

# **PRÉDICTION ET EXPLICABILITÉ DE LA DÉGRADATION DES BATTERIES**

**MODÈLES DE PRÉDICTION DU SOH  
BASÉS SUR LES DONNÉES DE BATTERIES DU MIT**

ZOÉ MARQUIS  
CHARLOTTE KRUZIC

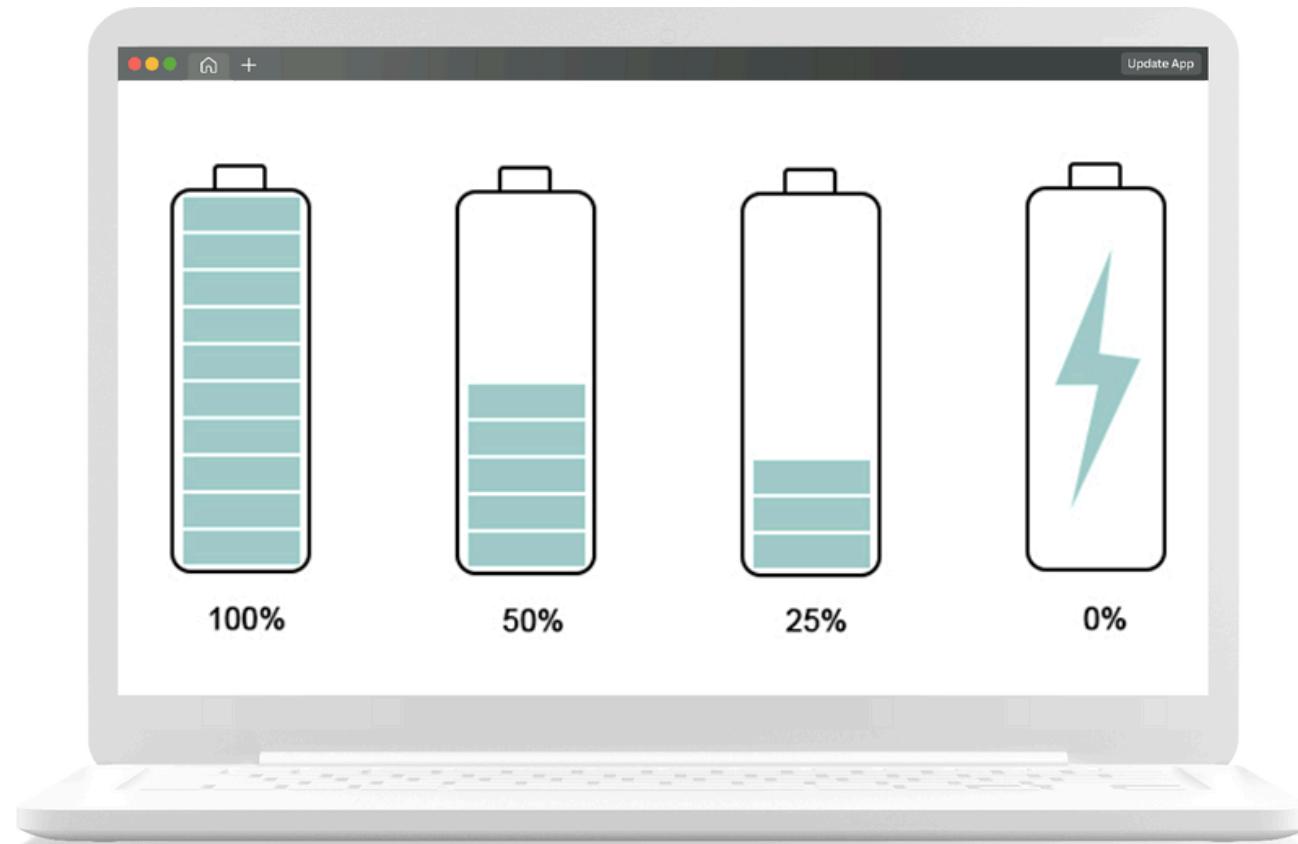
# Introduction

## Objectifs

- Développer un modèle prédictif
- Identifier les facteurs influençant

## Méthodes clés

- Prédiction de séries temporelles
- Techniques d'explicabilité

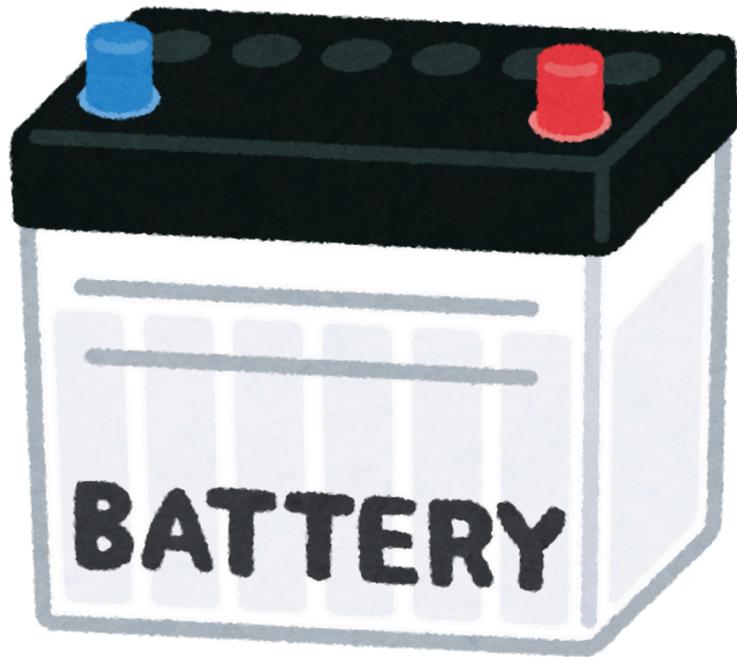


Batteries du MIT  
• 3 batchs  
• 124 batteries

# Méthodologie

1. Préparation des données
2. Création de fenêtres glissantes
3. Entraînement et optimisation
4. Comparaison des performances
5. Application des techniques d'explicabilité

