

*Projet de THESE CIFRE*

## ***Modèles d'Apprentissage automatique et d'Aide à la Maintenance Prédicative de Systèmes Avions***

**Lieu :** Clermont-Ferrand :

- HOP ! (Air France)
- LMBP (Laboratoire de Mathématiques Blaise Pascal)

**Encadrants :** Julie Trespeuch (HOP), Anne-Françoise Yao (LMBP), Nourddine Azzaoui(LMBP).

### **I. Contexte**

Ce projet de thèse s'inscrit dans la thématique de recherche MPSA (Maintenance Prédicative pour Système sur la flotte d'avion de Hop), collaboration LMBP-HOP en entamé en 2018 pour le développement d'outils de maintenance prédictive sur la flotte Embraer de Hop.

L'objectif de l'équipe de recherche LMBP-HOP est d'élaborer des solutions innovantes : modèles mathématiques, algorithmes d'apprentissage machine et/ou statistique ou d'intelligence artificielle pour une aide à la maintenance prédictive. Nos modèles sont construits à partir de différentes sources de données et informations fournis par les équipes techniques. Les données peuvent issues d'enregistreurs Fault **History Data Base** (FHDB), **Quick Access Recorder** (QAR) ou d'autres types des données telles que les rapports d'incident pilote, rapports de maintenance, vieillissement des pièces, etc.... Outre des données numériques classiques, les données peuvent être du texte, des séries temporelles, des images, une combinaison ou bien d'autres types de données. Ainsi le challenge de l'équipe est donc de répondre à des problématiques spécifiques pour chaque système avion. L'exploitation de ce type de données pour la maintenance prédictive reste marginale dans ce contexte. Notre démarche conduira à :

- 1) La modélisation des lois de probabilités des variables aléatoires (éventuellement à valeurs dans des espaces fonctions ou de formes) dont les observations sont fournies par les QAR. Ces modèles nous permettront de construire des tests statistiques pour la détection d'alerte et donc de prévision des pannes et par conséquent d'aide à la décision à la maintenance prédictive.
- 2) Une fois ces modèles éprouvés, une interface de maintenance prédictive sera ouverte pour permettre de guider la maintenance sur avion.

Ce travail devra être réalisé pour chaque système avion, par ordre de criticité au regard de nos objectifs primordiaux, à savoir la prédiction des défaillances des systèmes sur avion afin :

- De limiter les risques liés à la sécurité des vols commerciaux
- De réduire les immobilisations coûteuses des avions et pénalisantes commercialement.

D'un point de vue méthodologique, ces problématiques nous ont conduit à implémenter ou adapter des méthodes de détection de rupture, d'anomalie (isolation forest), clustering, analyse des messages.

Pour la suite de nos travaux, nous souhaitons aller plus loin et proposer des algorithmes d'IA, des chatbots en gardant de façons sous-jacent l'innovations mathématiques à apporter dans le domaine de l'IA et bien entendu faire des publications.

## **II. Les verrous à lever**

La situation de la recherche actuelle autour des problématiques abordées ne nous permet pas de répondre à certaines questions. Ainsi, notre recherche devrait permettre de lever quelques verrous aux questions suivantes :

### **1. Quelles sont les données pertinentes qui permettent d'anticiper une panne sur un système avion ?**

En fait, comme signifier précédemment, actuellement, la maintenance est curative puisque concentrée sur les réparations post-panne, mais également préventive puisque certains systèmes sont remis à neuf à intervalles constants prédéfinis. Dans cette configuration, les pannes surviennent sans avoir été anticipées, ce qui engendre des conséquences sur l'exploitation mais également un risque pour la sécurité des vols.

De ce fait, et comme évoqué précédemment, les données issues de l'avion ne nous ont encore jamais permis de prédire une panne (quelques soit le système avion), nous ne disposons pas des listes de données pertinentes à étudier pour détecter une panne proche sur chaque système avion.

C'est pourquoi un aspect de nos travaux de recherche va être de constituer un socle de données qui seront nos variables significatives sur chaque système étudié (incertitude 1). Ces données seront de divers types (chiffrées, courbes, discrètes, etc.) nécessitant éventuellement des outils d'analyse, machine learning à partir de données textuelles s'ils s'avèrent pertinent d'intégrer les rapports d'incidents pilotes.

### **2. Comment exploiter les données pour proposer des alertes cohérentes sur l'état d'un système ?**

Une fois les données pertinentes identifiées dans la partie précédente, nous devons étudier la structure des variables aléatoires d'intérêt pour estimer leurs lois de probabilités (incertitude 2).

Puis, nous devons déterminer les tests à mettre en place à partir de ces modèles (incertitude 3) ainsi que les niveaux d'alerte (incertitude 4). Outre les tests sur l'historique, notre plus grande incertitude sera levée lors de la première dépose de pièce prédictive.

Nous devons prendre nos précautions quant à l'exploitation des données à plusieurs niveaux :

- Certains types de vols peuvent modifier nos paramètres (ex : les vols salins).
- Les conditions météorologiques peuvent également impacter nos analyses et plus largement les saisons.
- Les phases de vols (dans l'ordre : taxi, décollage, montée, croisière, descente, atterrissage, taxi) peuvent également impacter.
- D'autres aléas pourront également être découverts à l'exploitation des données et l'utilisation des modèles.

