

Intitulé français du sujet de thèse (limité à 250 caractères espaces compris) 114

Comptage des caractères contenus dans la cellule ci-dessous : 114

Détection d'anomalies et Indicateur de Dégradation par Apprentissage automatique: approches en Matrices Aléatoires

Nom du Laboratoire LIST3N
Equipe de recherche Modélisation Stochastique, Apprentissage & Décision (MSAD)

NOM et Prénom du directeur de thèse KHAROUF Malika
Courriel du directeur de thèse malika.kharouf@utt.fr

NOM et Prénom du co-directeur de thèse HUYNH Khac Tuan
Courriel du co-directeur de thèse tuan.huynh@utt.fr

Mots clés
Maintenance prévisionnelle, détection d'anomalies, intelligence artificielle, deep learning, matrices aléatoires, système intel

Champs scientifique
(à choisir OBLIGATOIREMENT parmi la liste "Champs scientifiques" en feuille 3 de ce fichier)
Science des données, Statistique et Probabilité, Sciences de l'ingénieur

Description du sujet de thèse (détaillant le poste à pourvoir)
(maximum 4000 caractères espaces compris)

Comptage des caractères contenus dans la cellule ci-dessous : 3203

L'optimisation de la maintenance des systèmes de production industrielle est une préoccupation majeure des responsables de la maintenance qui souhaitent implanter les politiques les plus pertinentes aux plans techniques et économiques. Des politiques de maintenance préventive basée sur l'âge existant d'ores et déjà sur le marché commencent à ne plus répondre au besoin personnalisé et éco-responsable en exploitant au plus juste les ressources. Dans ce contexte, l'évolution de la maintenance classique à une maintenance « intelligente » est devenue un sujet important. La maintenance prévisionnelle s'impose comme une solution efficace car elle permet non seulement d'anticiper des pannes à l'avance grâce à la surveillance du fonctionnement du système et à la prédiction de son état de santé, mais aussi d'optimiser des ressources humaines et matérielles pour diminuer le coût de la maintenance et la gravité des conséquences d'une panne.

Avec la diffusion d'aujourd'hui des nouvelles technologies numériques utilisant des objets connectés, l'internet des objets, le big data, l'intelligence artificielle et plus généralement la science des données, la maintenance prévisionnelle est de plus en plus adoptée et adaptée et fait émerger dans des actions stratégiques majeures au sein des entreprises. À titre d'exemple, Air France - KLM a exploité les historiques de vol des A-380 et des atterrissages à Paris pour établir un programme Big Data permettant de détecter une panne possible en moins d'une heure et d'établir son diagnostic en cinq minutes. Un projet Big Data de SNCF Transiliens vise à automatiser le diagnostic des pannes à partir de données issues des rames connectées en temps réel, offrant ainsi une vision plus précise et complète de l'état du matériel.

Malgré ses attentes prometteuses, la mise en œuvre de la maintenance prévisionnelle reste encore un grand défi par manque de connaissance en temps réel sur l'état de santé du système. La construction en ligne des indicateurs de dégradation d'un système à partir des données de surveillance est donc une question centrale de la maintenance prévisionnelle. Il s'agit de prendre en compte des données de plus en plus abondantes et de types différents (Big Data) issues de réseaux de capteurs dans la construction de l'indicateur. Cet indicateur doit être simple à dimension réduite et ayant pour but d'estimer le temps de vie résiduel du système. Cette approche nécessite souvent de se munir d'un grand nombre de capteurs, ce qui peut être coûteux. