

UNIVERSITE OFFICIELLE DE BUKAVU



ECOLE DES MINES

PRESENTATION N°2 : INITIATION L'IA ET GENIE LOGICIEL

ENERGY EFFICIENCY APP

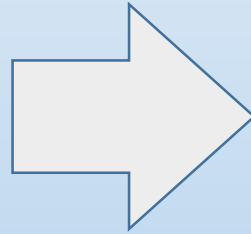
Membres du groupe

1. ASHUZA KULIMUSHI Ozias
2. MAPENDANO NTUGULO Lebon
3. USINDI RUGAMIKA Jonathan

Plan sommaire

Partie I : Initiation à l'IA

- Contexte du projet
- Présentation des données
- Outils et Méthodes
- Résultats et Discussions
- Recommandations



Partie II : Génie logiciel

- Contexte du projet
- Présentation de l'application
- Outils et Méthodes
- Résultats et Discussions
- Recommandations

Initiation à l'IA

Contexte du projet

- L'absence de la machine idéale est une contrainte pour l'industrie (minière)
- La maîtrise des coûts et leurs variations est le sens même de la planification minière
- La maîtrise des coûts liés à l'énergie commence par la maîtrise des besoins en énergie.



MODELE O'HARA



Machine Learning

Présentation des données

Energy Efficiency (UCI) : un tableau de 10 colonnes de données, avec 8 caractéristiques (X_i) et 2 cibles (Y_j) et au total 768 échantillons.

- X_1 : Compacité relative
- X_2 : Aire de surface
- X_3 : Surface murale
- X_4 : Surface de toit
- X_5 : Hauteur totale
- X_6 : Orientation
- X_7 : Zone vitrée
- X_8 : Distribution de la surface vitrée
- Y_1 : Charge de chauffage
- Y_2 : Charge de refroidissement

Initiation à l'IA

Présentation des données

Dataset statistics		Variable types	
Number of variables	10	Numeric	6
Number of observations	768	Categorical	4
Missing cells	0		
Missing cells (%)	0.0%		
Duplicate rows	0		
Duplicate rows (%)	0.0%		
Total size in memory	60.1 KiB		
Average record size in memory	80.2 B		

Outils et méthodes

Outils :

Les outils utilisés pour cette partie sont :

- Jupyter Notebook
- Des bibliothèques nécessaires de sklearn (voir le code)
- La fonction MSE, Adam, Tangente hyperbolique,...
- Le modèle GBM et le modèle de réseau de neurones.

Méthode :

La méthode adoptée consiste à entraîner le modèle sur les données d'entraînement, le tester avec les données de test et l'évaluer (MSE et r^2). Ensuite, optimiser les paramètres du modèle pour minimiser le MSE et maximiser le r^2 .

Résultats et discussions

Linear Regression: MSE = 8059206937214674748309504.000, R2 = -84244242288389363073024.000

Ridge: MSE = 12.270, R2 = 0.869

Lasso: MSE = 12.620, R2 = 0.865

Elastic Net: MSE = 19.559, R2 = 0.792

Polynomial Regression: MSE = 20075.704, R2 = -212.999

Decision Tree: MSE = 59.613, R2 = 0.365

Random Forest: MSE = 24.269, R2 = 0.741

XGBoost: MSE = 72.440, R2 = 0.232

KNN: MSE = 8.233, R2 = 0.911

OLS: MSE = 12.260, R2 = 0.869

Gaussian Process: MSE = 13.436, R2 = 0.855

Bagging: MSE = 24.386, R2 = 0.739

Réseau des neurones: MSE = 0.455, R2 = 0.995

Résultats et discussions

CIBLE 1

Gradient Boosting Machines (GBM)

CIBLE 2

Erreur quadratique moyenne

```
mse = mean_squared_error(y_test_1d, y_pred)
print(f'Erreur quadratique moyenne: {mse}')
```

Coefficient de détermination

```
r2 = r2_score(y_test_1d, y_pred)
print(f'Coefficient de détermination: {r2}')
```

Erreur quadratique moyenne: 0.10884208458340759
Coefficient de détermination: 0.9989162729362436

Erreur quadratique moyenne

```
mse2 = mean_squared_error(y_test_1d2, y_pred2)
print(f'Erreur quadratique moyenne: {mse2}')
```

