

# **Document de Cadrage : Cockpit de Maintenance Prédictive.**

## SOMMAIRE :

1.	Présentation du Projet et Contexte Métier .....	2
2.	Identification des Utilisateurs et Scénarios d'Usage.....	3
3.	Traduction en Problème IA.....	3
1.	Modèle 1 : Classification binaire (Alerte imminente) .....	4
2.	Modèle 2 : Régression (Estimation du RUL) .....	4
3.	Logique de Scoring : Catégorisation du Risque .....	4
4.	Besoins en Données et Fréquence .....	4
5.	Objectifs Mesurables et Critères d'Acceptation.....	5

## 1. Présentation du Projet et Contexte Métier

EnergiTech est une société anonyme fondée le 12 mars 2012 et filiale du groupe GreenPower Holding. Son siège social se situe à Paris. En plus du siège, l'entreprise possède trois sites opérationnels : un centre de production à Lyon, un parc solaire à Nantes et un hub dédié aux données à Strasbourg. Ces implantations permettent à EnergiTech de couvrir l'ensemble du territoire français tout en restant proche de ses principaux clients industriels et institutionnels.

Outre la vente d'électricité, EnergiTech propose des services de maintenance prédictive aux exploitants de ses parcs. La feuille de route d'EnergiTech prévoit de doubler sa capacité installée pour atteindre 250 MW d'ici 2028, tout en automatisant 80 % des opérations de maintenance grâce à l'introduction de modèles d'intelligence artificielle capables de prédire les pannes. L'objectif ultime est d'atteindre la neutralité carbone sur l'ensemble de ses sites d'ici 2030, en combinant optimisation de la production, stockage d'énergie et amélioration continue des processus grâce à l'analyse de données.

Actuellement, l'usage de l'IA chez EnergiTech est artisanal et fragmenté : les modèles de classification et de régression sont exploités manuellement au cas par cas via des interfaces de test simplifiées. Cette approche empêche d'avoir une vision globale du parc, rendant impossible la priorisation objective des interventions techniques. Sur le plan opérationnel, l'absence de monitoring et d'observabilité signifie que les erreurs de calcul ou les indisponibilités du service ne sont pas détectées en temps réel. Sans chaîne CI/CD, le déploiement de nouvelles versions reste risqué et non reproductible. Enfin, ce manque d'industrialisation bloque l'objectif stratégique de l'entreprise : automatiser 80% de sa maintenance d'ici 2028.

## 2. Identification des Utilisateurs et Scénarios d'Usage

Les principaux utilisateurs cibles sont les membres du service de maintenance :

- Responsable Maintenance : Doit avoir une vue d'ensemble du parc pour allouer les ressources et prioriser les sites.
- Chef d'équipe / Technicien : Doit consulter la liste des éoliennes à risque élevé pour la semaine à venir et programmer les interventions techniques.

- Scénario principal : "Consulter le tableau de bord quotidien pour identifier les machines présentant un risque critique et justifier la priorité d'intervention via le score de risque et le temps estimé avant panne".

## 3. Traduction en Problème IA

### 1. Modèle 1 : Classification binaire (Alerte imminente)

Ce modèle analyse les flux de données des capteurs pour déterminer si une éolienne risque de tomber en panne dans un délai strict de 7 jours. Il fournit une probabilité de défaillance qui sert de signal d'alarme immédiat pour les équipes opérationnelles.

C'est l'indicateur de "danger à court terme" qui permet de déclencher des inspections de sécurité en urgence avant une casse majeure.

### 2. Modèle 2 : Régression (Estimation du RUL)

Ce modèle calcule la "Remaining Useful Life" (RUL), soit le nombre précis de jours ou d'heures restant avant la panne théorique de la turbine. Contrairement à la classification, il offre une vision quantitative qui facilite la planification logistique et l'ordonnancement des pièces de rechange. Il permet de passer d'une maintenance réactive à une stratégie d'anticipation basée sur l'usure réelle des composants.

### 3. Logique de Scoring : Catégorisation du Risque

Le système orchestre les deux modèles pour classer chaque machine selon une hiérarchie de priorité:

- Critique : Cumul d'une prédition de panne sous 7 jours et d'une RUL très faible (ex: < 3 jours), nécessitant une intervention immédiate.
- Élevé : Probabilité de panne significative ou RUL comprise entre 4 et 10 jours, demandant une programmation prioritaire dans la semaine.
- Modéré / Faible : Aucun signal de panne à court terme et RUL confortable, nécessitant une simple surveillance standard via le tableau de bord.

## 4. Besoins en Données et Fréquence

- **Données d'entrée (Capteurs et Logs) :** L'application ingère les flux massifs provenant du système SCADA (température, vitesse du vent, puissance

produite) ainsi que les journaux de supervision technique. Avec un volume critique de 12 To de données générées annuellement, l'architecture doit inclure une étape de prétraitement robuste pour nettoyer et normaliser ces informations avant qu'elles n'alimentent les modèles de classification et de régression.

- **Fréquence de mise à jour (Batch quotidien)** : Le cycle de décision de la maintenance ne nécessite pas un monitoring à la seconde, mais une analyse fraîche chaque matin. Le scoring du parc est donc réalisé via un traitement par lot (batch) nocturne. Ce processus automatise l'extraction des données de la veille, lance les prédictions et met à jour le cockpit de pilotage pour que les équipes opérationnelles disposent d'un état des lieux précis dès leur prise de poste.

## 5. Objectifs Mesurables et Critères d'Acceptation

- **Indicateurs fonctionnels (Le "Top N" à risque)** : Le cockpit doit fournir en temps réel une liste ordonnée des N éoliennes présentant le risque de panne le plus élevé. Cet indicateur transforme des probabilités mathématiques en un outil d'aide à la décision concret, permettant aux responsables de maintenance de visualiser immédiatement quels actifs nécessitent une inspection prioritaire pour éviter une casse coûteuse.
- **Contraintes techniques (Performance et Latence)** : L'industrialisation impose des limites de performance : le temps de calcul total du scoring pour l'ensemble du parc doit rester inférieur à une durée définie (ex: 2 heures) pour garantir la fraîcheur des données. Par ailleurs, l'API de consultation doit offrir un temps de réponse optimal pour assurer une expérience fluide aux utilisateurs, même lors de l'accès à des volumes de données historiques importants.
- **Critères d'acceptation (Précision et Confiance)** : La solution sera validée si elle parvient à identifier les turbines ayant un temps avant panne (RUL) inférieur à X jours avec un taux de fiabilité défini lors de la phase de test. Un niveau de confiance minimal (ex: 85% ou 90%) est exigé pour chaque alerte afin de crédibiliser l'outil auprès des techniciens et d'éviter les interventions inutiles déclenchées par des faux positifs.