

# Cahier des Charges

## Analyse Prédictive pour l'Accès au Crédit en Afrique de l'Ouest

### 1 Contexte

L'accès au financement constitue un levier crucial pour le développement des petites et moyennes entreprises (PME) ainsi que du secteur agricole en Afrique de l'Ouest. Toutefois, l'absence d'évaluation objective de la solvabilité constitue un frein pour l'accès au crédit, en particulier pour les entrepreneurs non bancarisés.

Ce projet vise à concevoir un système intelligent capable d'évaluer automatiquement la solvabilité d'un demandeur de crédit, afin de soutenir l'inclusion financière dans la région.

### 2 Objectif du Projet

Développer un modèle d'intelligence artificielle permettant de prédire la solvabilité des petites entreprises et des agriculteurs, à partir de leurs données économiques et financières, et de proposer une API accessible aux institutions de microfinance pour tester un profil emprunteur.

### 3 Périmètre du Projet

- Collecte et traitement de données économiques, financières et sectorielles.
- Conception et entraînement d'un modèle de classification binaire (solvable / non solvable).
- Intégration du modèle dans une API web accessible.
- Documentation technique et rapport d'analyse.

### 4 Technologies et Outils

- **Langage** : Python
- **Librairies** : Pandas, NumPy, Scikit-learn, TensorFlow ou PyTorch, Matplotlib, Seaborn
- **Base de données** : PostgreSQL ou MongoDB
- **API** : FastAPI
- **Versioning** : Git/GitHub

## 5 Livrables Attendus

- Jeu de données nettoyé et prêt à l'emploi
- Rapport d'analyse exploratoire (EDA) et visualisations
- Modèle de machine learning entraîné et évalué
- Script d'exportation du modèle (pickle ou joblib)
- API REST permettant de tester la solvabilité d'un utilisateur via une requête JSON
- Rapport final du projet (PDF)

## 6 Méthodologie

### 6.1 Étape 1 : Collecte et Préparation des Données

- Collecte ou simulation de données économiques et financières
- Nettoyage, traitement des valeurs manquantes et encodage
- Normalisation ou standardisation des variables numériques

### 6.2 Étape 2 : Analyse Exploratoire

- Analyse statistique des variables
- Visualisation des corrélations et des distributions
- Analyse de la balance des classes

### 6.3 Étape 3 : Modélisation

- Sélection des variables pertinentes
- Entraînement de plusieurs modèles (régression logistique, arbres, forêts, réseaux de neurones)
- Évaluation à l'aide de métriques : précision, rappel, F1-score, AUC-ROC
- Interprétabilité du modèle (SHAP, LIME)

### 6.4 Étape 4 : Intégration et Déploiement

- Création d'une API FastAPI avec un endpoint /predict
- Test local ou hébergement léger
- Documentation de l'API (Swagger UI)

### 6.5 Étape 5 : Rédaction du Rapport

- Description des étapes, choix méthodologiques et résultats
- Présentation des performances du modèle
- Capture d'écrans ou exemples d'utilisation de l'API

## 7 Impact Attendu

- Réduction du risque pour les institutions financières
- Inclusion financière de populations non couvertes par les systèmes bancaires classiques
- Base technologique réutilisable dans d'autres contextes géographiques similaires