



université
virtuelle
Burkina ★ Faso

Licence : Génie Logiciel

Rapport de projet

(Système de détection de plagiat)

Réalisé par :
PAFADNAM Ibrahim

Sommaire

1. Introduction
2. Cadre du travail
3. Analyse et spécification des besoins
4. Conception UML
5. Développement de l'application
6. Difficulté et solutions apportés
7. Conclusion
1. [Lien vers le code source](#)

Introduction

Au cours de notre parcours universitaire en Licence 3 Génie Logiciel, nous avons acquis de nombreuses connaissances en programmation, développement web, machine learning, etc. Ce rapport présente le travail réalisé dans le cadre de notre projet de fin d'études, qui porte sur le développement d'un système de détection de plagiat.

Pour rendre compte de notre projet, nous allons vous exposer ses principaux axes en suivant le plan indiqué ci-dessus.

Cadre du travail

Problématique : Le plagiat, de plus en plus répandu en raison de l'accès facile aux ressources numériques, est difficile à détecter manuellement dans les milieux académiques et professionnels, nécessitant une solution automatisée.

Solution proposée : La solution proposée est un **Système de Détection de Plagiat** basé sur **Flask** (backend) et **PostgreSQL** (base de données), capable de comparer des fichiers soumis par l'utilisateur à des documents existants et des sources en ligne

Analyse et spécification des besoins

Objectif de l'application :

L'objectif de ce système est de permettre aux enseignants d'analyser les devoirs de leurs étudiants afin d'évaluer la

similarité des documents, et ainsi détecter les éventuels cas de plagiat.

Besoins :

Soumission de fichiers : Permettre aux utilisateurs de télécharger des fichiers aux formats **PDF**, **DOCX** et **TXT**.

Extraction et traitement du texte : en utilisant NLTK

Analyse de similarité :

- Comparaison inter-documents : Lors de la soumission de plusieurs fichiers, le système doit comparer leur similarité à l'aide d'un modèle pré-entraîné (Universal Sentence Encoder).
- Recherche en ligne : Si un seul fichier est soumis, il doit être comparé aux ressources disponibles en ligne via une API de recherche.

Rapport de plagiat : Le système doit générer un rapport détaillé indiquant le pourcentage de similarité, les sections plagiées et les liens vers les sources détectées

Stockage des résultats : Les résultats des analyses doivent être enregistrés dans une base de données pour un suivi futur.

Interface utilisateur moderne : Utilisation de **Bootstrap** et **AJAX** pour créer une interface intuitive et réactive, facilitant l'interaction avec le système.

