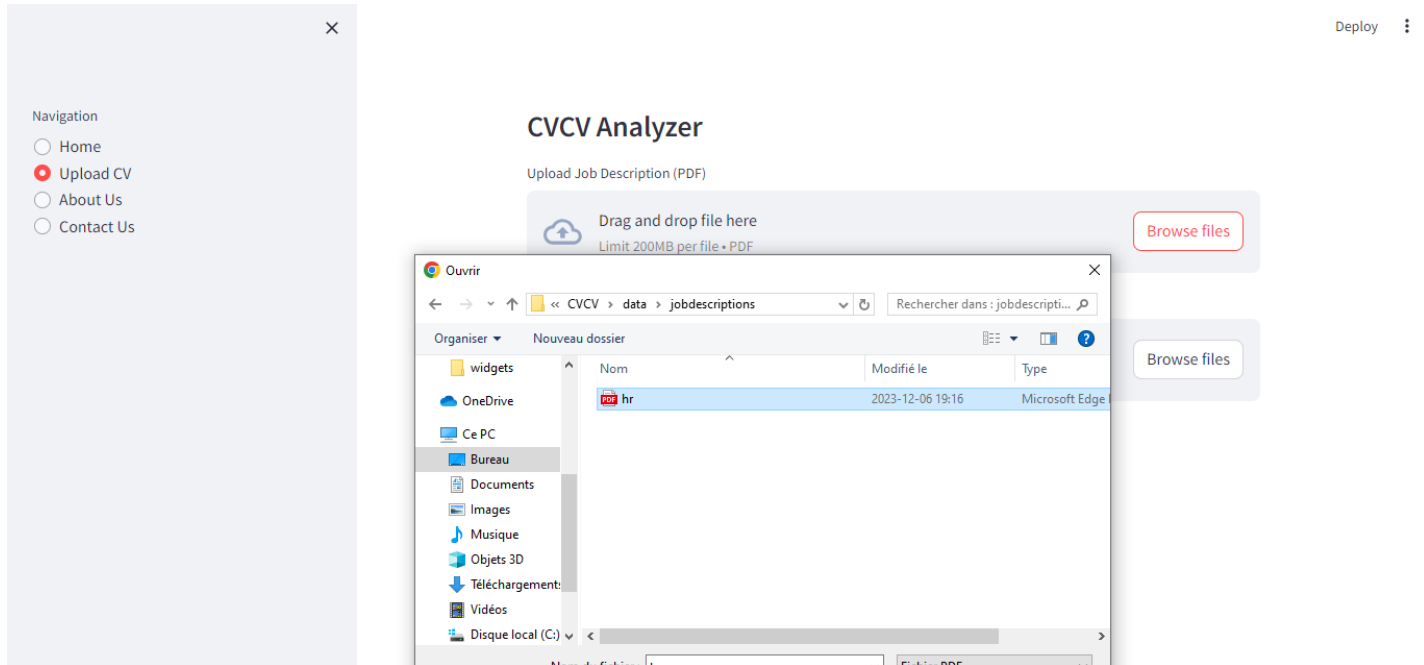


Figure 30: Job Description Selection Interface

L'Utilisateur ajoute la Job Description souhaitée

## 2.2. Sélection d'un CV :

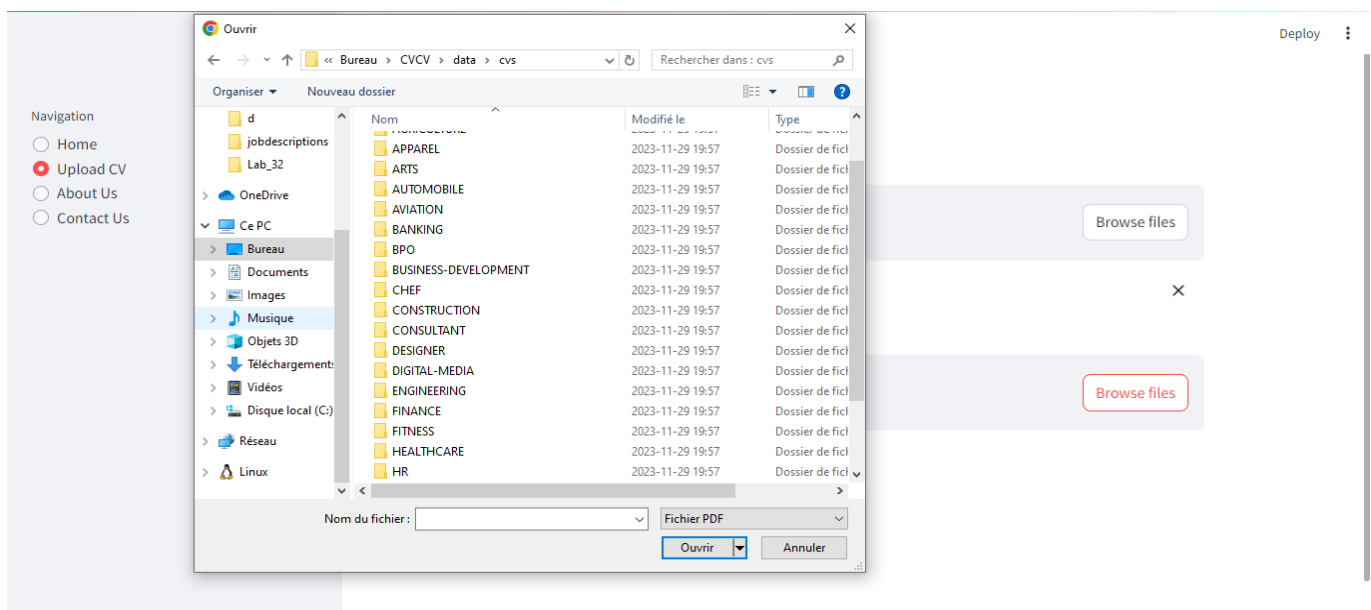


Figure 31: CV Selection Interface

Même chose que la sélection des Jobs Descriptions, il peut sélectionner le CV sur lequel il souhaite tester la correspondance.