# Structure des données

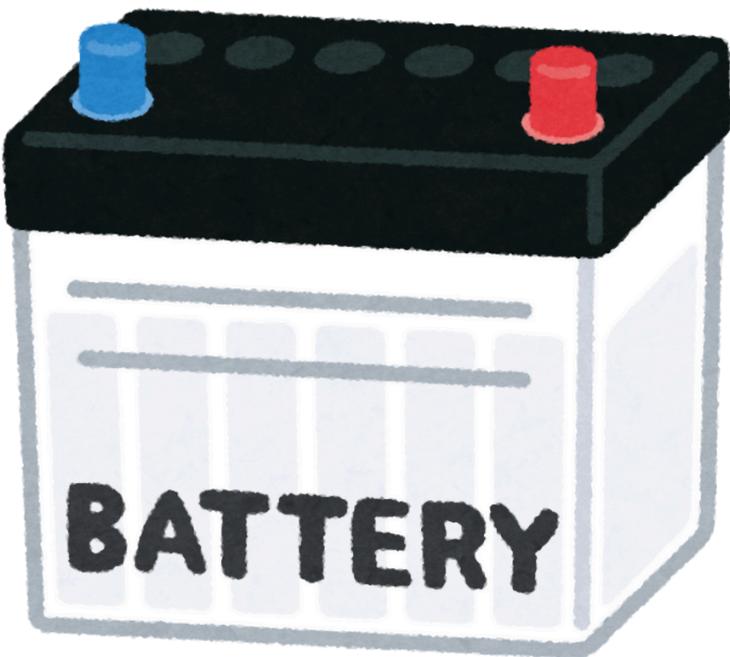


## Données stables

Colonnes à valeurs fixes dans les données d'une batterie

- `cycle_life`
- `charge_policy` : C1, C2, Q1
- `barcode`
- `channel_id`

# Structure des données

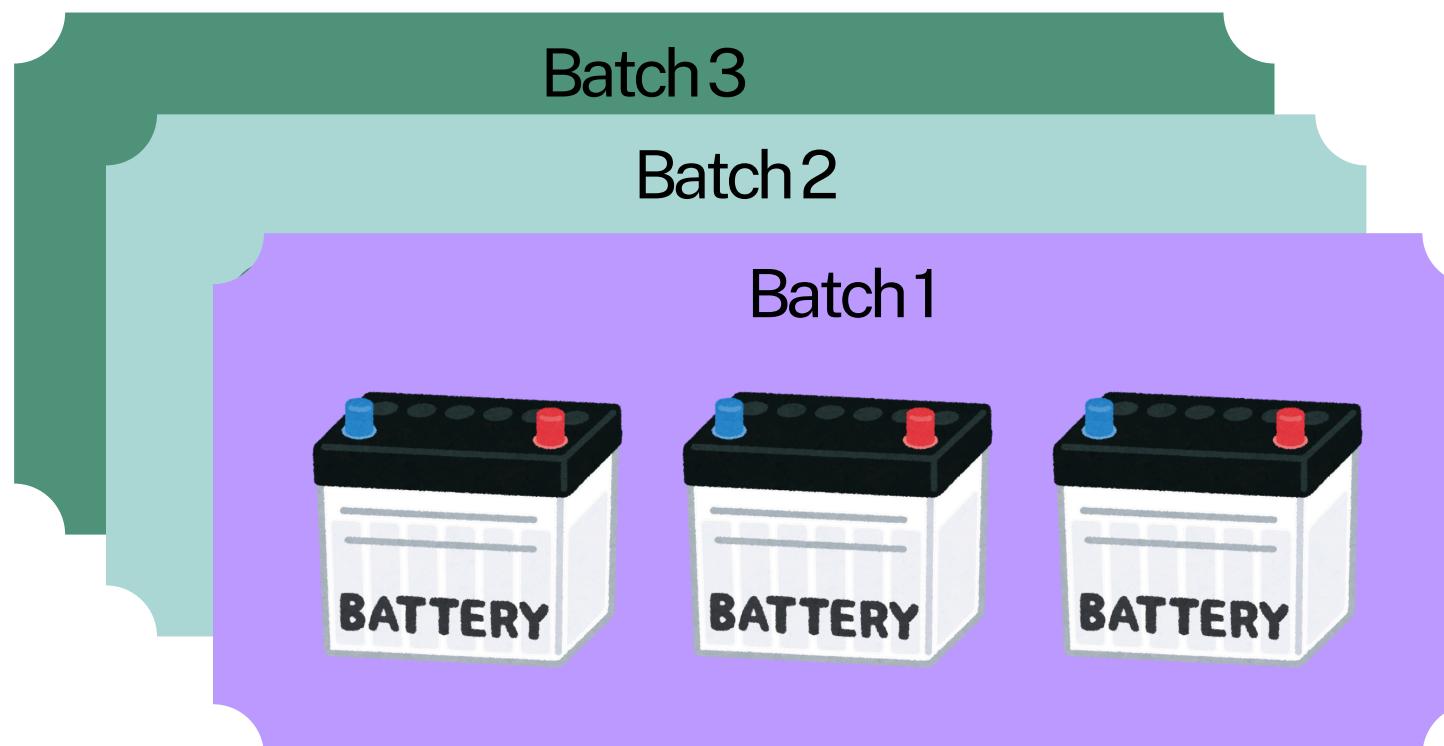


## Données temporelles

Colonnes évolutives d'une batterie

cycle (numéro de cycle)	IR	QC	Tavg	Tmin	Tmax	chargetime	SOH
1							
2							
3							

# Structure des données



## Données communes au batch

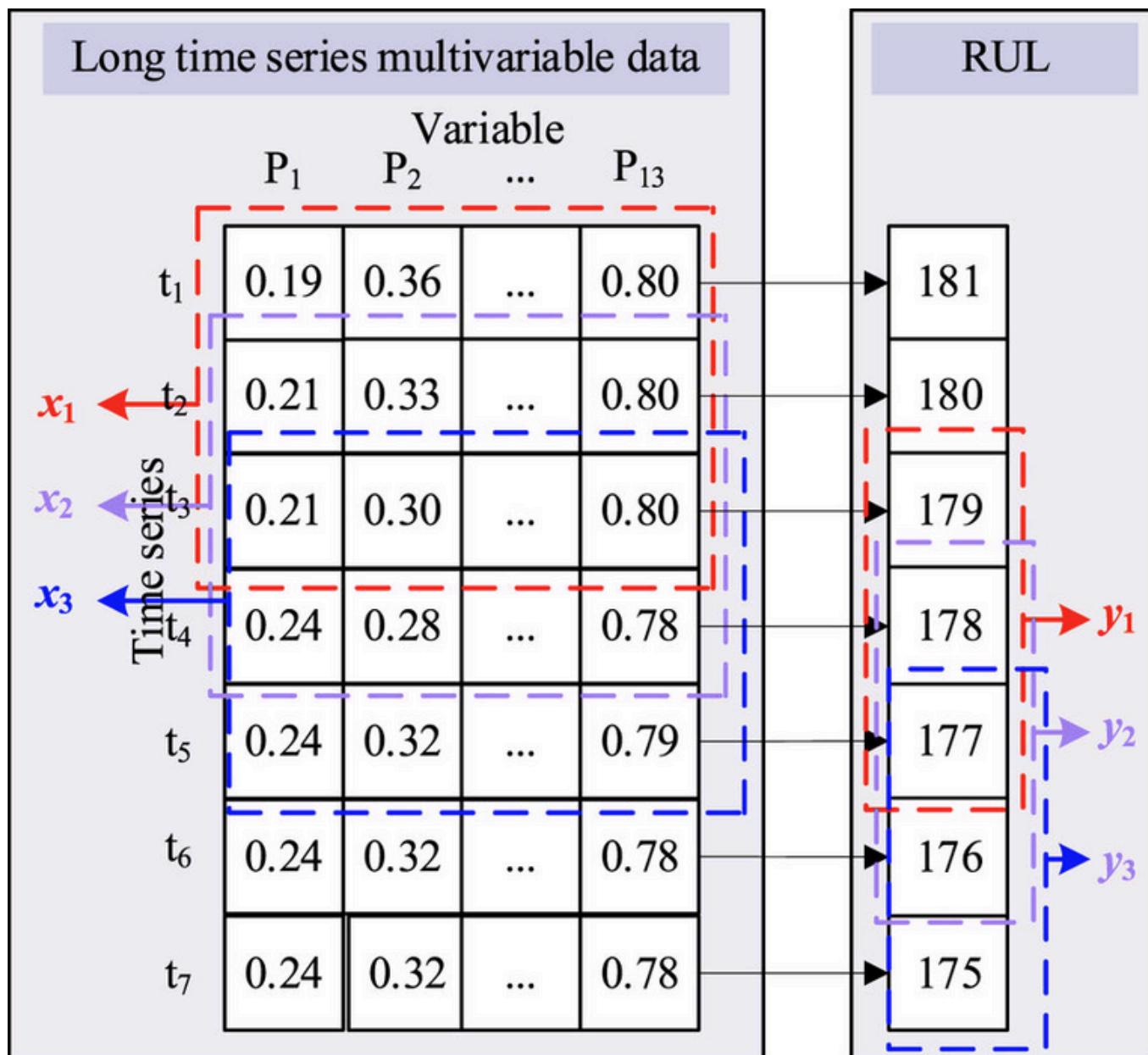
(extraites de la documentation)

- pulse\_width
- cycle\_to\_x% (of nominal capacity)
- cutoff\_currents
- charging\_rest
- after\_discharging\_rest
- before\_discharging\_rest
- IR\_rest

# Préparation des données

- Gestion des valeurs aberrantes
- Normalisation
- Encodage avec 2 méthodes

# Fenêtres glissantes



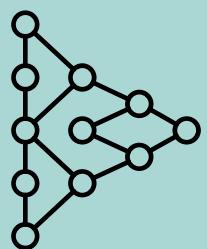
## Utilité

- Découper les séries temporelles

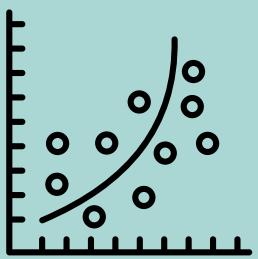
## Taille optimale

- Taille fenêtre et pas (stride)
- par itération avec Random Forest
- Taille :
  - Fenêtre : 20 timesteps (cycles)
  - Stride : 5