Diagrammes UML

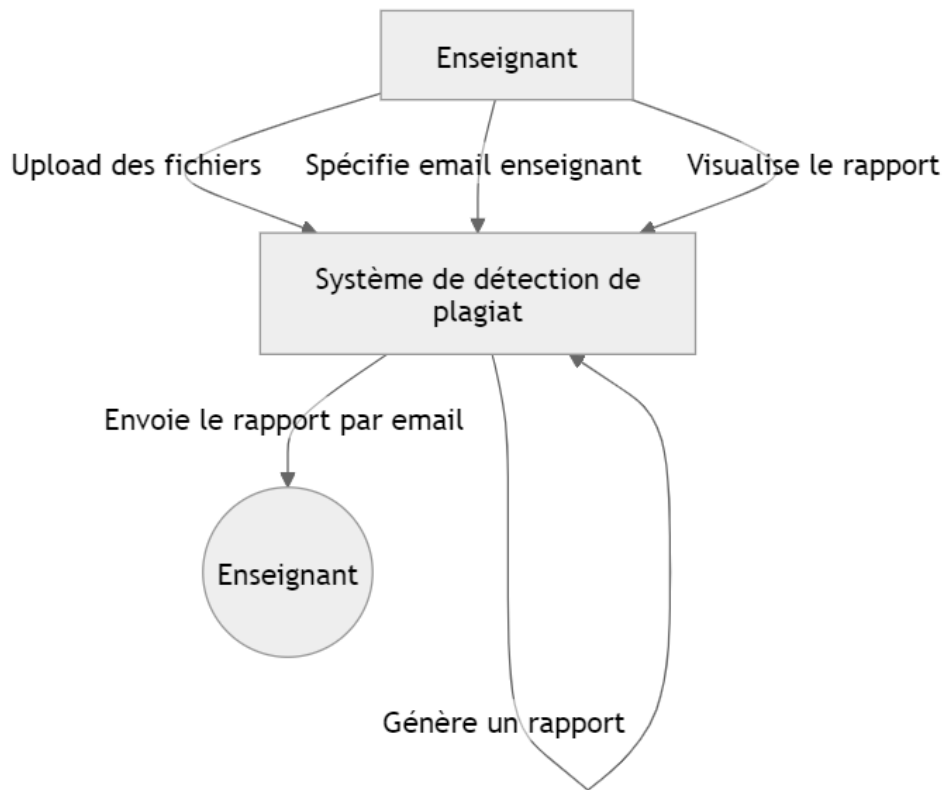


Figure 2 Diagramme de cas d'utilisation

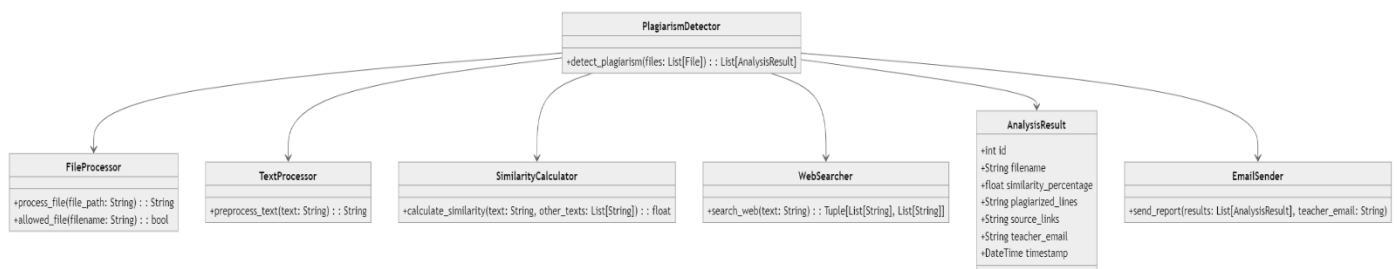


Figure 1 Diagramme de classe

Spécifications logicielles

Backend	Flask
Frontend	Bootstrap,html,js, AJAX
API	Google Search Custom
Machine learning	TensorFlow
Traitement du texte	NLTK
Modèle pour calculer la similarité	Universal Sentence Encoder
Base de données	PostgreSQL (dans Render cloud)
Gestion de versions décentralisé	Git, github

