

Maintenance préventive

Data challenge 2024



gwenlake.



enedis



BARRAUD Lorenzo
BRIAND Léo
MIRZA Simon
SONDEJI KARL

Contexte

- Le réseau électrique en Bretagne représente **107 792 km**.
- Dans un contexte **environnemental responsable**, allouer de manière optimale les différentes interventions de maintenance du réseau est un enjeu essentiel.
- L'objectif est de proposer une approche statistique classique ainsi qu'une méthode par Machine Learning afin de couvrir l'aspect de **l'IA responsable**.

01

INTRODUCTION

02

MODÉLISATION

03

CONCLUSION



1. INTRODUCTION

1.1 Présentation de la base

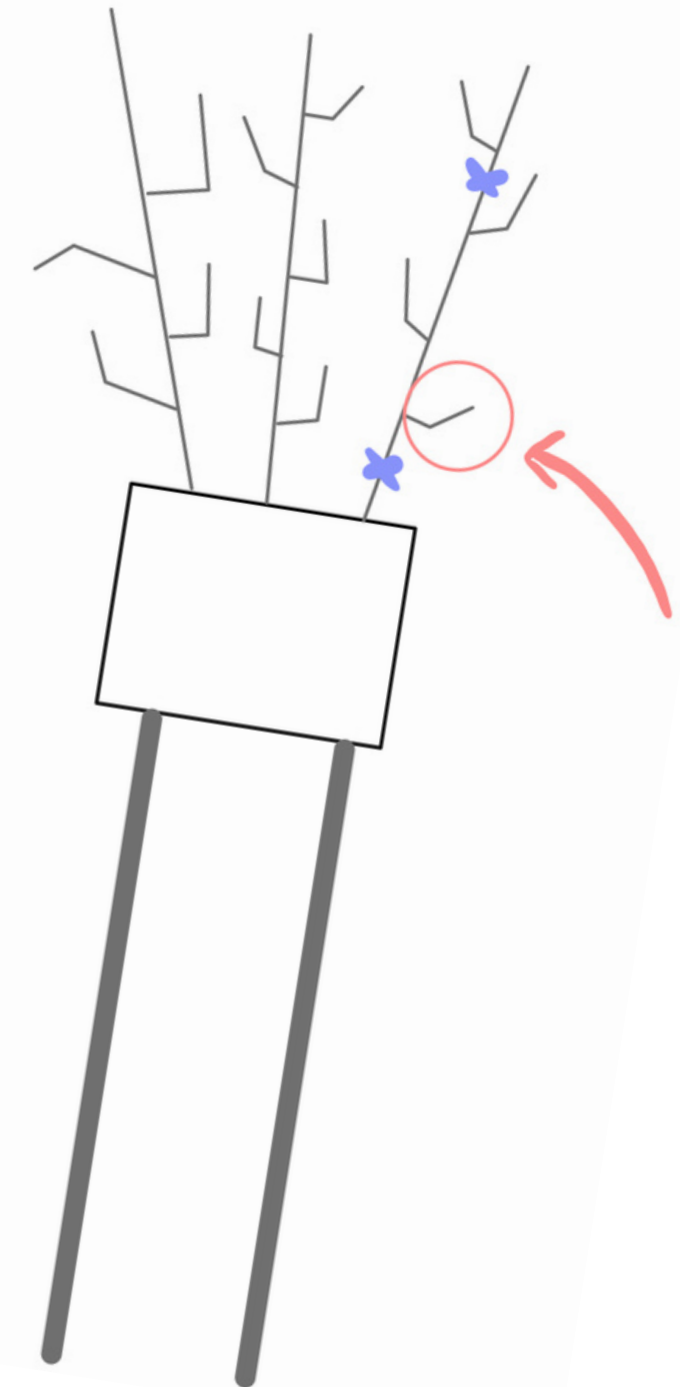
1.2 Exploration de la base

1.1 Présentation de la base

- Complexité : composée de 9 variables dont 3 variables date
- La variable que l'on cherche à prédire est la **probabilité de survenue d'un incident sur un tronçon**

Nettoyage des données :