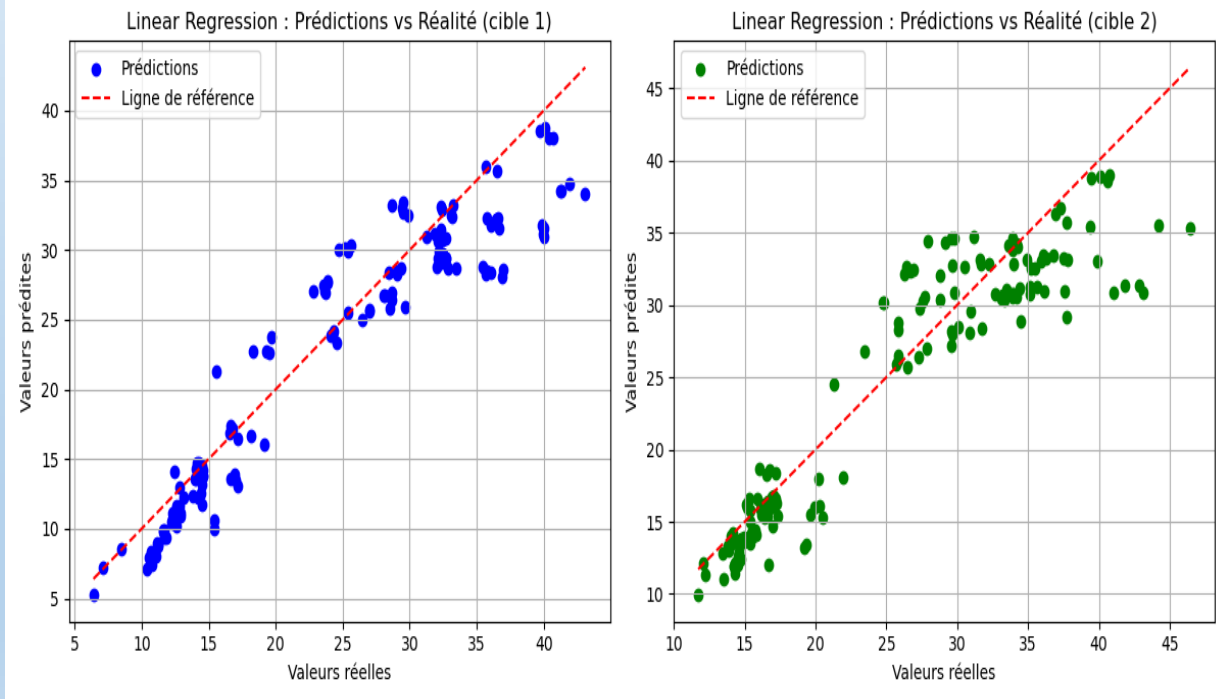
Coefficient de détermination

```
r2_2 = r2_score(y_test_1d2, y_pred2)
print(f'Coefficient de détermination: {r2_2}')
```

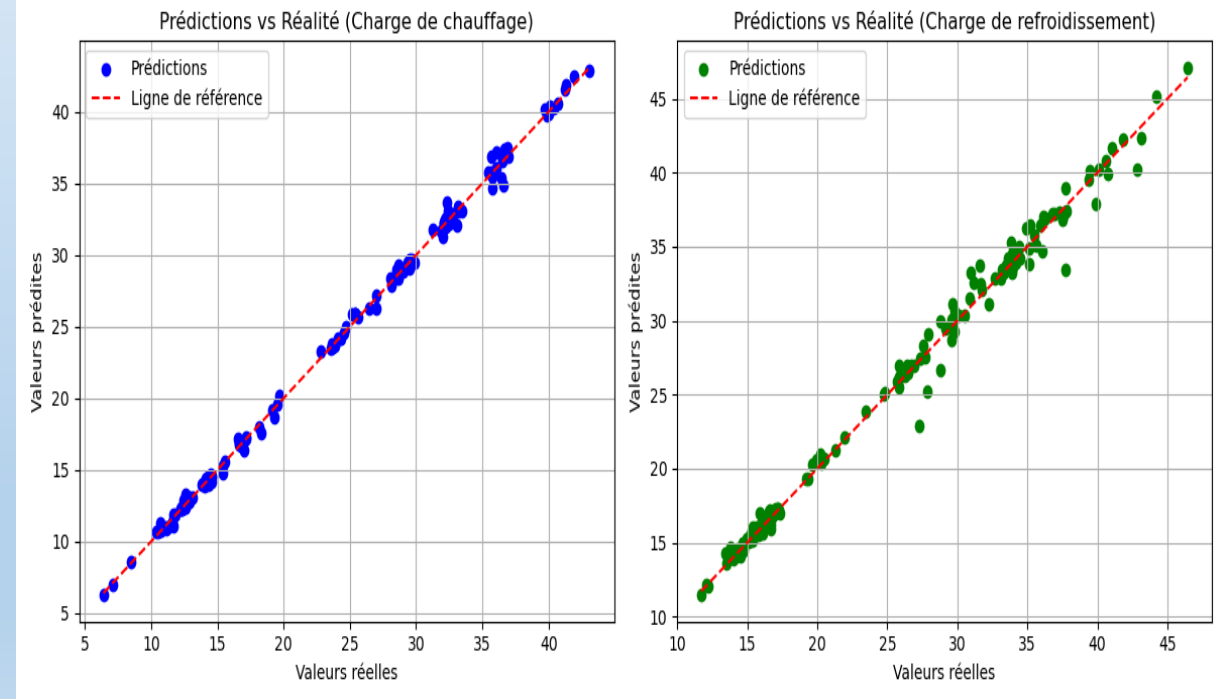
Erreur quadratique moyenne: 0.3467723401895603
Coefficient de détermination: 0.9960562376685818