Présentation de l'application

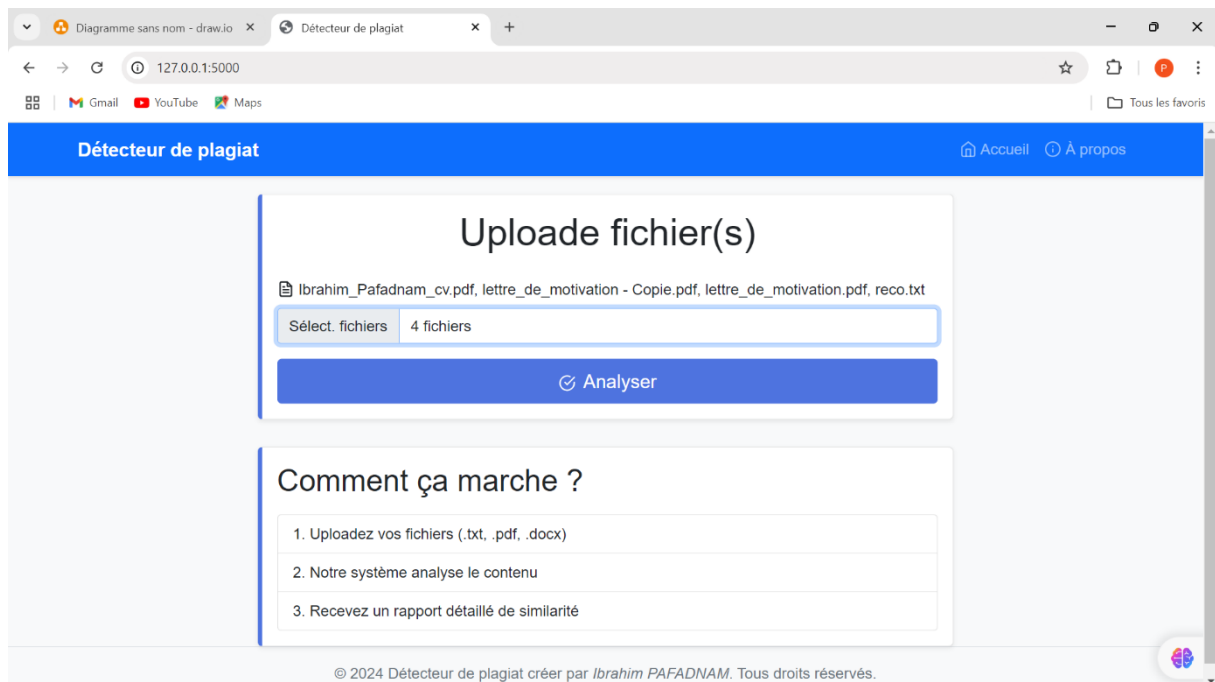


Figure 3 La page d'accueil

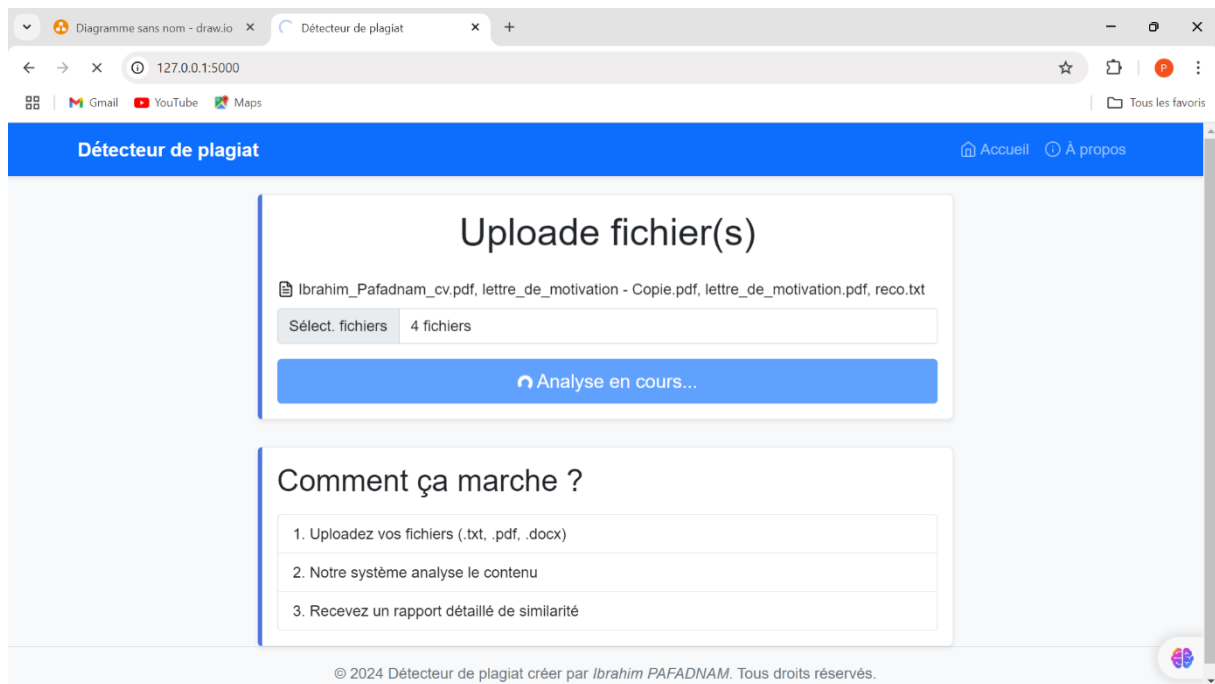