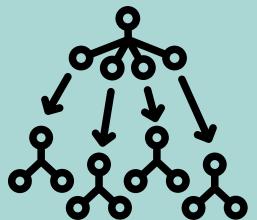
# Modèles Développés



- Réseaux de neurones récurrents (RNN) :
- SimpleRNN
  - LSTM
  - CNN 1D

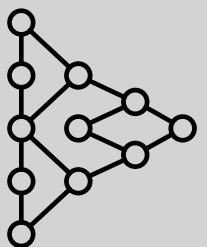


- Modèles de régression :
- Linear
  - Ridge
  - Lasso
  - Elastic Net

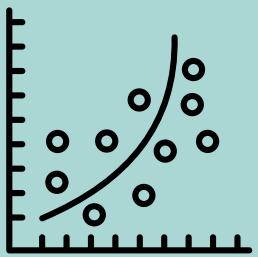


- Modèles basés sur des arbres de décision :
- Random Forest
  - XGBoost

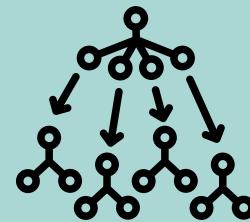
# Optimisation Grid search



- Réseaux de neurones récurrents (RNN) :
- SimpleRNN
  - LSTM
  - CNN 1D

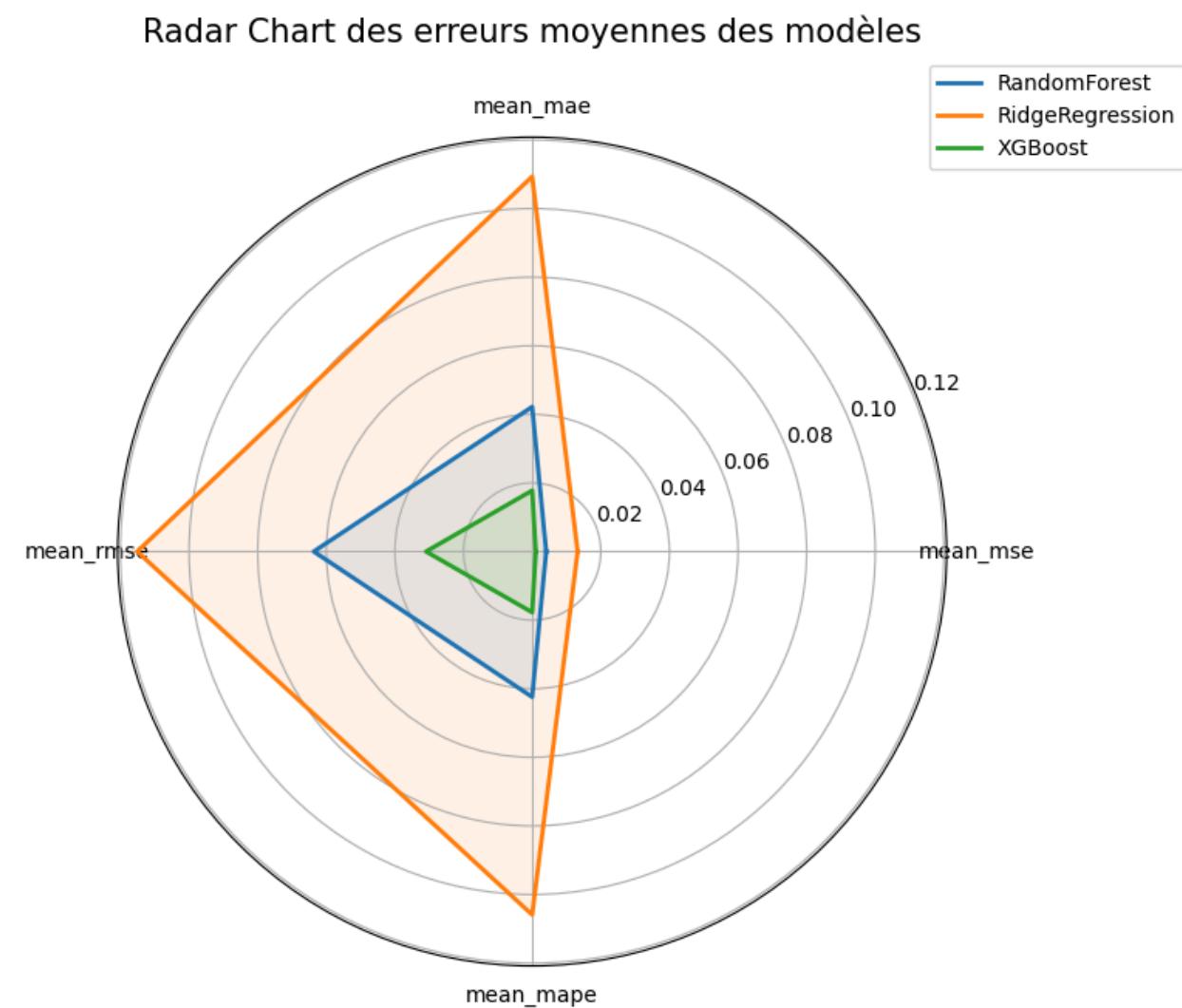
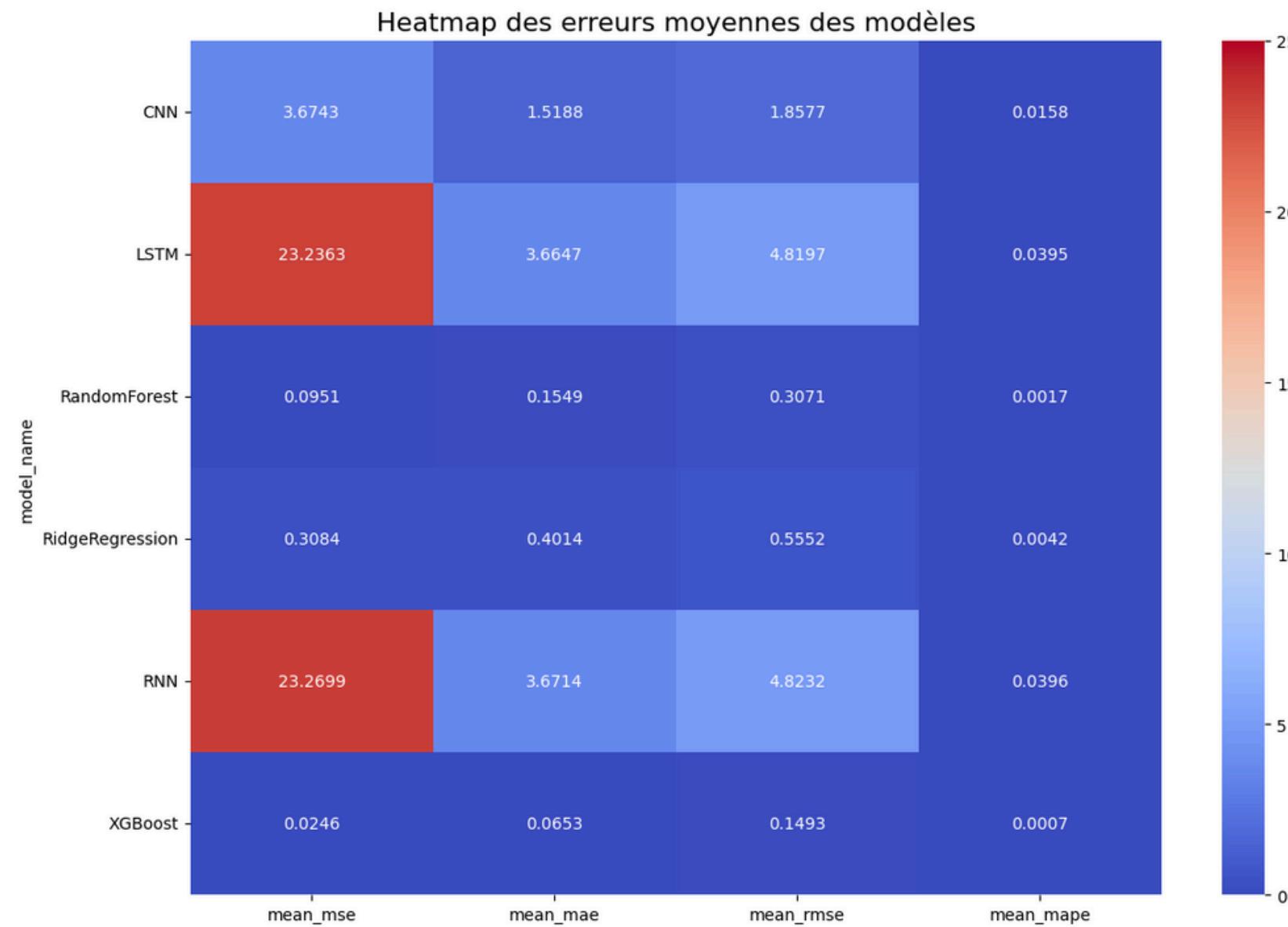


