

RAPPORT PROJET GPI :

MAINTENANCE PRÉDICTIVE

2022 - 2023



Réalisé Par :

- Matouk Afaf
- Guerrab Mouna
- El Jaouhari Mohamed
- El Amrani Ilyas

Encadré Par :

Pr. ER-RATBY Mohamed

Sommaire :

1.Introduction

2.Science des données et
maintenance predictive

3.Jeu des données utilisées

4.Implémentation

5.Visualisation des résultats

6.Vision

7.Conclusion

Annex A : Prétraitement & Modèle

Annex B : KPIs

1. Introduction :

La maintenance prédictive utilise un ensemble des méthodes pour prédire quand une machine ou un équipement est susceptible de tomber en panne, et planifier de manière proactive la maintenance pour éviter cette panne. Cette approche proactive de la maintenance permet de réduire les temps d'arrêt, d'augmenter l'efficacité et de prolonger la durée de vie de l'équipement.

Traditionnellement, la maintenance était effectuée selon des calendriers établis ou lorsque l'équipement présentait des signes de défaillance. Cette approche passive peut entraîner des temps d'arrêt inutiles, une augmentation des coûts et une perte de productivité. La maintenance prédictive, quant à elle, utilise les données collectées à partir de capteurs, de journaux d'appareils et d'autres sources pour identifier les signes avant-coureurs d'une défaillance de l'appareil et prévoir quand une maintenance est nécessaire.

La maintenance prédictive peut être appliquée à une variété d'industries, y compris la fabrication, le transport, l'énergie et la santé, et gagne en popularité avec l'avènement de l'internet des objets (IoT) et la disponibilité des outils d'analyse de données volumineuses (Big data analytics).

2. Science des données & maintenance predictive :

La science des données joue un rôle clé dans la maintenance prédictive. La maintenance prédictive repose sur l'analyse de grandes quantités de données provenant de diverses sources telles que les capteurs de périphérique, les journaux de maintenance et les données de performances historiques. Les scientifiques des données peuvent analyser ces données à l'aide d'algorithmes d'apprentissage automatique et de techniques analytiques avancées pour identifier les modèles et les anomalies qui peuvent indiquer une défaillance imminente de l'équipement.

L'application de la science des données à la maintenance prédictive aide les organisations à mieux comprendre l'état des équipements, à identifier les signes avant-coureurs de problèmes potentiels et à déterminer à quel moment la maintenance doit être effectuée pour éviter les pannes d'équipement. La science des données aide également à optimiser les calendriers de maintenance et à effectuer la maintenance au bon moment pour minimiser les temps d'arrêt et maximiser l'efficacité.

3. Jeu des données utilisées :

Notre jeu de donnée se manifeste en une entreprise qui dispose un appareil qui envoie quotidiennement des lectures de capteurs. Elle souhaite développer une solution de maintenance prédictive qui identifie de manière proactive le moment où la maintenance doit être effectuée. Cette approche promet des économies par rapport à la maintenance préventive régulière ou basée sur le temps. En effet, la tâche ne s'exécutera que si elle est légitime.

Nos données sont disponible sous forme d'un fichier Excel contenant la date de chaque enregistrement, avec des mesures captées lors d'exécution de tache, accompagné par une colonne présentant si la machine a échoué ou non pendant la journée.

à propos des dates, nos enregistrement sont propres à l'année 2015, à partir de 01/01 jusqu'à 01/12.

4. Implémentation :

Pour pouvoir implémenter nos idées et construire ce système, nous allons découper notre projet en 5 étapes :

- Analyser et évaluer notre jeu de données
- Développer les fonctionnalités nécessaires pour le fonctionnement du système.
- Développer le modèle final.
- Calculer les indicateurs de performance.

Le but de cette partie est d'expliquer comment implémenter informatiquement les idées discutées précédemment, ainsi qu'expliquer les solutions proposées pour faire face aux différentes contraintes du problème en terms de données, de calcul et des technologies utilisées.

