



Code UE: PVPS10
INTITULE DE L'UE: Infrastructure Big data
INTITULE ECUE: [Projet Big Data](#)
MASSE HORAIRE: 21H
CREDIT ECUE: 2
ENSEIGNANT: Dr. Olouladé Babatoundé Moctard
b.oloulade@outlook.com
METHODE D'ENSEIGNEMENT: Travaux de recherche de l'étudiant
Restructuration de connaissances
METHODE D'EVALUATION: Projet par groupe de 2/3 personnes

Titre du projet 1: Analyse des données de fabrication de ciment

Objectifs :

- Utiliser les méthodes de découverte de connaissances, de cloud computing et d'exploration de données pour analyser des données issues du processus de fabrication de ciment.
- Appliquer des techniques d'apprentissage automatique pour nettoyer, analyser et interpréter les données.
- Fournir des recommandations pratiques basées sur les résultats obtenus pour améliorer les prises de décision dans ce domaine industriel.

Travaux à réaliser :

- 1. Prétraitement des données :**
 - Identifier et traiter les valeurs manquantes et les données bruitées.
 - Supprimer les colonnes inutiles ou présentant des valeurs constantes.
 - Réaliser une transformation des variables pour standardiser l'ensemble des données.
- 2. Analyse exploratoire :**
 - Calculer les statistiques descriptives (moyenne, médiane, minimum, maximum, écart type) pour chaque variable si applicable.
 - Analyser les corrélations entre les différentes colonnes afin d'identifier des relations significatives.
- 3. Application des algorithmes :**
 - Implémenter des méthodes de machine learning (arbre de décision, régression par vecteurs de support, régression polynomiale, régression linéaire, CNN) pour combler les valeurs manquantes.
 - Comparer les performances des algorithmes à l'aide de métriques telles que le R^2 , l'erreur quadratique moyenne (RMSE) ou le coefficient de corrélation de Pearson r .
 - Justifier le choix de l'algorithme le plus pertinent pour les données étudiées.
- 4. Visualisation des résultats :**
 - Produire des graphiques de distribution des données pour certains paramètres clés.
 - Présenter les résultats des corrélations et des prédictions sous forme de tableaux ou de diagrammes clairs et professionnels.
- 5. Conclusion et recommandations :**

- Synthétiser les principales découvertes de votre analyse.
- Proposer des recommandations pratiques basées sur vos résultats pour aider à optimiser les processus de fabrication du ciment.

Livrables attendus :

- Un rapport structuré contenant :
 - Une introduction précisant le contexte et les objectifs de l'étude.
 - Une description des données et des étapes de prétraitement.
 - Les résultats obtenus et leur interprétation.
 - Une conclusion et des recommandations.
- Des annexes incluant les codes ou scripts utilisés.
- Un lien donnant accès au code et aux données nettoyées

Format et critères d'évaluation :

- Le rapport doit respecter les normes académiques en termes de structure et de présentation (introduction, méthodologie, résultats, conclusion, références).
- Une attention particulière sera accordée à la clarté de l'analyse, la rigueur méthodologique, la qualité des visualisations et la pertinence des recommandations.

Date limite de soumission : Avant le 20 Janvier 2024

NB : Faites preuve de créativité et de rigueur scientifique dans l'exécution de cette tâche.

Titre du projet 2 : Model de prédiction des récoltes agricoles

Problématique : Prévoir les rendements agricoles en fonction des conditions climatiques, des maladies et des types de cultures.

Tache: Concevoir un modèle de deep learning pouvant prédire les récoltes agricoles en utilisant des données de la FAO.

Livrables attendus : *Un rapport contenant :*

1. Rapport de revue de la littérature.
2. Rapport sur la collecte et la préparation des données.
3. Code Python pour la collecte et préparation des données. (Capture et lien)
4. Modèle de deep learning pour la prédiction des récoltes.
5. Rapport d'évaluation et validation du modèle.
6. Rapport final du projet.