- Modèles de régression :
- Linear
  - Ridge
  - Lasso
  - Elastic Net



- Modèles basés sur des arbres de décision :
- Random Forest
  - XGBoost

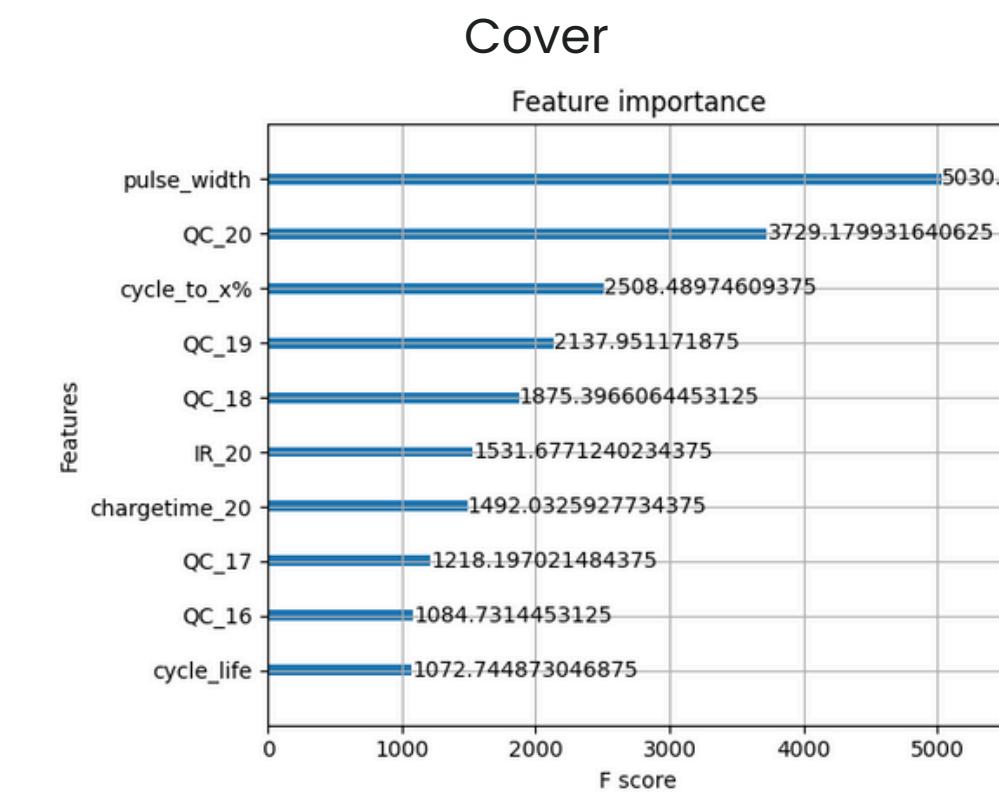
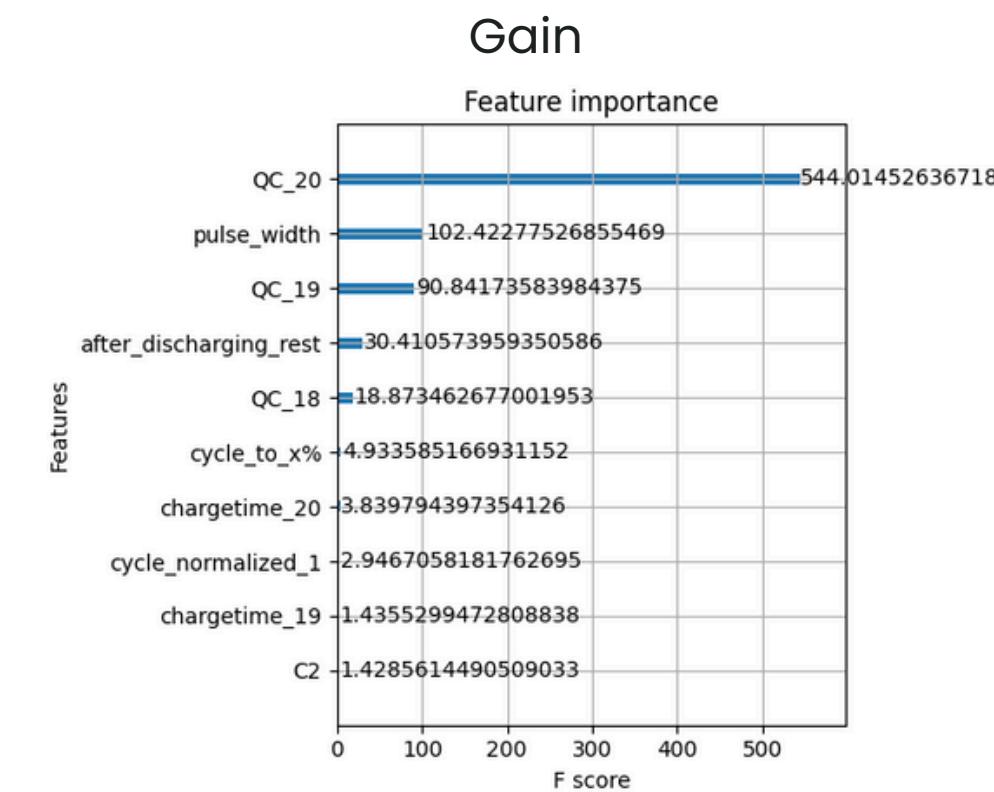
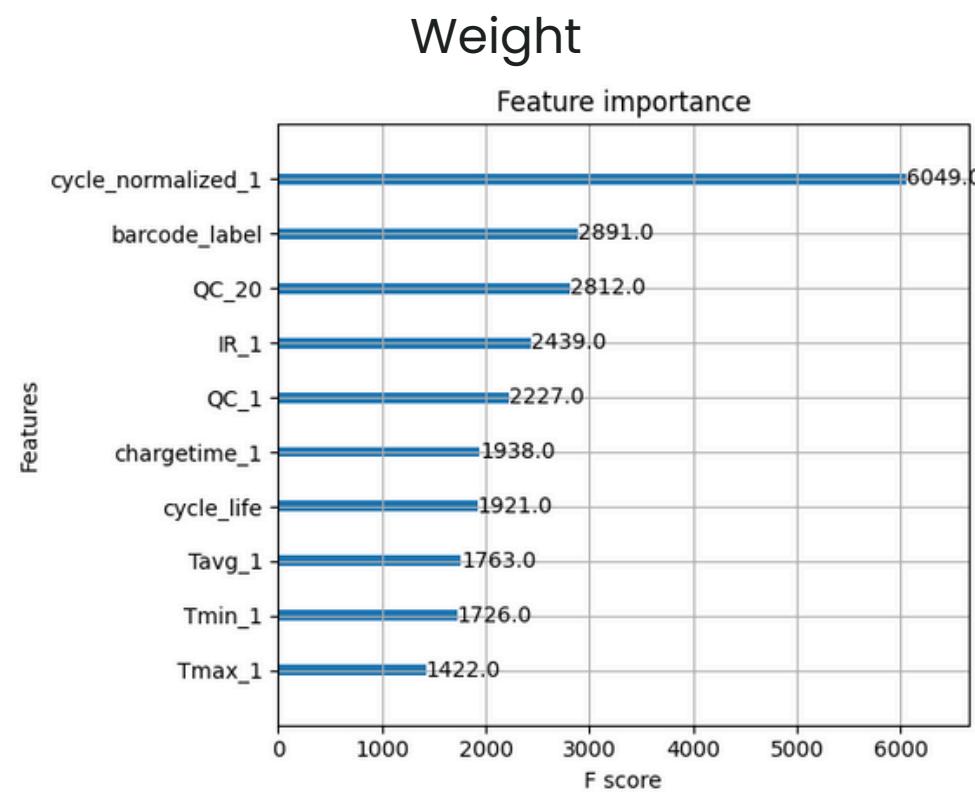
# Comparaison des performances