La suite est l'explication du notebook joint à ce document (annex A) , nous présenterons des commentaires et des explications tout au long de la lecture des différentes sections de la réalisation.

4. Implémentation :

4.1 Analyse et évaluation du jeu de données:

La première étape consiste à analyser et évaluer notre jeu de données pour savoir quels variables de décisions qu'on peut utiliser, cette partie concerne les sections 1 jusqu'à 4 (Data Transformation).

Nos données sont de forme tabulaire avec les colonnes "date", "device", "failure", "metric1", ..., "metric9", la colonne "date" décrit la date avec le format standard "JJ/MM/AAAA", la colonne "device" est le nom de l'appareil ou la machine dont on mesure les différents metrics de 1 à 9, finalement la colonne "failure" est une variable binaire prenant la valeur 1 s'il y avait une panne ce jour ou non.

La mesure des métriques est supposée quasi-constante durant toute la journée.

Les conclusions principales de notre analyse sont :

- La rareté des pannes : On a 106 enregistrement de pannes, contre 124388 enregistrement sans panne.
- Le graph du code [142] montre clairement que le nombre d'enregistrements sur les machines devient de moins en moins fréquent après chaque mois, confirmant les données numériques du tableau [141].

- Le code [145] montre la diminution de nombre d'enregistrements sur les machines le plus on avance dans l'axe des temps.
- Le code [144] montre que nos données sont daté entre le 1 Janvier 2015 et le 2 Novembre 2015
- Le code [155] permet de supprimer la colonne "metric8" puisque elle a des valeurs égaux à ceux de la colonne "metric7".
- Nous avons pu rendre nos données sous une forme plus intéressante pour un modèle de machine learning en exposant de nouvelles colonnes qu'on peut utiliser dans le but de prédire nos pannes comme la colonne "ActiveDays" qui permet de calculer en cumulation combien de jours la machine était elle active, "week_day" et "month" pour numérisé la date vue que le type date est difficile à traiter par un algorithme de machine learning sans transformation, et la colonne "failure_before", pour savoir si dans l'enregistrement actuelle la machine a-t-elle déjà tombé en panne au moins une fois avant ou non.

Pour la suite, nous allons développer un pipeline, qui est un algorithme simple, permettant de mettre dans la forme finale de cette partie nos données en entré.

4. Implémentation :

4.2 Outils nécessaires au fonctionnement de système:

Le pipeline est un outil qui permet d'extraire et transformer les données pour les mettre en une forme souhaitée, qui rend les données prêt à la consommation par un algorithme quelconque, de machine learning dans notre cas.

Dans cette partie (5-6), le pipeline est développé et tester avec un algorithme de machine learning simple 'KNN', qu'on va expliquer par la suite.

Dans la première partie nous nous inspirons des traitement effectuées dans la première partie pour construire la fonction pipeline, qui prend en paramètre la base de données utilisée, l'enregistrement à transformer sous la forme d'une liste de valeurs, et un objet 'scalar' qui permet de mettre à la même échelle les métriques de l'enregistrement avec ceux de la base.

La totalité des procédures effectuées est expliquée en commentaires sur le code.

Pour la deuxième partie, nous préparons nos données pour tester notre pipeline, nous construisons à l'occasion un modèle de machine learning appelé KNN (Les K voisins les plus proches). Après le test et l'entraînement de notre modèle nous obtenons un score d'environ 97%, par suite, nous utilisons le pipeline pour transformer un enregistrement et nous effectuons la prédiction avec le modèle entraîné qui prédit 0, c'est à dire pas de panne.

