

SaaS solution optimizing water-intensive production



33%

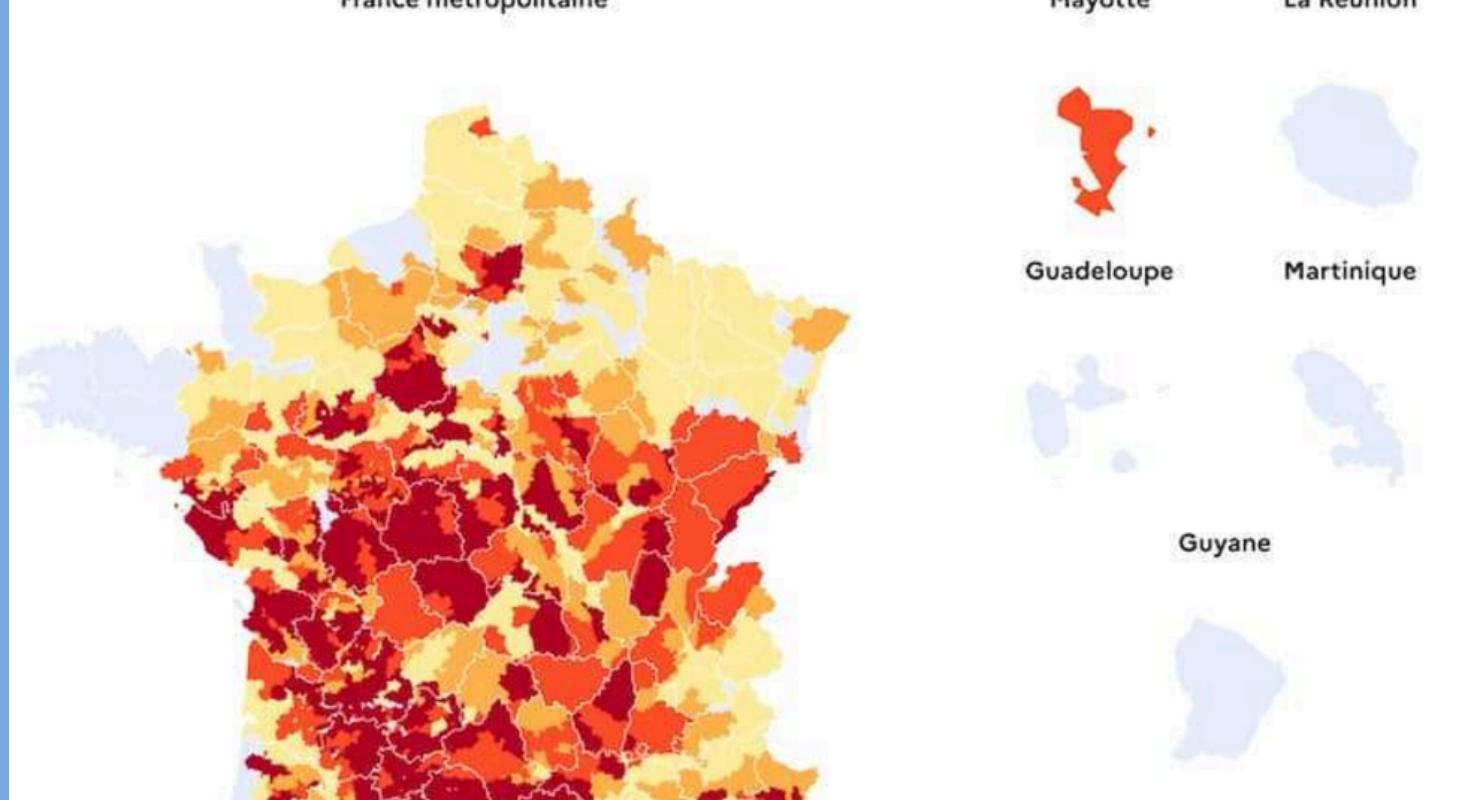
Industrial sector's need



1.01 tn

French GDP impacted

How Essential Groundwater is for water-intensive Industrials?