# EXPLICABILITÉ

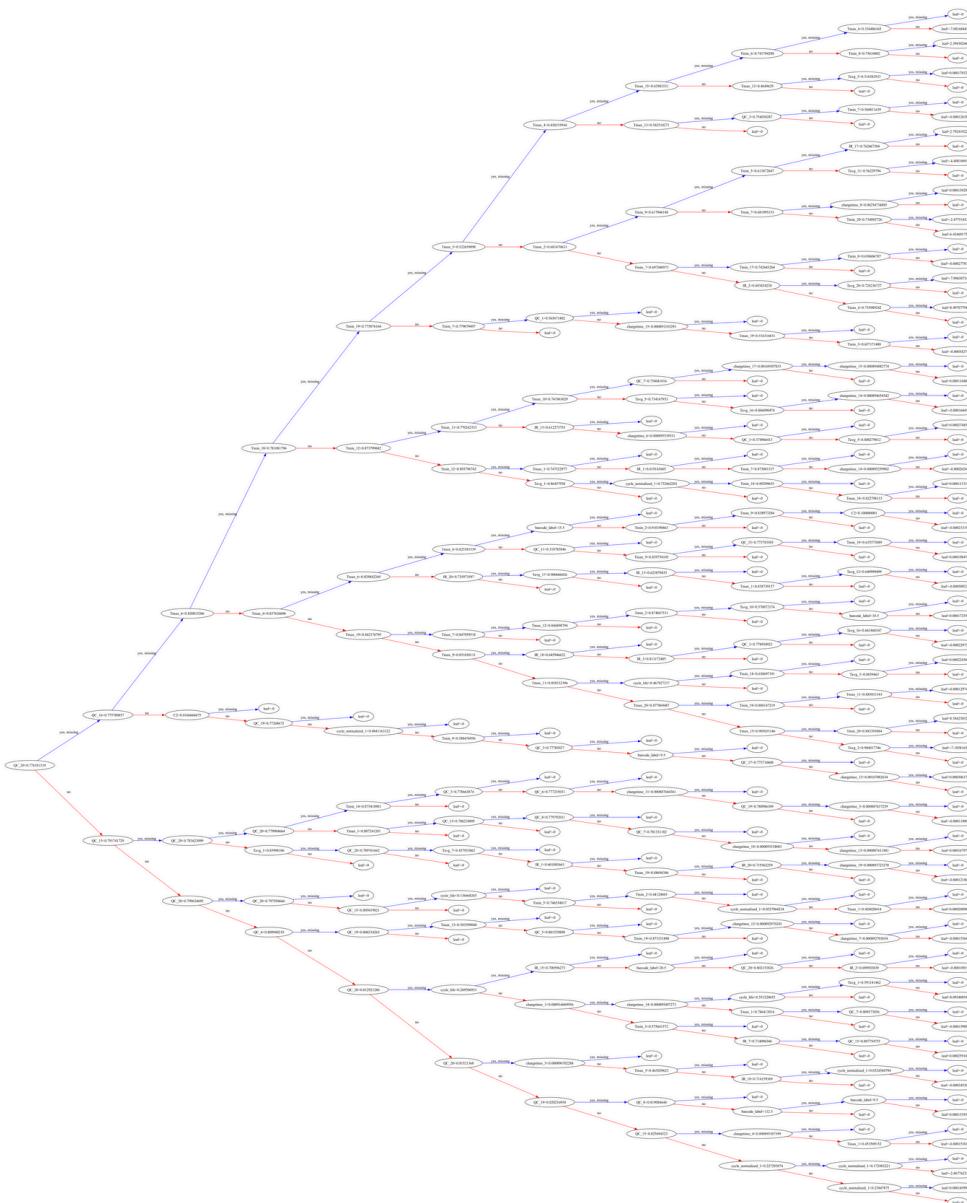
## Importance des features

- Weight : Fréquence d'utilisation des variables dans les arbres.
- Gain : Contribution des variables à la réduction de l'erreur.
- Cover : Nombre moyen d'observations couvertes par les branches d'un arbre.



# EXPLICABILITÉ

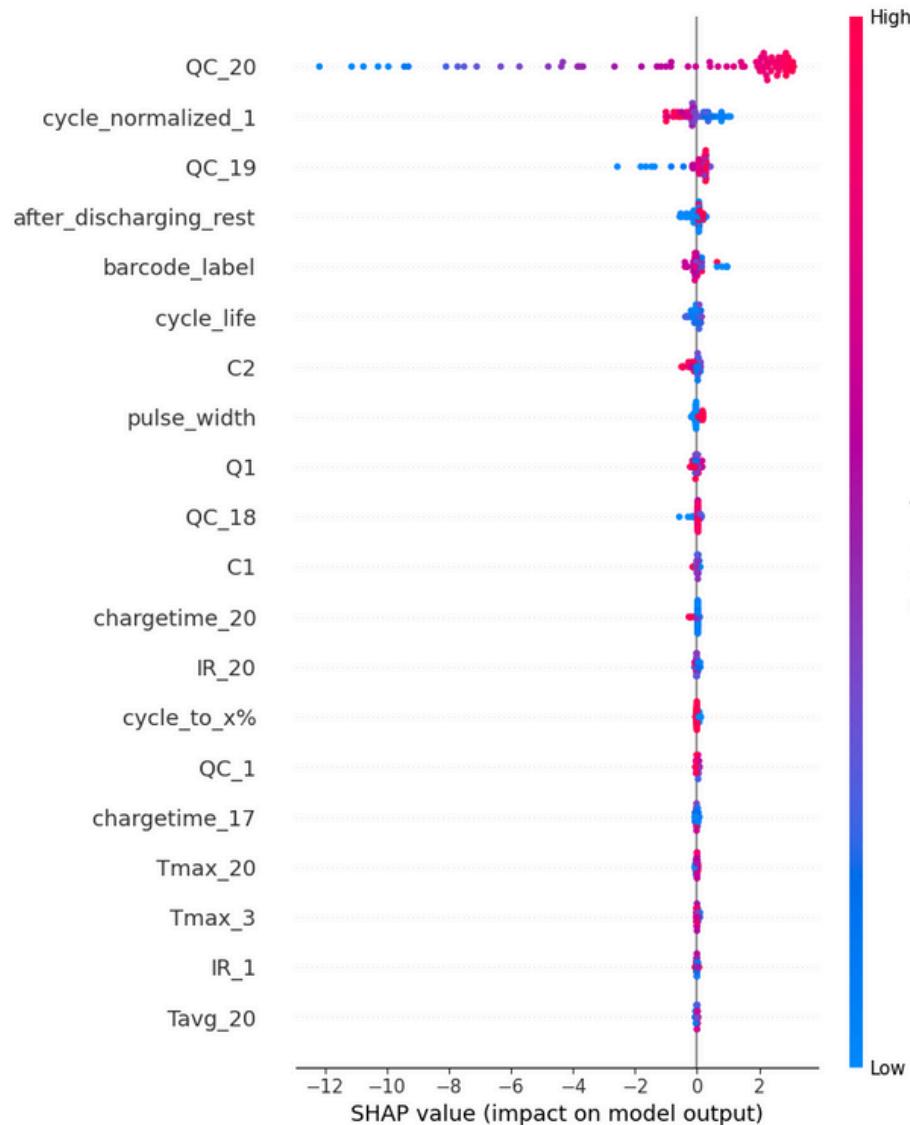
Visualisation du meilleur arbre de décision



- Split initial : Basé sur QC\_20, soulignant son importance dans la prise de décision.
- Splits suivants : Utilisation de caractéristiques comme QC\_15, QC\_16, et températures (max, avg, min).
- Temporalité : Forte contribution des variables aux timesteps avancés (autour de 20).
- Impact global : Variables QC, températures et charge time sont déterminantes.

