

Prédiction de la Production Électrique des Micro-Turbines à Gaz : Un Problème de Régression Complexe

Introduction du Problème

La production électrique des micro-turbines à gaz est un processus complexe influencé par plusieurs variables comme le temps et la tension. Ce projet aborde un problème de régression où l'objectif est de prédire la production électrique future à partir de données historiques, une tâche cruciale pour optimiser l'efficacité énergétique.

Présentation du Problème

Le défi principal réside dans la précision de la prévision de la production électrique. Avec des données dynamiques et non linéaires et temporelles, il est nécessaire de développer des modèles capables de capturer ces variations pour fournir des prévisions fiables.

Description du Problème

La production électrique (en watts) dépend de divers facteurs mesurés sur une échelle de temps. Le but est de prévoir la puissance future (P_{t+1}) à partir des décalages temporels (P_t , P_{t-1} , P_{t-2}), du temps et de ses dérivées tel que (l'année, le mois, la semaine, le jour, l'heure). Cette tâche est cruciale pour une gestion efficace des micro-turbines, qui doivent répondre à des demandes énergétiques fluctuantes.

Solution Proposée

Pour répondre à ce défi, une approche basée sur des modèles de régression avancés a été adoptée. L'idée est de construire un modèle capable de prévoir avec précision la production future en tenant compte des différentes variables temporelles et électriques.

Méthodologie

La méthodologie suivie comprend plusieurs étapes :

1. **Prétraitement des données** : Transformation et nettoyage des données pour éliminer les anomalies.
2. **Feature Engineering** : Sélection et transformation des caractéristiques pertinentes pour améliorer la performance du modèle.
3. **Choix des Modèles** : Test de plusieurs modèles de régression comme Régression Linéaire, Random Forest, XGBoost, LightGBM, et CatBoost.
4. **Optimisation des Hyperparamètres** : Utilisation de la méthode RandomizedSearchCV pour affiner les paramètres des modèles.

Outils et Technologies

Les outils utilisés incluent :

- **Google Colab** : Pour l'environnement de développement.
- **Bibliothèques Python** : Scikit-learn, NumPy, Pandas pour l'analyse des données et l'implémentation des modèles.
- **Librairies de Machine Learning** : XGBoost et LightGBM pour les modèles de boosting.

Résultats

Les résultats ont montré une performance remarquable, particulièrement avec le modèle Multi-Layer Perceptron (MLP) après optimisation, atteignant un coefficient de détermination (R^2) de 0.9999 et un MSE de 89.8137, ce qui indique une précision presque parfaite.

Discussion et Analyse

Les résultats révèlent l'efficacité des modèles choisis, en particulier après l'optimisation des hyperparamètres. Cependant, certaines techniques d'ensemble ont montré une baisse de performance, soulignant l'importance de choisir le bon modèle pour le bon contexte. Les résultats ont également montré que la robustesse des modèles peut être améliorée avec des méthodes de validation croisée.

Conclusion

Ce projet a démontré la faisabilité et l'efficacité des techniques de régression avancées pour prédire la production électrique des micro-turbines à gaz. Les prochaines étapes incluront l'exploration de modèles plus sophistiqués et des techniques de validation pour améliorer encore la précision et la robustesse des prévisions.