4. Implémentation :

4.3 Synthétisation des données et création du modèle final:

Pour résoudre le problème de non balance de notre base de données, et pour avoir une situation dans laquelle la maintenance prédictive aura un effet remarquable, nous allons synthétiser une base de données en se basant sur la base de données actuelle, avec une distribution plus adoptée pour engendrer plus de pannes, nous concentrons aussi sur un seul type de machines, qui est le type "S1F0" d'après le code [177]. On synthétise dans un premier lieux les données de l'année 2015, on entraîne, de la même manière qu'avant, le modèle KNN sur ces données (code [232]), par la suite, nous générons une dataset pour l'année 2016 pour la validation de notre modèle, sur laquelle nous effectuant nos études. Les codes [235] et [236] sont utilisés pour prédire avec notre modèle les pannes et enregistrer les données sous format CSV, les résultats sont regroupées dans un tableau avec les colonnes "month", "failure" : C'est le nombre de pannes prédites et "realFail" est le nombre des pannes réels qui vont se produire si on effectue pas de maintenance prédictive. Ces résultats seront utilisés par la suite dans le calcul des indicateurs de performance.

4. Implémentation :

4.4 Calcul des KPIs:

Pour calculer les indicateurs de performance, on se base sur le tableau récupéré dans l'étape précédente et on effectue le calcul des KPIs dans le cas où la maintenance prédictive est effectuée, et le cas où elle n'est pas effectuée.

L'annexe B comprend l'implémentation et l'exécution de cet algorithme, bien expliqué avec des commentaires.

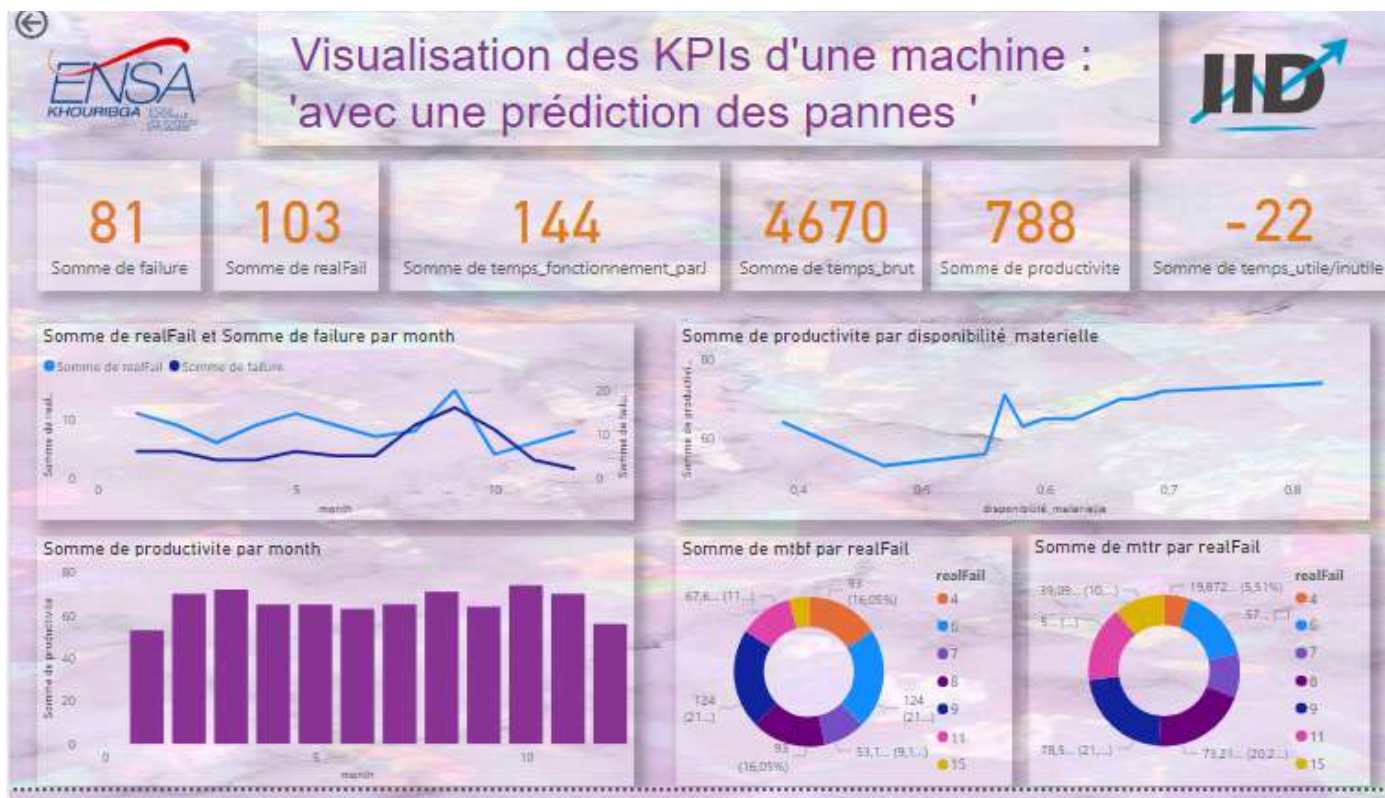
À l'issue de cette partie, nous disposons d'un fichier "CSV", de forme tabulaire, à utiliser pour créer notre tableau de bord.

