

Rapport de Projet

Introduction

Ce projet vise à analyser des données agricoles et écologiques pour fournir des recommandations sur les cultures et d'améliorer la compréhension des dynamiques saisonnières et de prédire les besoins agricoles, notamment en termes de gestion de l'eau et des nutriments, en fonction des conditions environnementales et des données satellitaires. En utilisant des outils tels que la régression linéaire, le clustering et l'analyse en composantes principales (ACP), le projet exploite des données sur les sols, les indices spectraux, et les mesures climatiques.

Partie 1 : Analyse des Données Agricoles

Objectif

L'objectif de cette étape est de comprendre les caractéristiques des données agricoles à travers des analyses exploratoires et des visualisations.

Déroulement

1. Exploration des Données :

- Importation et préparation du fichier agriculture_data_required.csv.
- Identification des moyennes générales et spécifiques aux différentes cultures pour des éléments comme l'azote (N), le phosphore (P), le potassium (K), l'humidité, et le pH.

2. Visualisation :

- Distribution des variables via des graphiques.
- Analyse des cultures ayant des besoins au-dessus ou en-dessous des moyennes globales.

3. Clustering :

- Application de K-Means pour regrouper les cultures en clusters basés sur leurs caractéristiques.
- Détermination du meilleur nombre de clusters via la méthode du coude.

4. Modélisation Prédictive :

- Utilisation de la régression logistique pour prédire les cultures adaptées à des conditions données.

Résultats

- Des schémas saisonniers émergent pour certaines cultures en fonction de la température et de l'humidité.
- Les seuils critiques pour chaque variable ont été identifiés, fournissant une base pour des recommandations agricoles.

- Les clusters de cultures montrent des regroupements naturels basés sur les caractéristiques du sol et les besoins climatiques.
- les résultats sont dans le script bien visualiser et l application du model

Partie 2 : Analyse des Images Satellitaires

Objectif

Collecter, prétraiter et analyser les indices spectraux à partir d'images Sentinel-2 pour détecter des variations environnementales sur plusieurs saisons.

Déroulement

1. **Extraction des Données Satellitaires :**
 - Accès à des images Sentinel-2
 - Filtrage des images avec une couverture nuageuse minimale entre juillet 2020 et décembre 2023..
 - Calcul des indices spectraux tels que NDVI, NDWI, et MNDWI.
2. **Nettoyage des données :** Traitement des valeurs manquantes et suppression des anomalies.
3. **Analyse des tendances :** Étude des variations mensuelles des indices via des visualisations comme les boxplots et heatmaps.
4. **Clustering des Mois :**
 - Utilisation de l'ACP et du clustering pour regrouper les mois selon leurs caractéristiques climatiques.
 - Identification des mois représentatifs pour chaque cluster.

Résultats

- Les indices montrent des tendances claires en fonction des saisons :
 - NDVI et NDWI atteignent des pics au printemps.
 - Des baisses significatives en été reflètent le stress hydrique.
- Les clusters temporels mettent en évidence des périodes critiques pour la gestion agricole.
- Les clusters identifiés sont les suivants :
 - Cluster 0 : Mois [1, 2, 3, 12] avec le mois le plus proche du centroïde étant le mois 2.
 - Cluster 1 : Mois [4, 5, 6, 9, 10, 11] avec le mois le plus proche du centroïde étant le mois 5.
 - Cluster 2 : Mois [7, 8] avec le mois le plus proche du centroïde étant le mois 7.

Partie 3 : Fusion des Données Terrain et Spectrales

Objectif

Fusionner les données de terrain avec les indices spectraux pour modéliser et comprendre les interactions entre les variables environnementales et spectrales.

Déroulement

1. **Chargement des données terrain** : Nettoyage et agrégation des mesures issues du fichier terrain_N045_or.csv.
2. **Extraction des pixels spectraux** : Utilisation des données Sentinel-2 pour extraire les valeurs des indices par pixel dans la zone d'intérêt définie par polygon3.geojson.
3. **Données Spatiales** :
 - o Utilisation d'un fichier GeoJSON pour définir la région d'intérêt (ROI).
 - o Extraction des indices spectraux au niveau des pixels de la ROI.
4. **Fusion des données** : Correspondance des données de terrain et spectrales basée sur une grille de 10 mètres pour une résolution commune..
5. **Analyse de corrélation** : Étude des relations entre les indices spectraux et les variables environnementales.
6. **Modèles de Régression** :
 - o Application de modèles pour prédire les éléments nutritifs (N, P, K) et d'autres variables environnementales (température, pH, humidité).

