
PROJET DATA

Présentation de l'entreprise Umbrella 4 U



Utilisation des données météorologiques pour optimiser l'emplacement

Analyse des données météorologiques



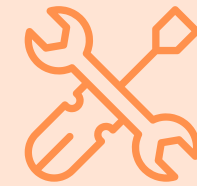
Placement stratégique des distributeurs dans les zones fréquentées

Gestion de l'emplacement



Optimisation du stock pour maximiser la disponibilité et les ventes

Gestion des stocks



Entretien régulier pour assurer le fonctionnement des distributeurs

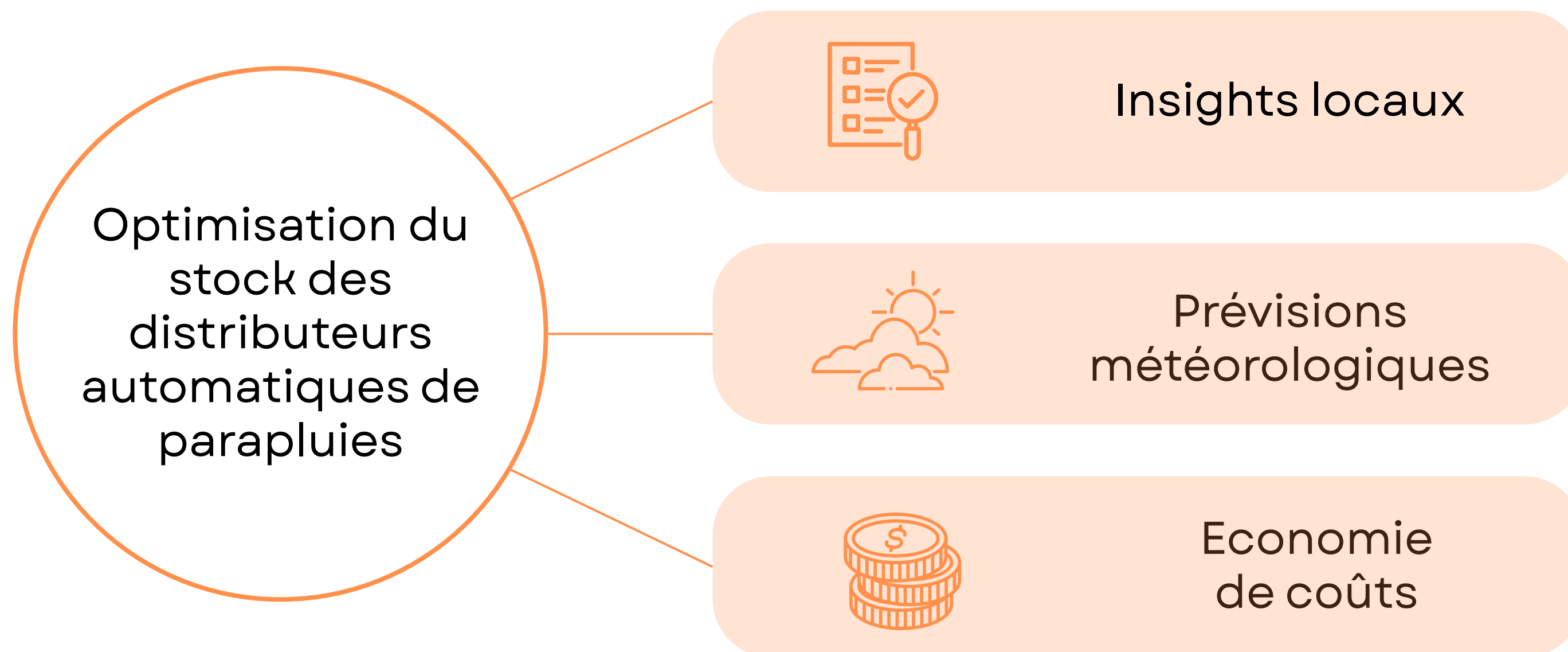