5. Visualisation des résultats:

Question analytique :

Notre Visualisation permet de répondre à cet ensemble des questions dans le cas de la prediction de pannes et de sans prediction de pannes :

- 1-quel est l'evolution des realFail et failure durant chaque mois ?
- 2-quel est la valeur de la productivité durant chaque mois?
- 3-quel est l'evolution de mtbf et la disponibilité materiel pendant chaque mois ?
- 4-quel est l'evolution de mtrr et la disponibilité materiel durant chaque mois ?
- 5-l'evolution de mtbf et mtrr durant chaque mois ?
- 6-quel est le temps net pour chaque realFail?
- 7-quel est le temps brut pour chaque realFail?



Visualisation des KPIs d'une machine 'sans prédiction des pannes'

Nombre de pannes

103

Nombre de fonctionnement par jour

144

Temps brut de fonctionnement

3931

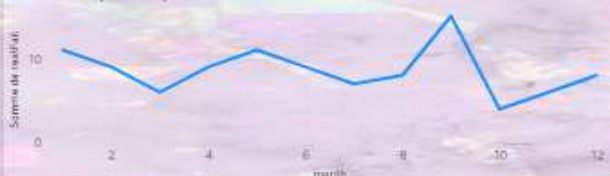
Productivité

697

Temps utile/inutile de fonctionnement

-103

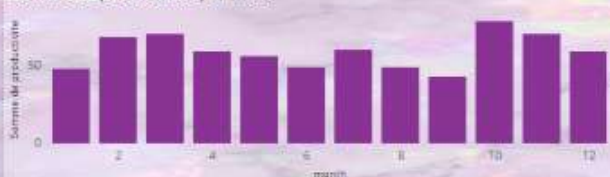
Somme de pannes par mois



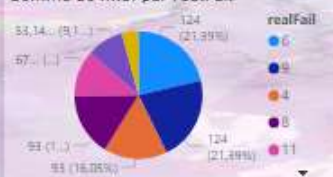
Somme de productivite par disponibilité_materielle