## 2.3. Résultat :

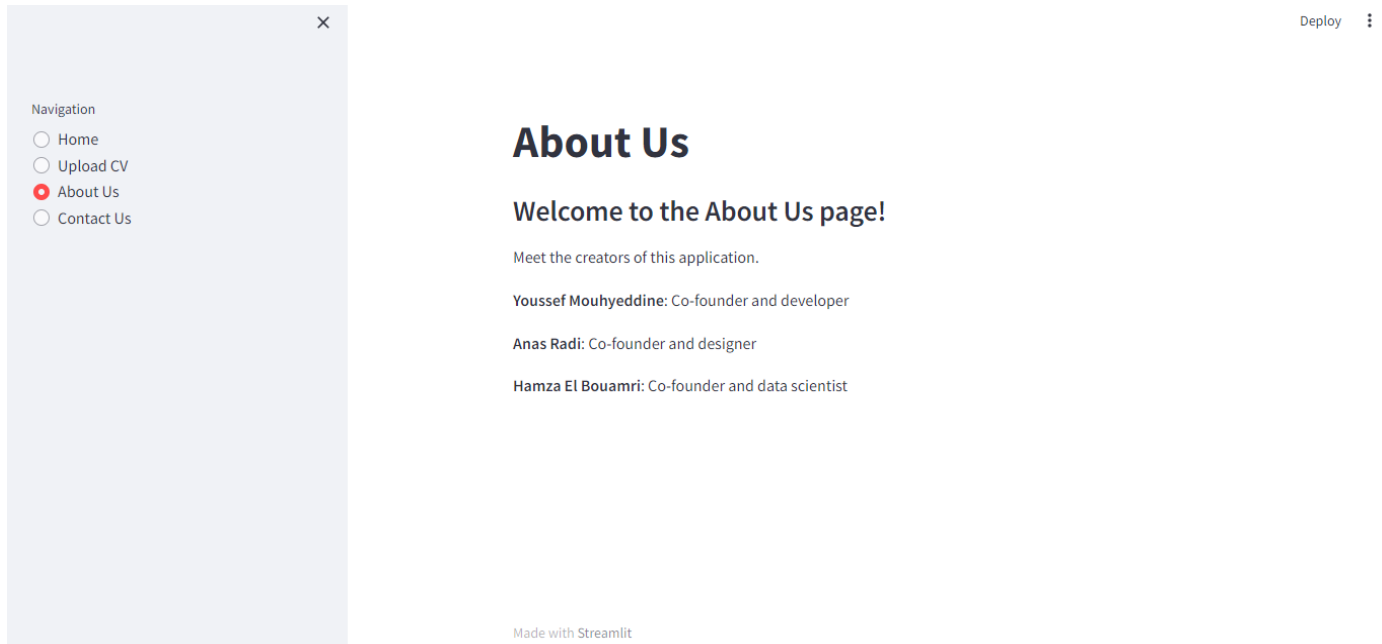


Figure 32: Résultat Upload CV Interface

Dans cette interface soigneusement conçue, nous avons mis l'accent sur la convivialité en regroupant de manière intuitive toutes les fonctionnalités essentielles pour les utilisateurs, ainsi que les résultats issus des analyses de nos modèles sophistiqués. L'objectif premier est de simplifier le processus d'évaluation des CV en fournissant un retour clair et immédiat sur la correspondance entre le CV téléchargé et la description de poste.

Lorsque les résultats dépassent le seuil critique de 70 %, nous indiquons avec enthousiasme que le candidat représente une correspondance significative avec les exigences du poste. Cependant, si les résultats sont en deçà