### **3. Comment intégrer nos modèles à la maintenance ?**

Nous sommes conscients qu'une vulgarisation des principes mathématiques sera nécessaire pour que nos modèles acquièrent une légitimité suffisante pour être utilisés au quotidien, d'autant plus que nos alertes engendreront des actions et des coûts non négligeables (ou des économies dans le meilleur des cas).

Ainsi des aléas humains peuvent être rencontrés. Néanmoins, nous devons faire de l'humain une force en lui offrant la possibilité d'optimiser le modèle par son expérience.

Outre la descente d'information nous utiliserons également une logique ascendante et devons mettre en place :

- Une interface de restitution des résultats des modèles algorithmiques permettant de consulter les alertes sur les systèmes ;
- Une interface de remontée de l'information lorsque le modèle génère des doutes ou que les équipes techniques souhaitent perfectionner l'outil / soumettre des idées.

### **III. Objectif spécifique de la thèse.**

La thèse vise à fournir des éléments de réponse aux questions précédentes au moins pour certains systèmes avions qui seront identifiés comme prioritaires. Notre démarche nous conduira à :

1) la modélisation des lois de probabilités des variables aléatoires, des processus temporelles, dont les observations sont fournies par les QAR. Ces modèles nous permettront de construire des tests statistiques pour la détection d'alerte et donc de prévision des pannes et par conséquent d'aide à la décision à la maintenance prédictive.

2) Une fois ces modèles éprouvés, une interface de maintenance prédictive sera intégrée aux outils de suivi standard qui permettra de compléter l'analyse préventive par l'analyse prédictive.

Ce travail sera réalisé pour chaque système avion, par ordre de criticité au regard de nos objectifs primordiaux, à savoir la prédiction des défaillances de système avion afin :

- De limiter les risques liés à la sécurité des vols commerciaux
- De réduire les immobilisations coûteuses des avions et pénalisantes commercialement.

#### **Algorithme à mettre en œuvre**

Ainsi, cette thèse devrait permettre la mise en place d'un algorithme (en que nous concertation avec les experts techniques) résumons par les étapes suivantes :

1. Sélection d'un système par l'équipe technique (criticité, sécurité, couts, fréquence des pannes, etc.)
2. Transmission du socle de connaissances suffisantes des experts techniques aux experts en traitement de données.
3. Sélection des données significatives et tests.
4. Recherche / construction de modèles statistiques fiables et des systèmes d'alertes cohérents (voir ci-dessous).
5. Adaptation des modèles au système (voir ci-dessous)
6. Mettre en place des seuils d'alertes.
7. Tests de nos modèles sur l'historique : validation du modèle ou le retour à l'étape 3.
8. Transmission des modèles sur une interface de test pour les experts techniques : validation du modèle ou le retour à l'étape 3.
9. Intégration des modèles à l'interface finale.
10. Dépot de l'équipement en prédictif : validation du modèle ou le retour à l'étape 3.
11. Validation définitive du modèle
12. Mettre en place un chatbot.

#### **Pour réaliser des étapes 4-5. Construction de modèles de détections de pannes, d'anomalies.**

D'un point de vue technique, ces étapes se traduisent par :

1. Analyser et structurer les données collectées par les méthodes de machine learning comme par text mining ou « *Natural Language Processing* » (*NLP*), clustering, classification, régression. Cette étape permettra d'identifier les variables significatives caractérisant le système étudié.
2. Construire des approches permettant de modéliser les lois de probabilités de ces variables qu'elles soient de type réel, vectoriel, textuel, fonctionnel ou processus temporel, ou autre ou une combinaison de chacun de ces types.
3. Proposer des tests de comparaisons de distributions pour ces variables en fonctionnement normal versus anormal. Intégrer ces tests dans un algorithme en fonction de la situation (type détection de rupture (Azzaoui et al (2013), Li & Ghosal (2018), Chiou et al. (2019)) « online » via des approches de CUSUM non-paramétrique (Sahki et al. (2019))).
4. Mettre en compétition les approches développées avec l'existant pour une détection optimale des pannes sur le système avion d'intérêt.
5. Intégrer cet algorithme à l'algorithme général décrit ci-dessus (étape 5).

Nous envisageons de déployer ces algorithmes au moins pour 2 systèmes avions durant la thèse.

**Mots clés :** Machine Learning, Intelligence Artificielle, Modèles probabilistes, Modélisation de processus temporels, détection de rupture, Text Mining, Natural Language Processing, détection d'anomalie, fiabilité.

#### **Profil cherché**

Le candidat devra avoir un M2 en mathématiques appliquées ou un diplôme équivalent (école d'ingénieurs), et une formation en statistique.

**Début souhaité : Automne 2024.**

**Contacts :** Envoyer CV et lettre de motivation à [anne.yao@uca.fr](mailto:anne.yao@uca.fr) et [jutrespeuch@hop.fr](mailto:jutrespeuch@hop.fr)