Figure 5 L'analyse en cours

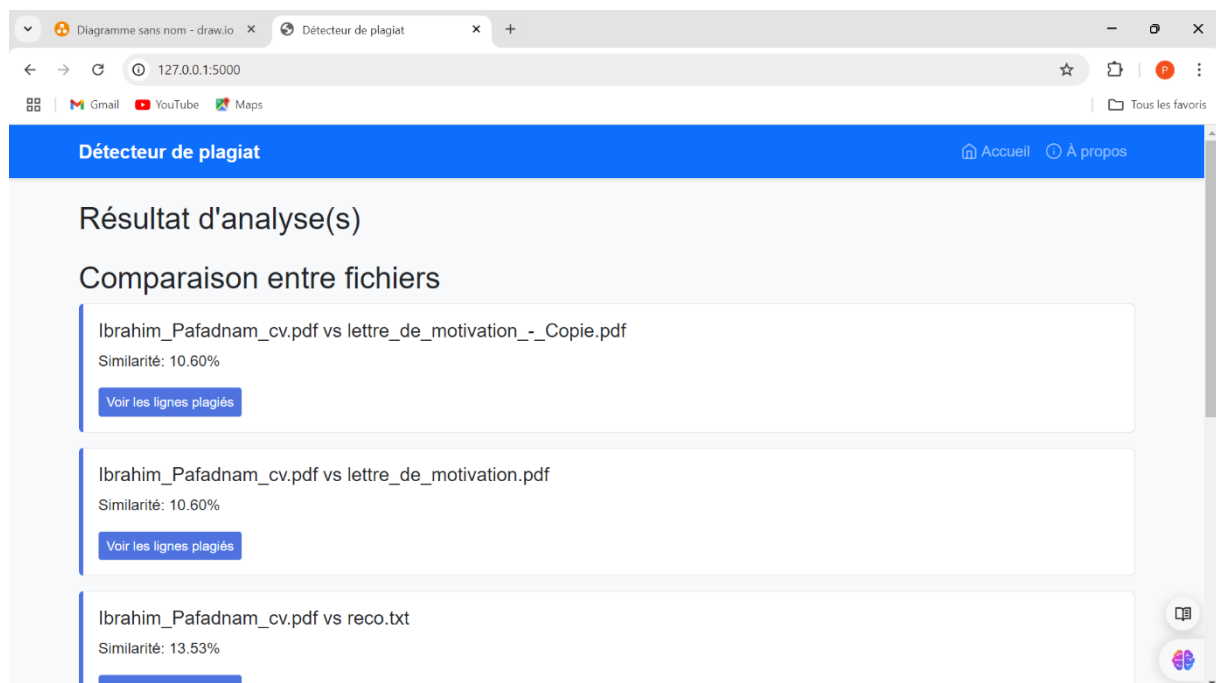


Figure 4 Rapport d'analyse(s)