Résultats et discussions : cas du réseau de neurones

Linear Regression

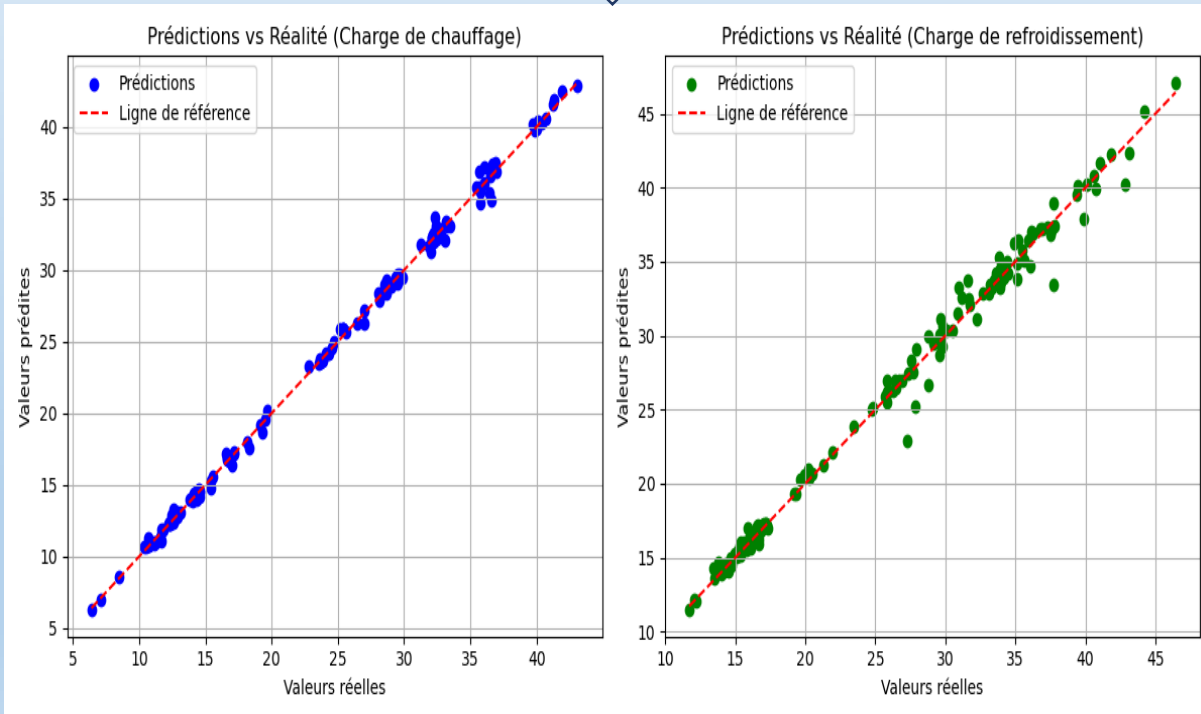


Réseau de neurones

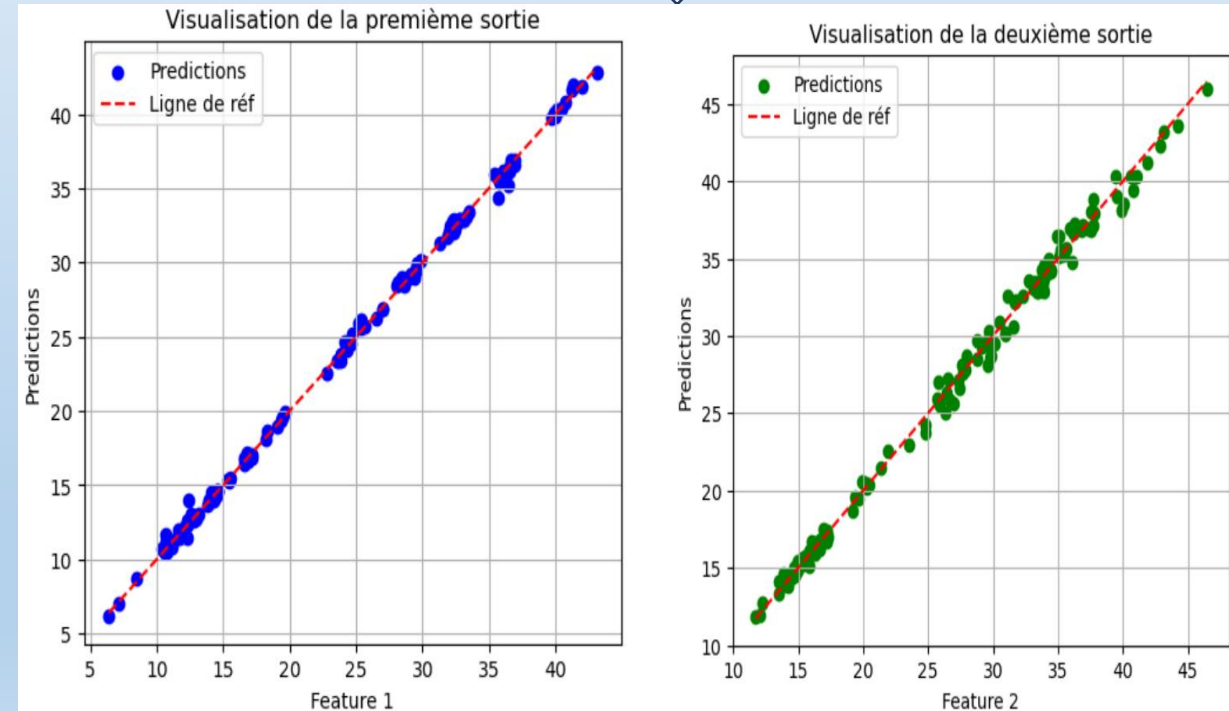


Résultats et discussions : cas du réseau de neurones

Réseau de neurones



Gradient Boosting Machines (GBM)