Maintenance



Réseau de distributeurs optimisé

Problématique business

Maximiser les ventes de parapluis grâce aux données météorologiques



Page 10 of 10



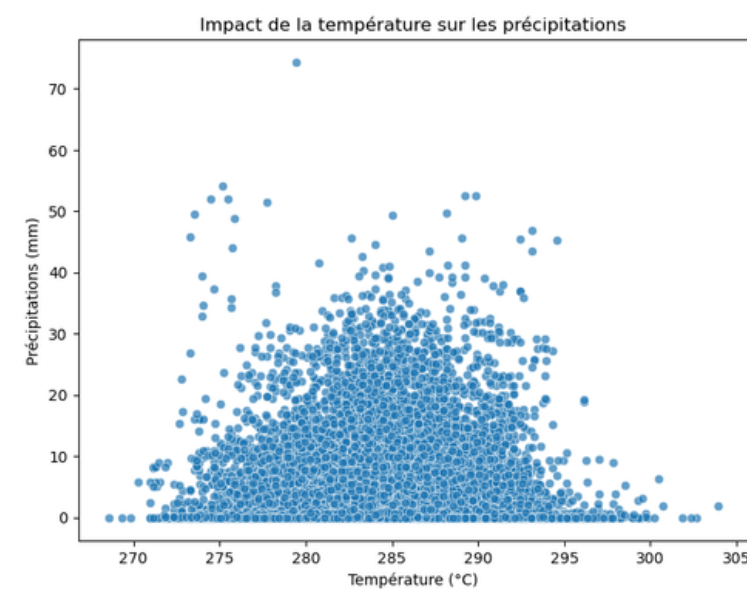
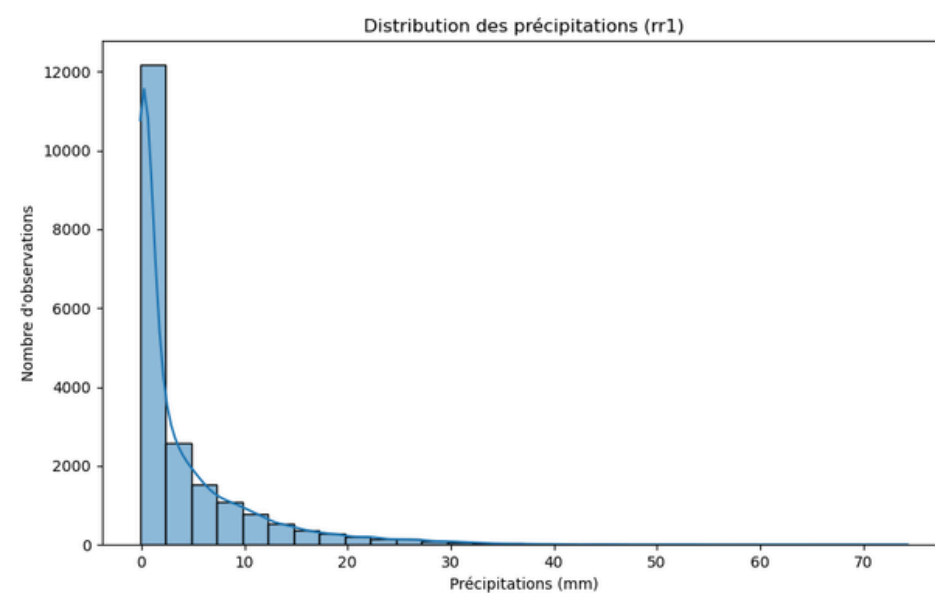
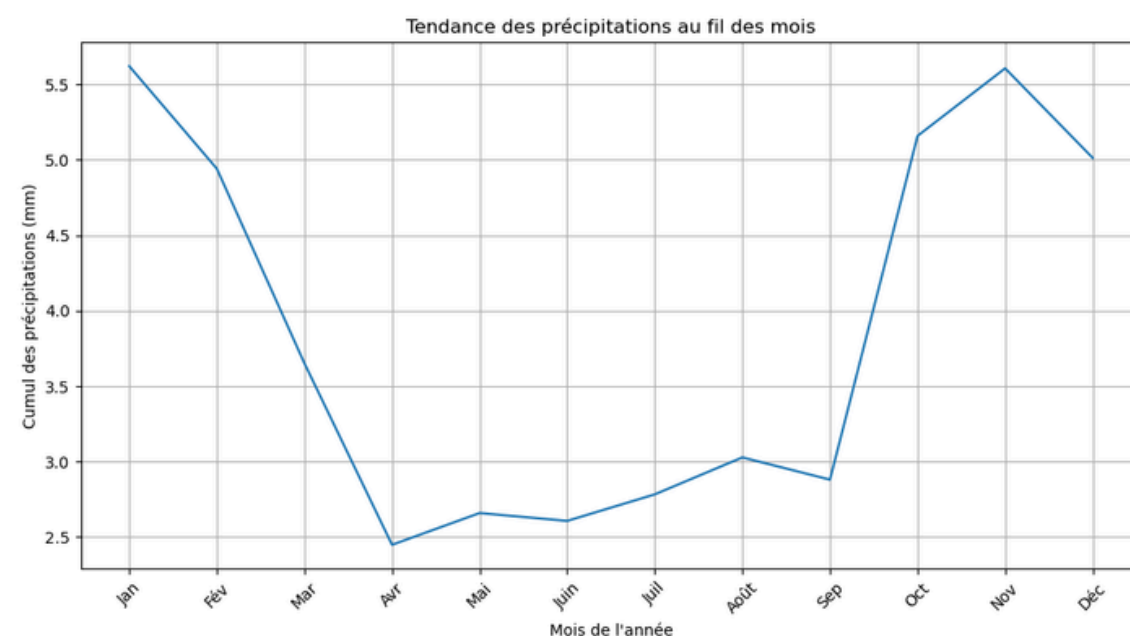
- Données d'observations issues des messages internationaux d'observation en surface (SYNOP) circulant sur le système mondial de télécommunication de l'Organisation Météorologique Mondiale.
- Mise à jour quotidienne