de ce seuil, cela signale que le CV téléchargé ne correspond pas pleinement à la description de poste ou aux besoins de l'entreprise. Dans cette situation, nous recommandons une révision attentive des critères pour garantir une adéquation optimale entre le candidat et le poste à pourvoir. Cette approche directe et informative vise à offrir une expérience transparente et utile aux utilisateurs, les guidant efficacement dans leurs décisions de recrutement.

### 3. About Us :



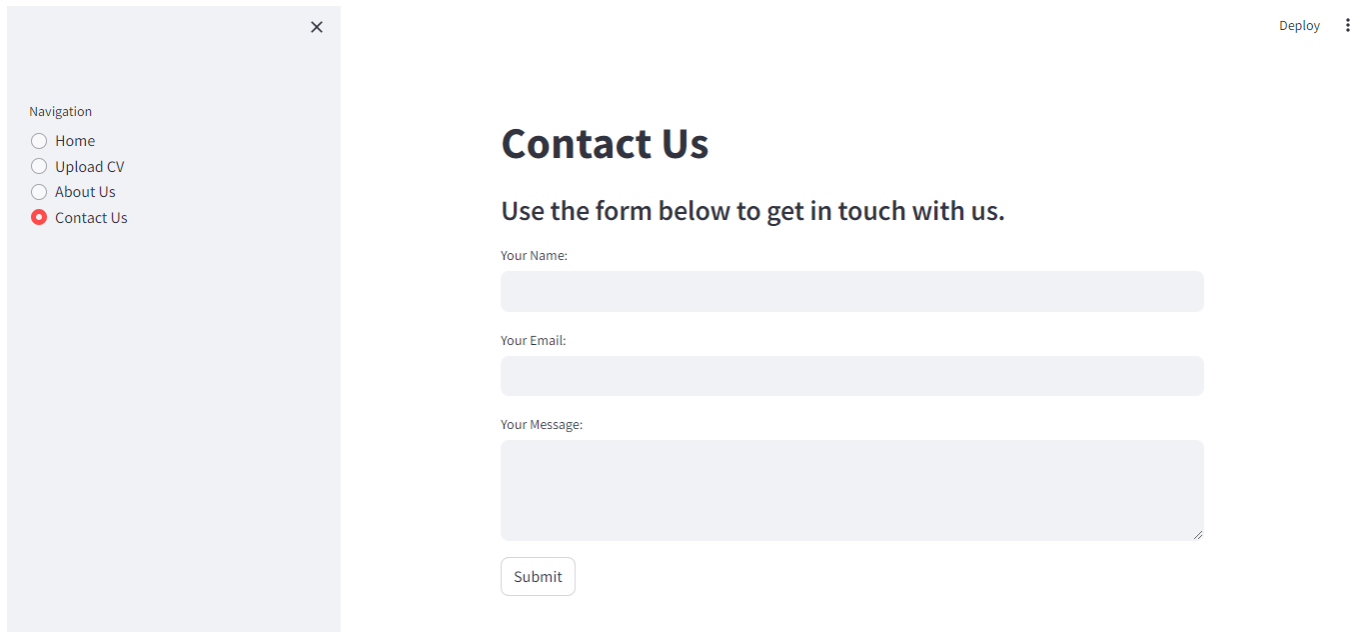
*Figure 33: About Us Interface*

Dans cette interface dédiée à la page "About Us", nous avons pris le soin de fournir aux utilisateurs les individus qui ont contribué à faire de ce projet une réalité grâce à leur dévouement et leur persévérance.