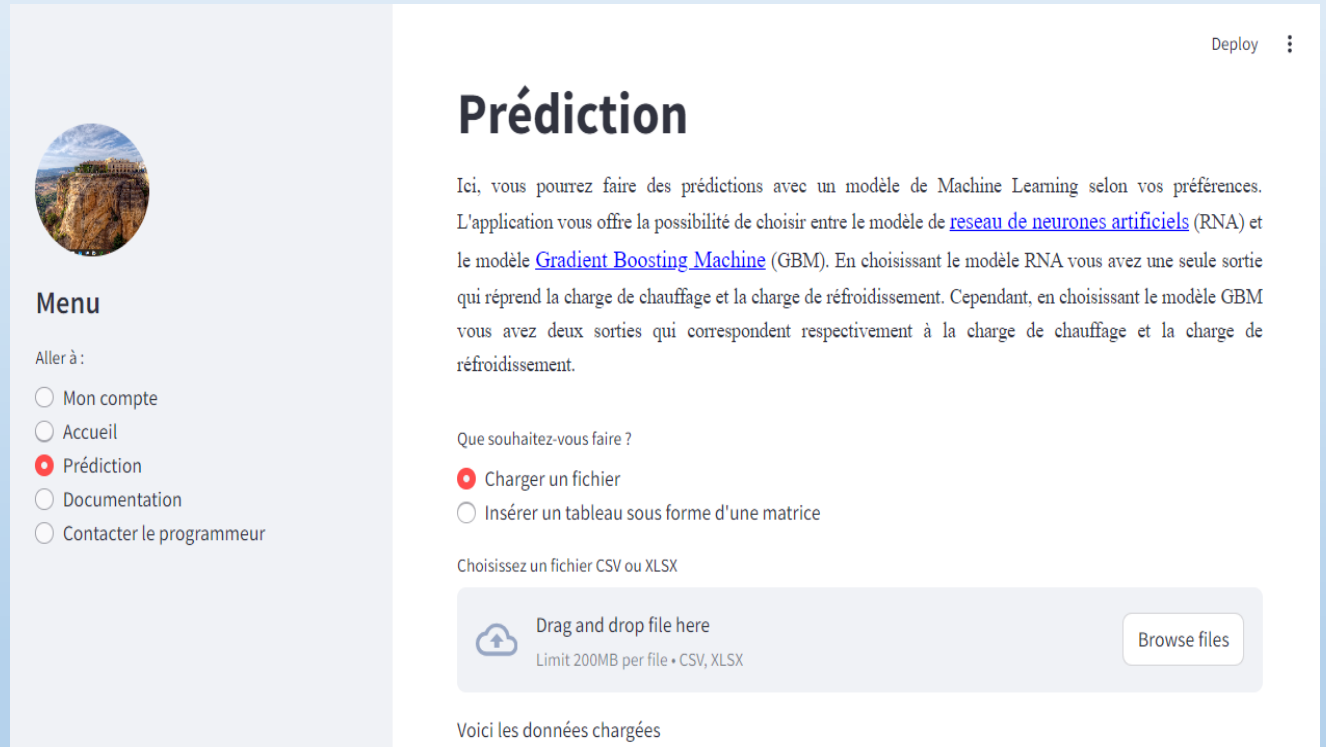


Partie II : Génie Logiciel

Contexte du projet

- Le public non expérimenté en programmation est celui visé par le projet de développement de l'application.
- Pour faciliter une utilisation des modèles de prédiction de charges énergétiques par les industriels et les entrepreneurs, il faut développer quelque chose de facile à manipuler.
- Une application web qui offre plusieurs possibilités à l'utilisateur et qui le guide pas à pas devient dès lors indispensable pour la vulgarisation du modèle de prédiction entraîné.

Présentation de l'application



The screenshot displays a web application interface. On the left, a sidebar contains a circular profile picture of a cliff and a 'Menu' section with links: 'Aller à:', 'Mon compte', 'Accueil', 'Prédiction' (highlighted with a red dot), 'Documentation', and 'Contacter le programmeur'. The main content area is titled 'Prédiction' and includes a 'Deploy' button in the top right. The text explains that users can make predictions using Machine Learning models (RNA or GBM). It details the outputs for each model: RNA provides heating and cooling charges, while GBM provides both. Below this, a section 'Que souhaitez-vous faire ?' offers two options: 'Charger un fichier' (selected with a red dot) and 'Insérer un tableau sous forme d'une matrice'. A file upload section follows, with the text 'Choisissez un fichier CSV ou XLSX', a 'Drag and drop file here' area with a cloud icon and 'Limit 200MB per file • CSV, XLSX', and a 'Browse files' button. At the bottom, it says 'Voici les données chargées'.

Partie II : Génie Logiciel

Outils et Méthodes

Outils :

- Python
- DB Browser SQLite
- Streamlit
- GitHub
- Modèles entraînés et enregistrés avec joblib
- PowerShell

Méthode :

La base de données est créée dans python avec les fonctions de Streamlit, gérée par le SGBD DB Browser SQLite. Le code est sauvegardé sur GitHub et l'application déployée sur Streamlit pour être consulté par le grand public.