Résultats

- Une forte corrélation a été observée entre NDVI et les nutriments du sol (N, P, K).
- NDWI présente une relation claire avec l'humidité, confirmant son utilité pour détecter les zones stressées par le manque d'eau.
- Les clusters de mois mettent en évidence des périodes optimales pour chaque indicateur.

Partie 4 : Modélisation et Prédictions

Objectif

Construire des modèles de régression linéaire pour prédire les besoins agricoles en fonction des indices spectraux et des données environnementales.

Déroulement

1. **Sélection des mois représentatifs** :
 - o Cluster 0 : Mois 2
 - o Cluster 1 : Mois 5
 - o Cluster 2 : Mois 7
2. **Préparation des données pour chaque cluster** : Extraction des variables environnementales pour les mois sélectionnés.
3. **Prédictions des cultures suggérées** : Utilisation d'un modèle prédictif pour déterminer la meilleure culture pour chaque cluster en fonction des variables environnementales.

Résultats

Les résultats finaux sont présentés dans le tableau ci-dessous :

Cluster	Mois	Mois le plus proche	Culture suggérée	N (mg/kg)	P (mg/kg)
0	1, 2, 3, 12	2	Banane	50.78	47.99
1	4, 5, 6, 9, 10, 11	5	Papaye	50.53	47.69
2	7, 8	7	Mangue	48.77	47.05

K (mg/kg)	Température (°C)	Humidité (%)	pH	Précipitations (mm)
33.14	23.31	54.98	6.90	68
33.08	29.57	54.87	6.99	68
33.71	34.08	52.16	6.86	68

Applications Clés

1. Généralisation à d'autres zones géographiques

Les clusters et modèles prédictifs peuvent être adaptés à toute zone ayant des caractéristiques similaires. Voici les principales opportunités :

- **Nouveaux territoires** : Avec des images satellitaires et la position géolocalisée d'une zone, il devient possible de prédire les variables environnementales pour une région non étudiée précédemment.
- **Réduction des coûts** : Cette méthode élimine le besoin d'investir dans des campagnes de collecte de données sur le terrain.
- **Précision spatiale** : La granularité de 10 mètres permet une analyse locale précise et pertinente pour des petites exploitations agricoles.

2. Planification des cultures saisonnières

Grâce aux clusters identifiés et aux cultures suggérées, il est possible de déterminer :

- **Les meilleures périodes de plantation** pour des cultures clés comme la banane, la papaye et la mangue.

- **Les ajustements des intrants agricoles** (azote, phosphore, potassium) en fonction des besoins prédictifs pour chaque cluster.
- **La gestion des ressources en eau** basée sur l'analyse du NDWI et des prévisions climatiques locales.

3. Assistance aux décideurs et aux ONG

Ce projet peut également servir de modèle pour des programmes de soutien à l'agriculture :

- **Prédictions à l'échelle régionale** : Les gouvernements peuvent utiliser cette technologie pour planifier des politiques agricoles évolutives.
- **Soutien à la sécurité alimentaire** : En prédisant les besoins en cultures et en ressources, il est possible d'anticiper les pénuries et de mieux gérer les risques.
- **Développement durable** : Cette méthode contribue à réduire le gaspillage de ressources en optimisant leur utilisation selon les besoins prédictifs.

Conclusion

Les résultats de ce projet offrent des possibilités d'application généralisables pour la gestion agricole dans diverses zones géographiques. En exploitant des images satellites et des modèles d'analyse des indices spectraux, il est possible de fournir des recommandations agricoles sans recourir à des données de terrain exhaustives. Cette approche peut transformer la manière dont les agriculteurs et les décideurs gèrent les ressources agricoles.

1. Prédire les valeurs des variables environnementales à partir d'une image satellitaire, grâce aux clusters définis dans la Partie 2.
2. Associer ces valeurs à des cultures optimales pour chaque saison et zone agricole étudiée.
3. Optimiser la planification des cultures en tenant compte des besoins spécifiques des plantes identifiés à travers les données environnementales (température, humidité, nutriments, etc.).
4. Réduire les coûts et les efforts liés à la collecte de données de terrain, permettant ainsi une prise de décision rapide et efficace pour les agriculteurs et les décideurs.

Ces résultats montrent que l'approche développée peut être appliquée à grande échelle, offrant une solution robuste pour améliorer la gestion agricole tout en assurant une meilleure utilisation des ressources et une durabilité accrue.