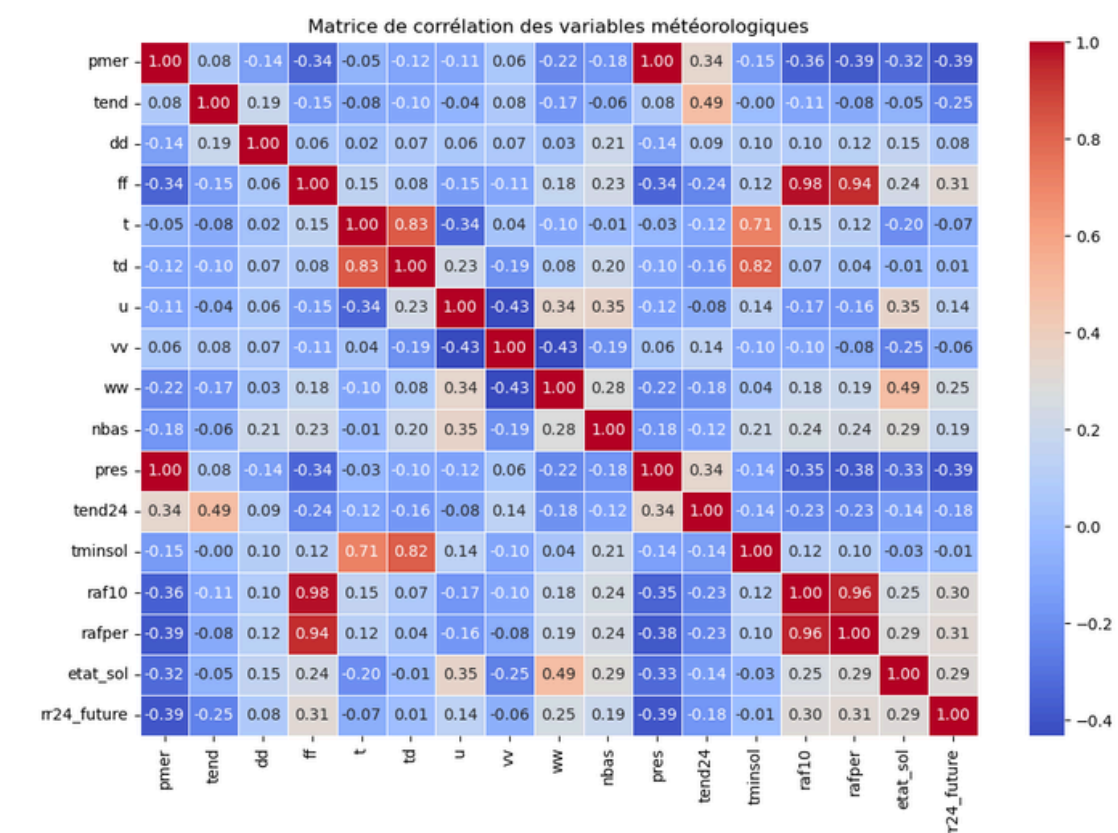
- Paramètres atmosphériques mesurés (température, humidité, direction et force du vent, pression atmosphérique, hauteur de précipitations) ou observés (temps sensible, description des nuages, visibilité) depuis la surface terrestre. Selon instrumentation et spécificités locales, d'autres paramètres peuvent être disponibles (hauteur de neige, état du sol, etc.)