Current Situation in France:

93 out of 96

departments with

Summer Water Restrictions

on industrial production



Save around 10% revenue potential loss with our SaaS solution!

Target Clients:

Most impacted industries according to UNESCO

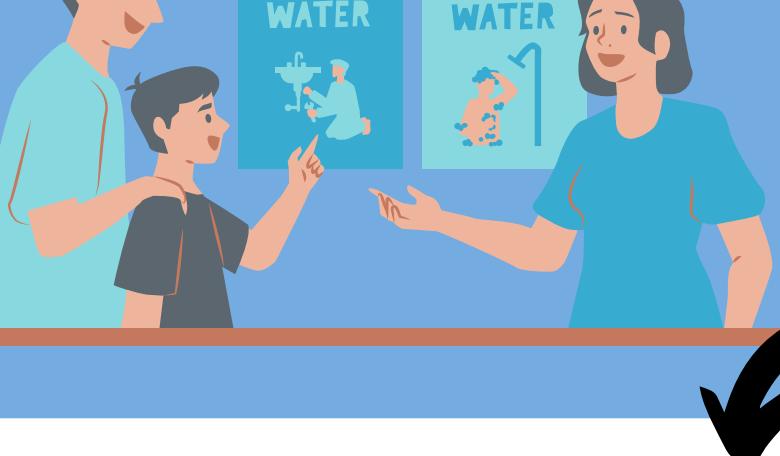
Construction **Energy** **Tech**
Mining **Generation**

Manufacturing
Oil & Gas
Chemicals



Business Solution

Predict groundwater shortages, plan smarter, cut losses, and secure sustainable production—all seamlessly integrated into your ERP.



01

Prediction of water shortage for the next quarter based on advanced machine learning technology

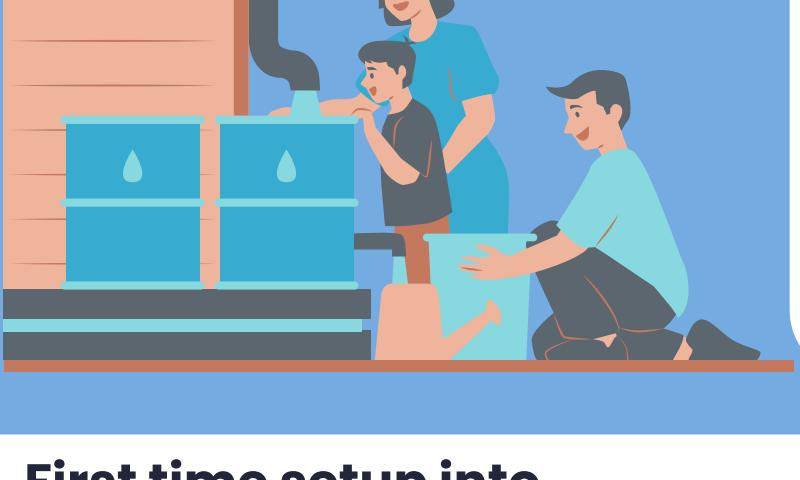
02

Optimized production planning based on all geographical location of factory sites with our SaaS integrated in your ERP system



03

Long-term sustainable solution of water recycling enhancing sustainable water management



First time setup into ERP system: € 50,000

3 engineers, 2-3 months onsite

Ongoing subscription cost: € 10,000/mo

- Water Shortage Prediction Map with the locations of all factory sites and optimized production planning
- Analysis on contingency and sensitivity cases, 1 data scientist to consult at all time
- New factory location picking simulation

Sustainable Water Management Solution: Ad-hoc cost

Pricing & Services

Targeting clients with water-intensive production

Our USPs



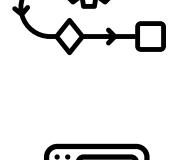
First Private Company apart from public players



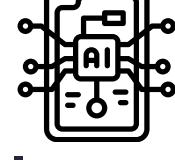
Affordable services with huge financial rewards



Long-term sustainable solution with positive ESG spillover effect



Fully automated of production plan across all factories



Advanced technologies with Machine Learning and Time-series Deep Learning Models

Scientific Approach

Data Overview

Data Cleaning

Model Selection

Best Model

Result & Discussion

Time-series data of watertable, meteo, hydrology, water withdrawal, economic and distance characters for underground water stations.