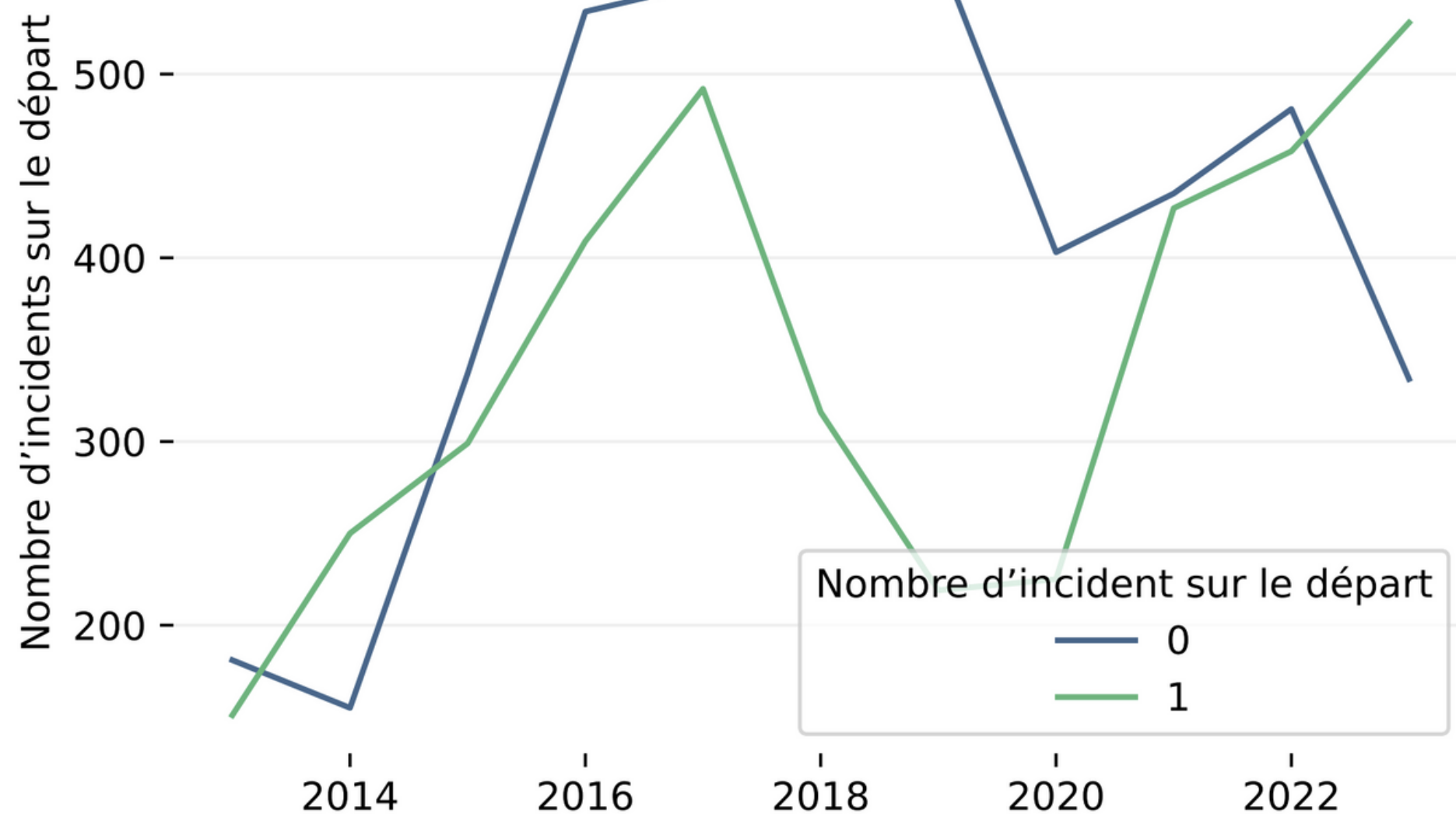
- Remplacement des **valeurs manquantes** par 0
- Variable cible transformée en binomiale
- Date de mise en service : transformée en ancienneté du tronçon
- Année du dernier vol sur le départ supprimée pour l'année 2022
- Remplacement des incidents classés 1 en 0 s'il n'y a pas eu de maintenance



1.2 Exploration de la base de données

- Plus il y a de maintenances effectuées en N-1 plus le nombre d'incidents l'année N est faible.
- On estime qu'une maintenance arrive à prévenir 50% des incidents.