Analyse exploratoire

Représentations graphiques



Matrice de corrélation



- Élimination de variables non pertinentes pour la prédiction rr24_future
- Variables dont la corrélation est inférieure à 25% éliminées. Sauf la variable mois

Pipeline de données

3 ÉTAPES MAJEURES



REQUÊTE API

Récupération des données JSON depuis une base publique



TRANSFORMATION

Conversion en DataFrame Pandas (tableau structuré)

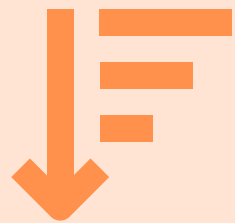


IMPLÉMENTATION

Rajout des dernières datas manquantes

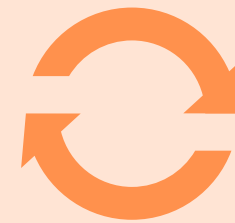
Le site ne permet pas de télécharger plus de 100 données à la fois : nécessité de ne prendre que les dernières données

Modèle de Machine Learning



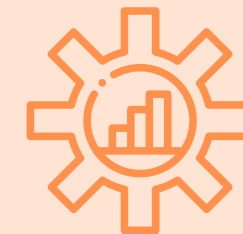
Collecte et préparation des données

- Chargement des données météorologiques
- Définition de la variable cible (rr24_future)
- Séparation des features et de la cible



Entraînement et optimisation

- Utilisation de Random Forest
- Optimisation des hyperparamètres via GridSearchCV
- Accuracy avec seuil de tolérance



Evaluation et sauvegarde du modèle

- Calcul de plusieurs métriques
- Enregistrement du modèle si il est meilleur que le précédent

Prévision des précipitations

Utiliser le modèle entraîné pour prédire la quantité de pluie dans 24h à partir des dernières observations météorologiques d'un département donné



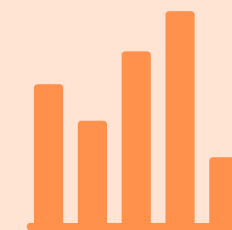
Téléchargement
et mise à jour les
données
météorologiques