Résultats et Discussions

- L'application fonctionne très bien, le CDC est respecté car l'application permet de prédire la charge de chauffage et charge de refroidissement
- Les possibilités offertes à l'utilisateur de profiter d'un maximum d'expérience utilisateur de l'application.
- Toutefois, le temps de réponse de l'application reste encore à revoir.
- Quant à l'application déployée, elle nécessite une révision surtout du côté importation de la bibliothèque joblib.

Recommandations

- **Optimisation des Modèles :**

- Continuer à affiner les hyper paramètres des modèles (GBM et réseau de neurones) pour améliorer la précision des prédictions.
- Tester d'autres algorithmes de machine Learning pour voir s'ils peuvent surpasser les performances des modèles actuellement utilisés.

- **Amélioration de l'Application :**

- Réduire le temps de réponse de l'application en optimisant le code et en envisageant des solutions de mise en cache.
- Intégrer des fonctionnalités supplémentaires basées sur les retours des utilisateurs pour améliorer l'expérience utilisateur.

Recommandations

- **Formation et Documentation :**

- Fournir des ressources de formation et de documentation pour les utilisateurs non expérimentés afin de les aider à mieux comprendre comment utiliser l'application.
- Créer des tutoriels vidéo ou des guides étape par étape pour faciliter la prise en main.

- **Suivi et Maintenance :**

- Mettre en place un système de suivi pour évaluer régulièrement la performance de l'application et des modèles.
- Planifier des mises à jour régulières pour corriger les bugs et améliorer les fonctionnalités.

- **Collaboration et Partenariats :**

- Envisager des collaborations avec des experts en énergie et en mines pour affiner les modèles et les rendre plus pertinents pour l'industrie.
- Explorer des partenariats avec des institutions ou des entreprises pour élargir la portée de l'application et obtenir des retours d'expérience supplémentaires.

MERCI POUR VOTRE ATTENTION !