Nombre d'incidents l'année de la dernière maintenance sur le départ



2. MODÉLISATION

2.1 Comparaison des résultats

2.2 Interprétations

2.3 Focus

2.1 Comparaison des résultats

	Logistique	XGBoost (ML)
AUC	0.59	0.79
F1	0.57	0.72

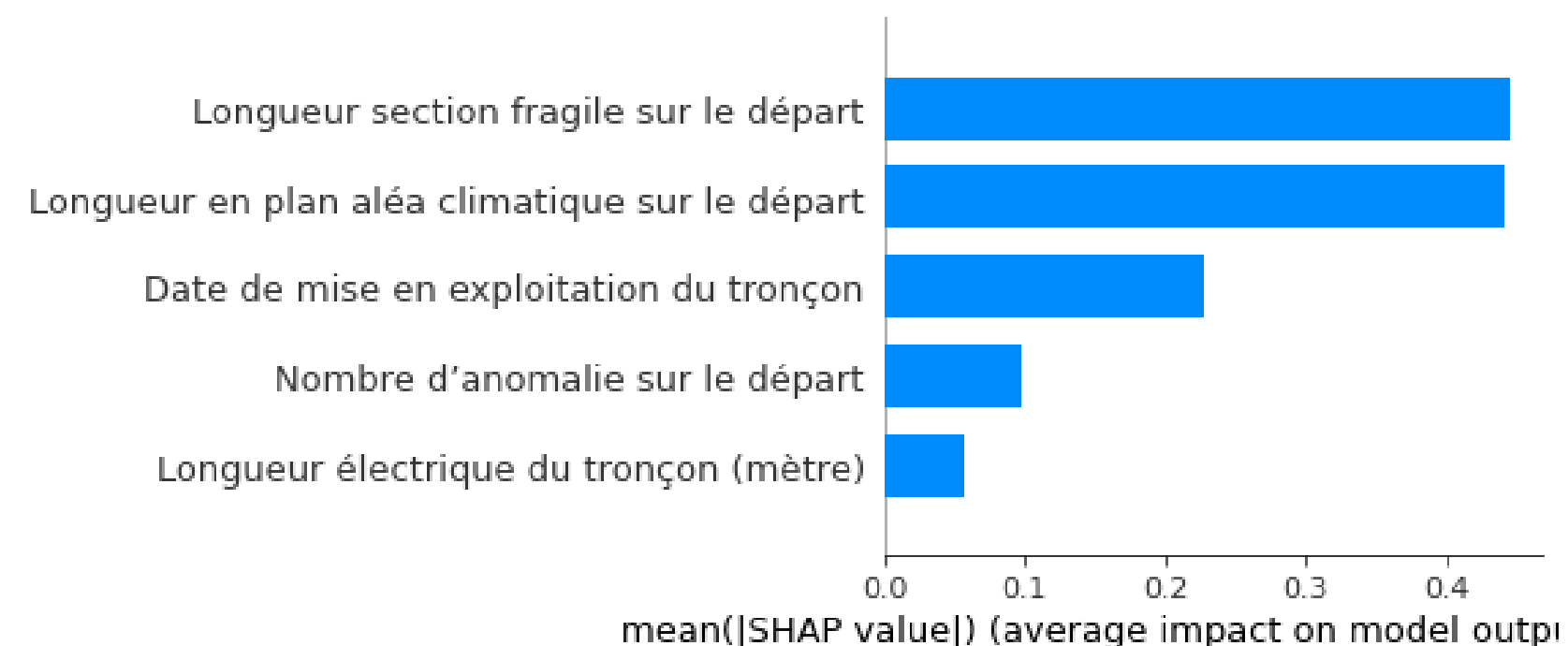
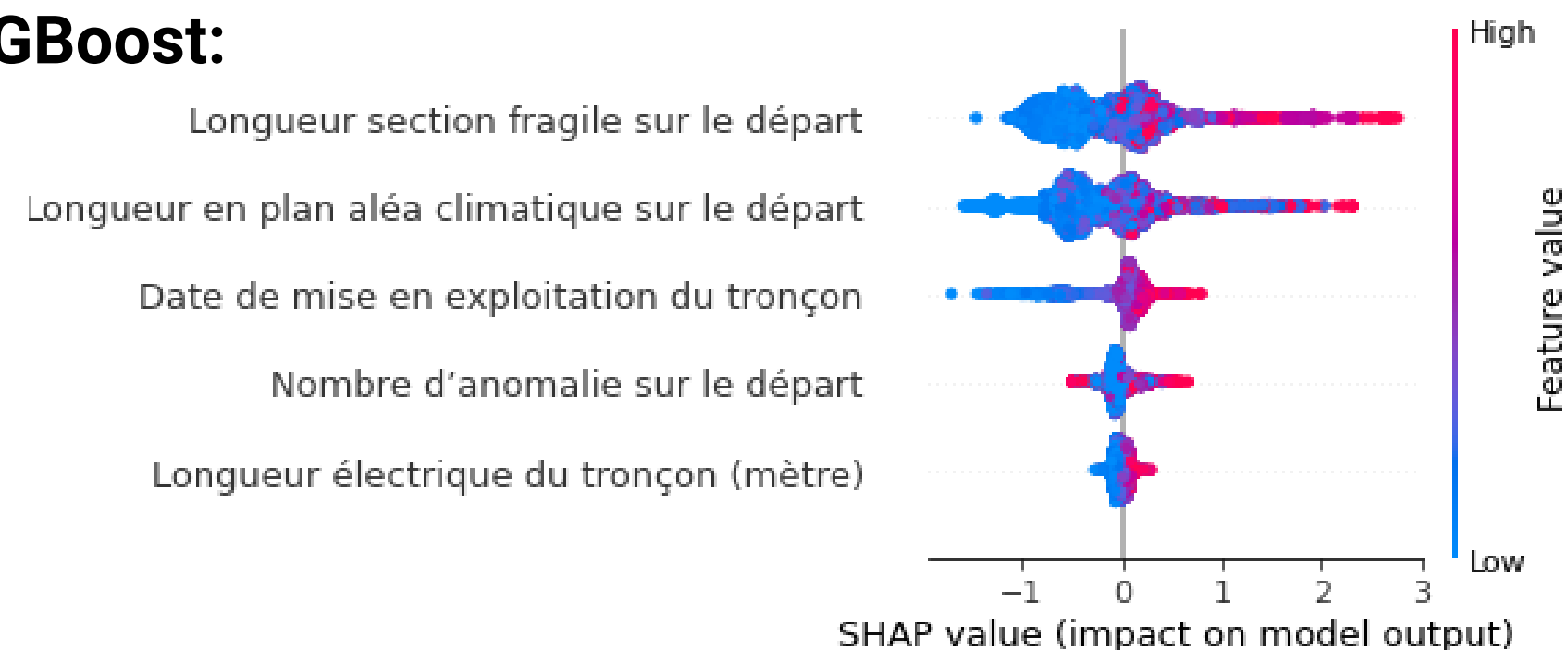
- Le taux d'incidents bien détectés par nos modèles dépend d'un **choix arbitraire**.
- Le **seuil de probabilité** auquel on juge qu'il y aura un incident peut dépendre de paramètres économiques par exemple.
- Le seuil peut aussi être déterminé avec une **limitation du nombre de km survolés**.