Chargement du
modèle et
récupération des
features utilisées

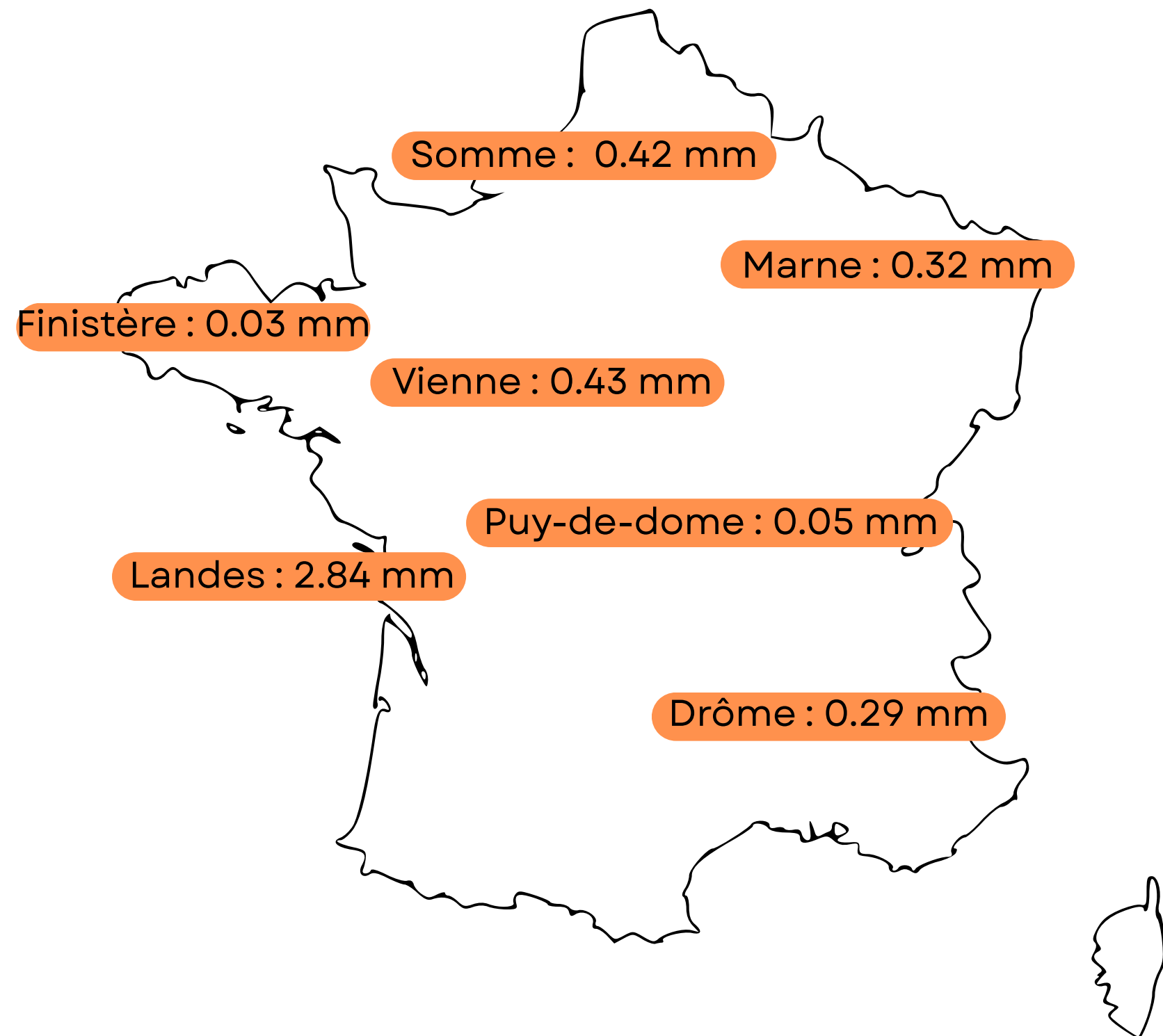


Vérification de la
présence des
features
nécessaires



Réalisation d'une
prédiction de la
quantité de pluie
prévue dans 24h

Résultats et performances



MÉTRIQUE CIBLE : ACCURACY

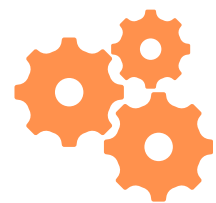
Modèle Landes	81.88 %
Modèle Vienne	86.78 %
Modèle Puy-de-dôme	89.34 %
Modèle Marne	87.94 %
Modèle Marne	84.21 %

Limites et perspectives

Limites du modèle



Lenteur de l'API liée à la base de données : le téléchargement est long



Possibilité d'utiliser des modèles de ML plus complexes

Pistes d'amélioration



Comparer la performance de RandomForest à celle d'autres modèles



Utilisation d'une base de données avec des localisations plus précises

Phase d'industrialisation

1

Récupération des données météorologiques - [pipeline.py](#)

- Collecte des données pour un département sous la forme d'un Dataframe
- Stockage dans un fichier csv sous BDD/data_<departement>.csv

2

Entraînement du modèle de Machine Learning - [ml.py](#)

- Préparation des données et entraînement du modèle
- Identification du modèle le plus performant (accuracy)