With data spanning from 2020-2023, excluding 2022 & 2023 summer months.

The prediction target is the groundwater (aka. watertable) level as a categorical variables with 5 values: Very high, High, Average, Low, Very low; of 2022 & 2023 summer months.

The goal is to predict the groundwater level in summer months with the least availability.

Two data cleaning solutions:

Solution I:

Step I: Removing features >20% missing value

Step II: Remove redundant columns refering to the same meaning (e.g. code & name)

Step III: Fill null value with mean for numerical varaibales, and mode for categorical variables

Narrow down 136 features to **60 features**

Among features left: Piezo 14; Meteo 16; Hydrology 9; Withdrawal 14; Economic 5; Distance 2.

Solution II:

Step I: Understanding the features item by item, combined with groundwater level management domain knowledge learnt shortly.

Step II: Removing features >60% missing value

Step III: Same method as above of filling null

Narrow down 136 features to **22 features**

Among features left: Piezo 6; Meteo 5; Hydrology 1; Withdrawal 4; Economic 4; Distance 2.

No duplicates exist.

- Categorical target prediction:

Model I: XGBoost

Pro: Robustness and accuracy in tabular data.

- Time-series neural network:

Model II: LSTM

Pro: Specialized for sequential data, robust predictions for dynamic

- Hybrid model:

Model III: Ensemble Learning

Underlying models: Random Forest, XGBoost, LSTM

Pro: Accurate and reliable due to multiscale feature extraction

Due to the limited time of Hackthon, Ensemble Learning is not trained

Main techniques: Cross-validation, Grid search for tuning

- XGBoost
- **Cross-validation** to prevent overfitting
 - **5-fold** cross validation
 - **Validation Set only include summer months data**
- **Hyperparameters through grid search:**
 - tree_method='hist', eval_metric='mlogloss', n_estimators=1000, max_depth=10, learning_rate=0.09, subsample=0.8, colsample_bytree=0.8, gamma=1, reg_alpha=0.1, reg_lambda=1
- **Evaluation metrics (F1 Score):**
 - 0.91 on training set
 - **0.57 on test set**

- Final Submition: **57% F1 score**

- Random Guess Benchmark: **20%**

Large boost in prediction peformance compared to benchmark.

Possible Reasons of prediction performance:

1. Limited information on summer groundwater level in the training set, might hinder prediction performance

2. Summer groundwater level is generally lower and has different patterns compared to other seasons.

3. Best model in literature review* has only 59% absolute error <0.5m in prediction considering complexity.

Future Potential

Well positioned to expand internationally & in multiple industries

- With core technology based on public dataset available for all countries, our startup is well-positioned to expand internationally.
- Targeting clients in water-intensive production industries (Energy generation, Tech, Construction, Chemicals, Oil & Gas, Mining, Manufacturing etc.), touching a large market of \$20.35bn global market.

Partnership with government and international institutions

- Collaborated with UNESCO World Water Forum to provide essential humanitarian aid to Africa .

Building Ensemble Learning Model to boost performance:

- Combining the strengths of both robustness non-linear tree models and sequential memory-based model



Our Team

Junior but ambitious & innovative team

2 M1 students
2 L3 students

Multicultural Background

2 French
1 Algerian
1 Chinese

Diverse Academic Background

3 Engineering
1 Data & Business

Gender Collaboration

3 males
1 female



**Best Team
Best Vibe**

Yiqin CAO
Yasser Bouhai
Alexandre Decavèle
Alexandre Foulquier