2.2 Interprétations

Logistique :

- Lorsque l'ancienneté, la longueur du plan climatique ou bien encore la longueur de la section fragile **augmentent, la probabilité d'avoir un incident augmente.**
- Toutes choses égales par ailleurs, si notre tronçon vieillit de **1 an**, la probabilité d'avoir un incident sur ce dernier augmente de **1,42%**.
- Lorsque la longueur en aléa climatique augmente de **1 mètre** pour un tronçon alors la probabilité d'avoir un incident augmente de **5.24%**.

XGBoost:



2.2 Focus

Voici un extrait des tronçons avec les probabilités de faire un incident **les plus élevées** :

- 89768
- 27531
- 25522
- 35189
- 74792
- 40234
- 59105
- 92246
- 88872
- 76138

De cette manière, nous pouvons choisir de cibler les tronçons les plus importants, de sorte à ne pas dépasser **25 km de diagnostique**.

3. CONCLUSION

3.1 Limites

3.1 Limites

- Des **données temporelles** aurait pu améliorer la performance de nos modèles. Il aurait été intéressant d'observer **l'évolution de l'état du réseau dans le temps**. Les modèles de **séries temporelles** permettent notamment de prendre en compte l'impact des événements passés sur le présent, et ce en maintenant une **forte interprétabilité**.
- L'intégration de **variables climatiques géolocalisées** donnerait potentiellement de l'information supplémentaire afin de cibler les zones davantage exposées aux **aléas climatiques**.



CONTACTS



Lorenzo BARRAUD

barraudlorenzopro@gmail.com



Léo BRIAND

leobriand35@gmail.com



Simon MIRZA

simon.mrza@gmail.com



Karl SONDEJI

karl.sondeji10@yahoo.com