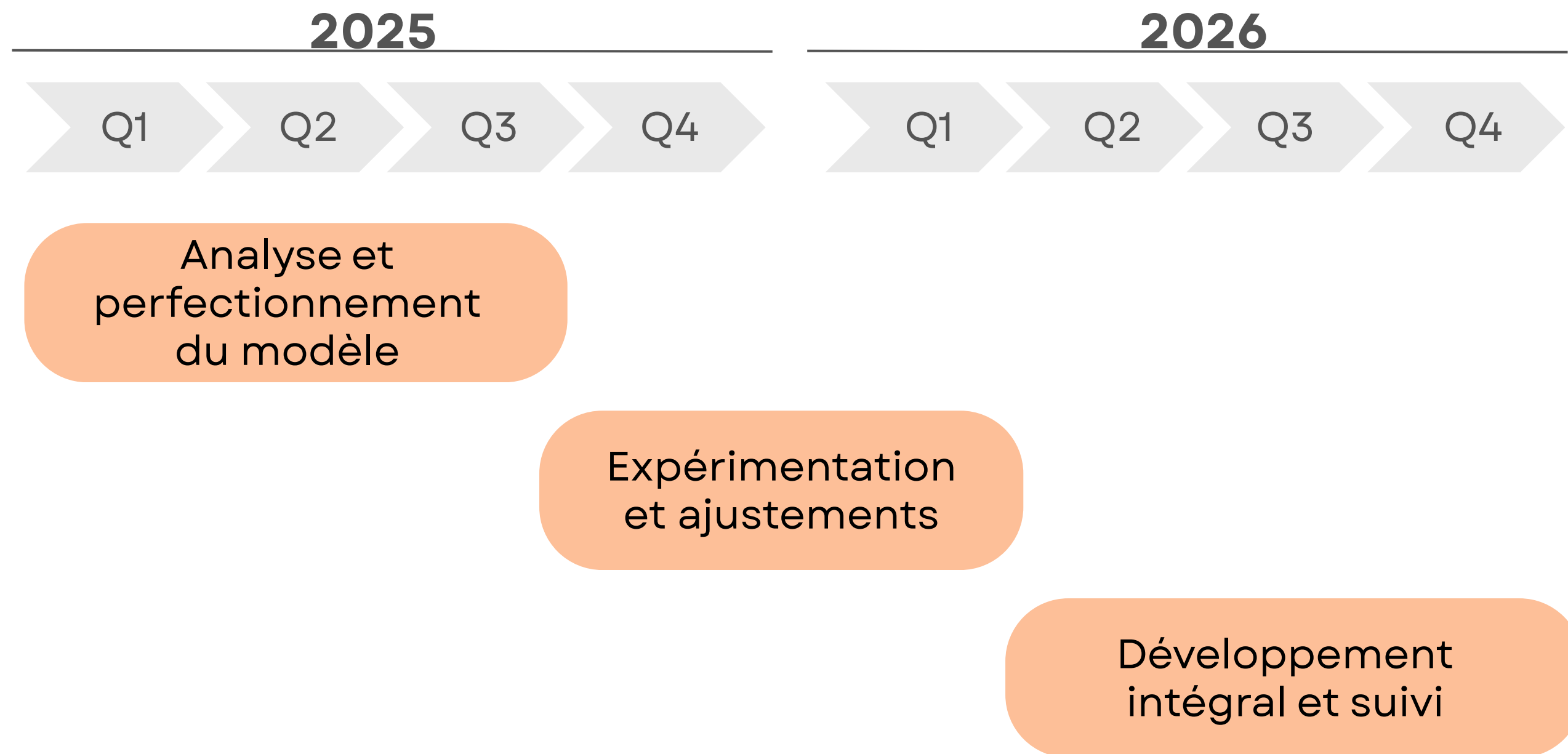
3

Prédiction des précipitations - [main.py](#)

- Prédiction sur les données les plus récentes avec le modèle entraîné
- Processus automatisé de la récupération des données à la prédiction

Exemple de sortie - “Prédiction de pluie dans 24h pour le Finistère : 6.4mm”

Implémentation et road map



Merci pour votre attention !