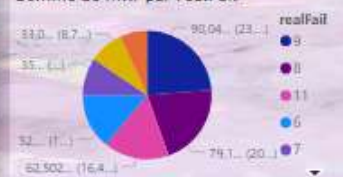
Somme de productivité par mois



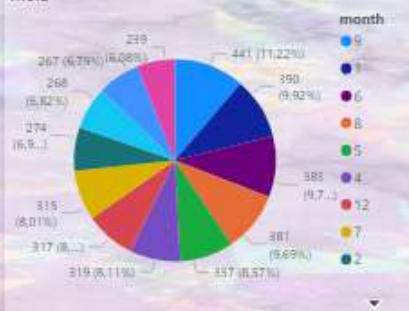
Somme de mtbf par realFail



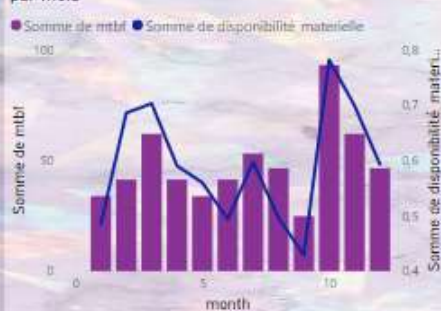
Somme de mtrr par realFail



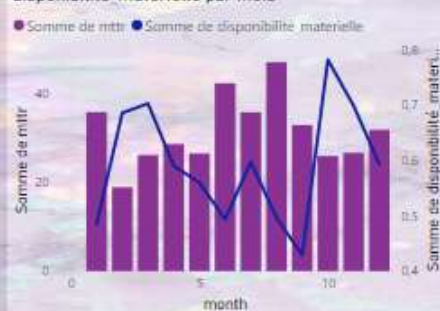
Somme de temps brut de fonctionnement par mois



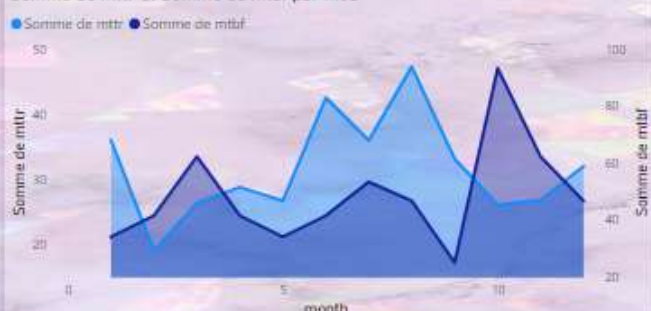
Somme de mtbf et Somme de disponibilité_materielle par mois



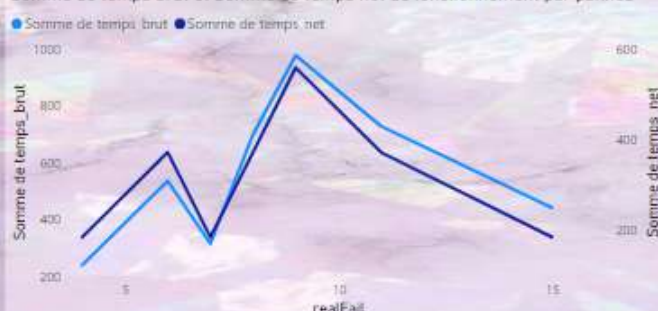
Somme de mtrr et Somme de disponibilité_materielle par mois



Somme de mtrr et Somme de mtbf par mois



Somme de temps brut et Somme de temps net de fonctionnement par pannes



6. Vision :

Ce projet est le premier pas pour la construction de tout un système, qui sera capable d'apprendre et s'adapter au plan de production et de maintenance de l'entreprise, un système plus complexe qui set de support aux responsables pour la prise des décisions et la planification des activités de maintenance et de la production.

Dans cette partie nous allons décrire une possibilité d'amélioration du projet et le processus avec lequel il peut être appliqué dans le domaine industriel réellement.