Nous présentons brièvement notre équipe créative et talentueuse, les esprits derrière cette application innovante. Les contributeurs, dont Youssef Mouhyeddine, co-fondateur et développeur, Anas Radi, co-fondateur et designer, ainsi que Hamza El Bouamri, co-fondateur et data scientist, sont présentés individuellement avec leurs rôles respectifs au sein du projet.

Cette approche vise à établir une connexion humaine avec les utilisateurs, en mettant en avant le travail collaboratif qui a donné vie à cette application.

## 4. Contact Us :



*Figure 34:Contact Us Interface*

À travers cette interface dédiée à la communication directe avec nos utilisateurs, l'expérience de contact est rendue à la fois simple et efficace. En utilisant le formulaire en ligne intégré, l'utilisateur peut aisément sélectionner son nom, renseigner son email, et transmettre son message. Cette approche épurée vise à éliminer toute complication inutile, offrant ainsi un moyen direct et rapide de nous contacter. Il vous suffit de remplir quelques détails, de rédiger votre message, puis de cliquer sur "SUMBIT". Notre processus de réponse rapide garantit que vos questions et retours seront traités promptement.

## Conclusion

En conclusion, la réalisation du projet CVCV a été fructueuse, combinant habilement des bibliothèques Python telles que Streamlit et scikit-learn. L'interface interactive a été développée avec Streamlit, et des techniques avancées d'apprentissage automatique ont été appliquées. Les choix technologiques, y compris l'utilisation de PyCharm et Lucidchart, ont facilité la collaboration et la planification. Les captures d'écran confirment le succès des tests, démontrant la fiabilité de l'application dans la correspondance entre les offres d'emploi et les CV. Un pas important vers une solution performante

# Conclusion Générale et perspective :

*Ce projet de fin d'année incarne la culmination harmonieuse des compétences et de la collaboration d'une équipe trinôme, symbolisant le pinacle de notre parcours académique. Nous avons plongé dans les méandres de l'innovation technologique appliquée à la sélection des candidatures, mettant particulièrement l'accent sur l'analyse de la concordance entre les descriptions de poste et les CV des candidats.*

*Grâce à l'exploitation de techniques avancées de traitement du langage naturel telles que Streamlit, pdfplumber et scikit-learn, l'application intègre une interface utilisateur interactive et applique des techniques avancées d'apprentissage automatique, et donc, notre équipe s'est employée à affiner la correspondance entre les attentes des recruteurs et les compétences des candidats. Ce faisant, nous avons simplifié la tâche complexe de la sélection, offrant une réponse technologique à un défi critique du processus de recrutement.*



# Webographie :

- <https://hive.blog/hive-139531/@geekgirl/pdfplumber-pdf-data-extracting-library-for-python-gets-a-big-update>
- <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1445304-python-definition-et-utilisation-de-ce-langage-informatique/>
- [https://en.m.wikipedia.org/wiki/File:Scikit\\_learn\\_logo\\_small.svg](https://en.m.wikipedia.org/wiki/File:Scikit_learn_logo_small.svg)
- [https://commons.wikimedia.org/wiki/File:NumPy\\_logo.svg](https://commons.wikimedia.org/wiki/File:NumPy_logo.svg)
- <https://www.jumpingrivers.com/blog/customising-matplotlib/>
- [https://en.wikipedia.org/wiki/File:Lucidchart\\_logo\\_%28September\\_2021%29.svg](https://en.wikipedia.org/wiki/File:Lucidchart_logo_%28September_2021%29.svg)
- <https://streamlit.io/brand>
- <https://1000logos.net/python-logo/>
- <https://lucid.app/documents>
- [https://en.m.wikipedia.org/wiki/File:PyCharm\\_Icon.svg](https://en.m.wikipedia.org/wiki/File:PyCharm_Icon.svg)
- <https://www.atlassian.com>