# EXPLICABILITÉ

## SHAP : Beeswarm Plot



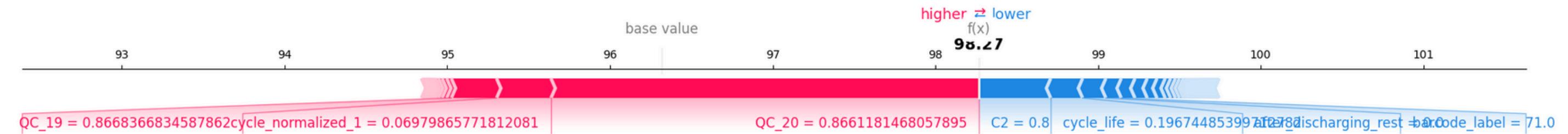
Montrer l'impact global des variables sur les prédictions du modèle.

- QC\_20, QC\_19 : Impacts majeurs, avec une relation inverse entre leurs valeurs et la prédiction du SOH.
- cycle\_normalized\_1 : Relation stable mais inversée (valeurs élevées → prédictions faibles).
- Facteurs clés : Derniers timesteps (QC, IR, températures) sont essentiels pour prédire l'état de santé de la batterie.
- Caractéristiques secondaires : Stratégies de charge et variables liées aux premiers cycles influencent également le SOH.

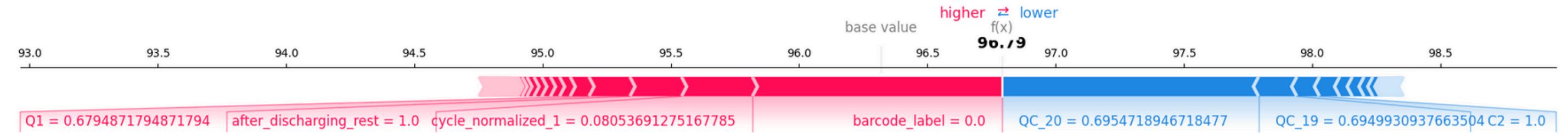
# EXPLICABILITÉ

## SHAP : Force Plot

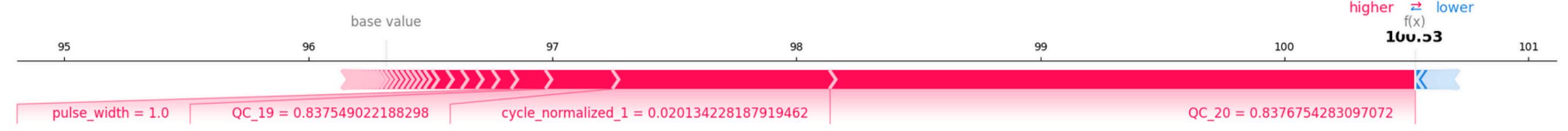
1



2



3



Courbe représentative de SOH pour chaque batch (lissage triple-exponentiel)



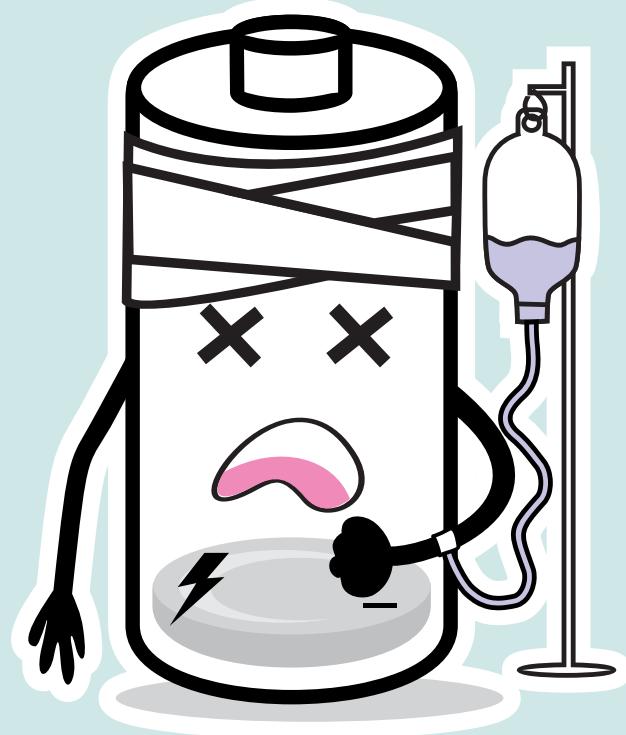
# TENTATIVES

- Récupération des données des cycles
- TimeSHAP : Analyse de l'impact des séquences temporelles sur les prédictions
- PyXAI : Explicabilité des modèles basés sur des arbres



# Conclusion

## Récapitulatif des objectifs



- Développer un modèle prédictif capable d'anticiper la courbe de SOH des batteries à partir des valeurs précédentes.
- Offrir une interprétation des facteurs clés via des techniques d'explicabilité comme SHAP ou LIME.
- Guider les décisions de maintenance et d'optimisation des batteries grâce à des résultats interprétables.

# Conclusion

Résultats obtenus