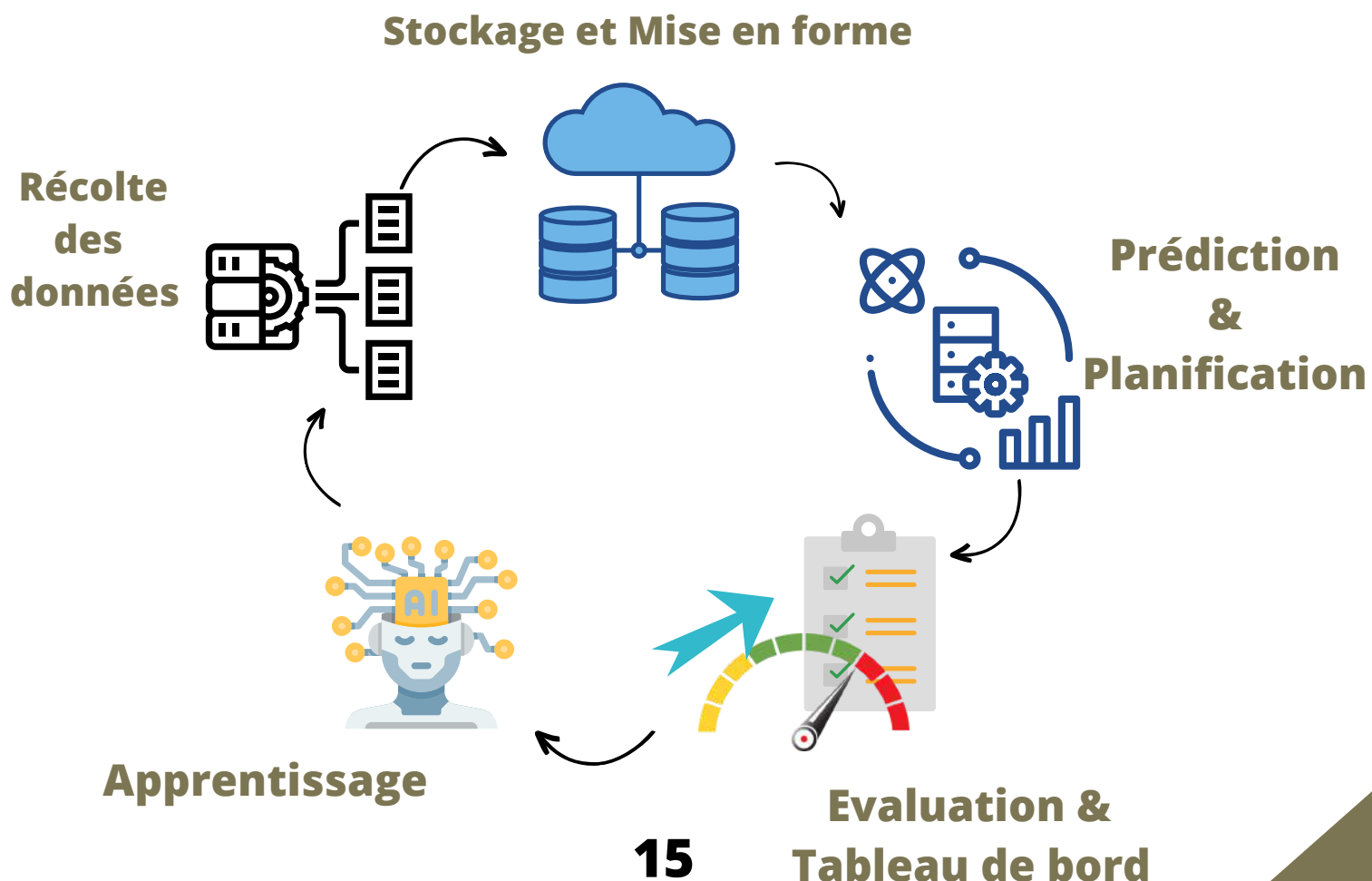
Le but de ce système et d'utiliser les sciences de données et l'informatique décisionnelle pour aider les responsables à décider sur les plans de maintenance prédictive et préventive au but d'améliorer la productivité.

Son mode de fonctionnement est constituer généralement de :

- Récolte des données : à l'aide des capteurs sur chaque machine, des rapports quotidiens et les contrôles effectués.
- Stockage et traitement des données: Au but de nettoyage et mise en forme et adaptation pour le modèle de machine learning (fait par le pipeline développé précédemment par exemple).

6. Vision :

- La prédiction : à l'aide du modèle de machine learning des pannes, et la réalisation des plans de maintenance.
- L'établissement des indicateurs de performance provisoires : en se basant sur les prédictions.
- Evaluation des résultats des prédictions et leurs effets: En mettons à jours le tableau de bords et les rapports.
- Entraînement du modèle de machine learning sur les résultats réels.



7. Conclusion :

La maintenance prédictive est devenu dernièrement une des piliers majeurs pour optimiser la maintenance de l'environnement industriel comme le cas des machines étudiée dans notre projet.

En la fusionnant avec la science des données, l'optimisation devient plus efficiente et plus économique à l'aide des outils mathématiques et informatiques avancées.

Annex A : Préparation des données et création du modèle