



Présentation PFE

Julien Romero

Objectifs du Projet

- Concevoir une application web complète avec un focus sur l'intelligence artificielle.
- Répondre à un besoin concret en simulant une utilisation professionnelle.
- Intégrer plusieurs compétences : traitement des données, apprentissage automatique, interfaces web, MLOps.

Composantes Clés du Projet

1. **Pipeline de données** : Extraction, transformation, et préparation des données pour l'entraînement.
2. **Système de stockage** : Gestion des données générées par l'application.
3. **Simulation de données** : Génération de données réalistes pour les tests.
4. **Modèles d'IA** :
 - a. Modèle de machine learning simple (classification, régression, etc.).
 - b. Modèle de deep learning (préentraîné ou entraîné).
5. **Interface web** : Accès aux fonctionnalités pour les utilisateurs.
6. **API** : Interface d'exploitation des modèles pour d'autres applications.
7. **Pipeline CI/CD et MLOps** : Automatisation des tests et déploiement.
8. **Tableau de bord** : Visualisation des performances et contrôle de l'application.

Organisation du Projet

- Travail en groupe de 2-3 étudiants.
- Livrables principaux :
 - Soutenance intermédiaire (30% de la note).
 - Soutenance finale (40% de la note).
 - Rapport écrit (30% de la note).

Planning Approximatif

Mois 1 : Planification et Préparation

- Comprendre le sujet, choisir un projet.
- Rechercher des datasets pertinents.
- Concevoir l'architecture initiale (pipeline, modèles, stockage).

Planning Approximatif

Mois 2 : Développement des Composants Fondamentaux

- Implémenter le pipeline de données et les simulations.
- Développer et entraîner le modèle de machine learning.

Planning Approximatif

Mois 3 : Développement Avancé et Intégration

- Soutenance intermédiaire pour retour et ajustements.
- Intégrer le modèle de deep learning.
- Développer l'interface web et l'API.

Planning Approximatif

Mois 4 : Finalisation et Livraison

- Mettre en place le CI/CD et MLOps.
- Développer le tableau de bord et finaliser les tests.
- Soutenance finale et remise du rapport.

Livrables et Évaluation

1. Soutenance Intermédiaire (30%)

- a. Objectif : Montrer les progrès et obtenir des retours.
- b. Contenu :
 - i. Progrès sur la pipeline de données et le ML.
 - ii. Aperçu de l'architecture globale.

2. Soutenance Finale (40%)

- a. Objectif : Démontrer une solution fonctionnelle.
- b. Contenu :
 - i. Application complète avec CI/CD, interface web, tableau de bord.
 - ii. Discussion sur les choix techniques et les améliorations possibles.

3. Rapport Écrit (30%)

- a. Contenu :
 - i. Description du projet et des besoins répondus.
 - ii. Détails techniques (modèles, pipeline, architecture).
 - iii. Analyse des performances et leçons apprises.

Exemples de sujets

Assistant virtuel pour le recrutement

- *Données collectées* : CVs, profils LinkedIn, évaluations d'entretiens.
- *Datasets* : Utilisation de Kaggle pour des datasets anonymisés.
- *Clients potentiels* : Agences de recrutement, entreprises en quête de talents.
- *Modèles* : Classification des candidats (ML) et analyse du langage naturel (DL).

Exemples de sujets

Générateur automatique de contenus pédagogiques

- *Données collectées* : Ressources éducatives existantes, feedbacks d'utilisateurs.
- *Datasets* : Open Educational Resources (OER).
- *Clients potentiels* : Écoles, plateformes e-learning.
- *Modèles* : Recommandation personnalisée (ML) et génération automatique de texte (DL).

Exemples de sujets

Outil de détection d'erreurs sur chaînes de production

- *Données collectées* : Images des produits, historiques de production, dessins industriels.
- *Datasets* : À créer si besoin.
- *Clients potentiels* : Industries manufacturières.
- *Modèles* : Classification d'images (ML) et détection d'anomalies (DL).

Exemples de sujets

Plateforme d'évaluation automatique des compétences techniques

- *Données collectées* : Résultats d'examens, évaluations par pairs.
- *Datasets* : Données anonymisées provenant d'universités ou plateformes MOOC.
- *Clients potentiels* : Universités, entreprises souhaitant évaluer leurs employés.
- *Modèles* : Systèmes de notation automatisés (ML) et analyse sémantique (DL).

Exemples de sujets

Plateforme d'apprentissage d'une langue

- *Données collectées* : Réponses à des questions.
- *Datasets* : Duolingo dataset.
- *Clients potentiels* : Personnes voulant apprendre une nouvelle langue.
- *Modèles* : Système de recommandation de problème (ML) et analyse textuelle (DL).

Exemples de sujets

Plateforme collaborative pour la gestion de projets open source

- ***Données collectées*** : Contributions utilisateurs, feedbacks sur projets existants.
- ***Datasets*** : Données provenant de GitHub ou autres plateformes open source publiques.
- ***Clients potentiels*** : Développeurs open source, entreprises cherchant à collaborer sur projets communs.
- ***Modèles*** : Systèmes recommandation pour contributions pertinentes (ML) et analyse de code (DL).

Exemples de sujets

Générateur d'histoires interactives pour enfants

- **Données collectées** : Choix des utilisateurs, préférences de genre, vocabulaire souhaité.
- **Datasets** : Livres pour enfants disponibles sur des plateformes comme Project Gutenberg.
- **Clients potentiels** : Écoles, parents, plateformes d'éducation en ligne.
- **Modèles** : Génération de texte pour créer des histoires personnalisées (DL) et suggestions de vocabulaire (ML). L'application permet aux enfants de choisir des personnages et des thèmes, tout en apprenant de nouveaux mots au fur et à mesure.

Exemples de sujets

Assistant virtuel pour l'écriture collaborative d'histoires

- **Données collectées** : Contributions des utilisateurs, styles d'écriture, thèmes préférés.
- **Datasets** : Textes littéraires classiques et contemporains pour l'inspiration.
- **Clients potentiels** : Écrivains en herbe, clubs de lecture, écoles.
- **Modèles** : Analyse de style d'écriture (ML) et génération de suggestions narratives (DL). Les utilisateurs peuvent co-créeer des récits avec l'assistance de l'IA.

Exemples de sujets

Application d'apprentissage du dessin avec feedback instantané

- **Description** : Une application qui guide les utilisateurs à travers des leçons de dessin étape par étape, fournissant un retour en temps réel sur leurs traits et techniques.
- **Données collectées** : Dessins des utilisateurs, progrès au fil du temps.
- **Clients potentiels** : Étudiants, artistes en herbe, écoles d'art.
- **Modèles** : Reconnaissance d'image pour évaluer les dessins (DL) et suggestions d'amélioration (ML).

Exemples de sujets

Outil d'automatisation de la saisie des factures

- **Description** : Une application qui utilise l'IA pour extraire automatiquement les données des factures et les saisir dans le logiciel de comptabilité, réduisant ainsi les erreurs humaines.
- **Données collectées** : Factures reçues, corrections apportées par les utilisateurs.
- **Clients potentiels** : PME, départements comptables d'entreprises.
- **Modèles** : Reconnaissance optique de caractères (OCR) et apprentissage automatique pour améliorer l'extraction.