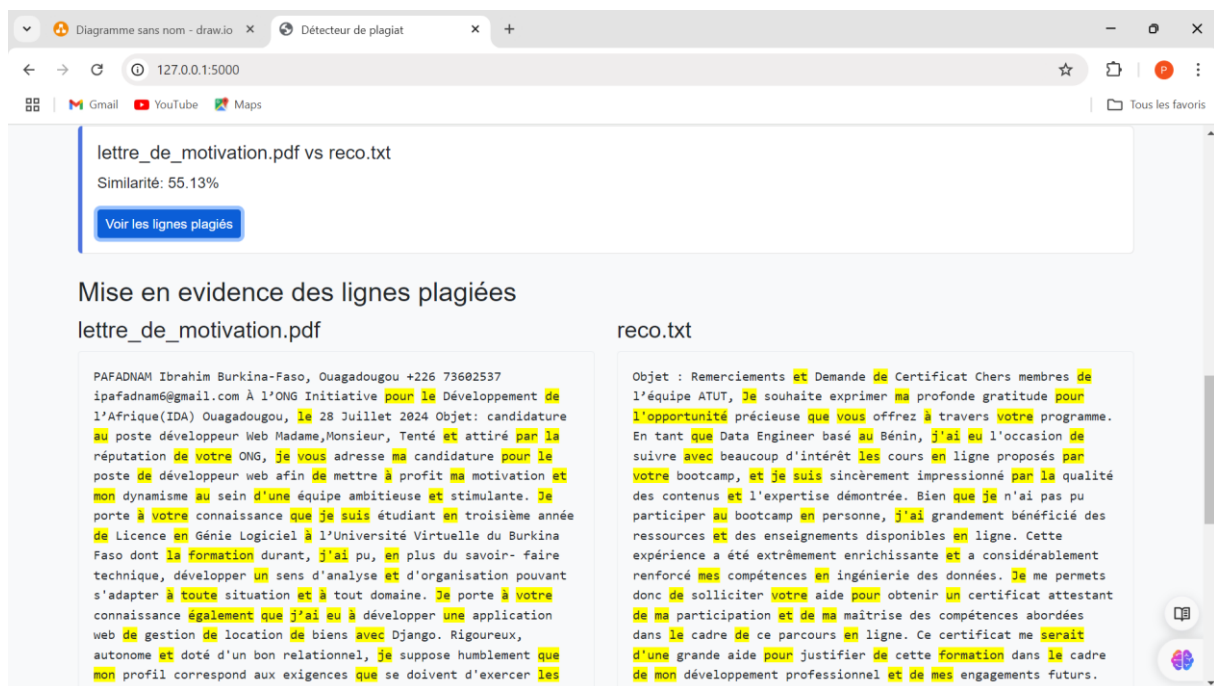


Figure 6 les lignes plagiées

Difficultés et solutions apportées

Le développement du système de détection de plagiat a été contraint par des délais serrés, ce qui a rendu difficile l'implémentation complète de toutes les fonctionnalités prévues. En particulier, certaines fonctionnalités avancées, comme l'intégration complète de la recherche en ligne via l'API Google Custom Search, l'envoi du rapport par email n'ont pas pu être finalisées et aussi l'hébergement.

Pour faire face à ces contraintes, j'ai priorisé les fonctionnalités essentielles comme l'extraction de texte, le prétraitement linguistique, et la comparaison entre documents. Cela permet au système de fonctionner de manière basique et d'offrir un premier niveau de détection de plagiat. Néanmoins la plateforme reste ouverte à des améliorations futures

Conclusion

Nous tenons à vous remercier pour cette initiative qui nous a permis de travailler sur un projet concret et utile pour la société dans le domaine de l'éducation.

Ce système de détection de plagiat développé répond aux exigences initiales et offre une solution robuste et flexible. Il constitue une base solide pour la détection de plagiat dans un contexte éducatif ou professionnel. Avec les améliorations suggérées, il a le potentiel de devenir un outil encore plus puissant et polyvalent.

Lien vers le code source :

<https://github.com/IbrahimPafadnam/detecteur-plagiat>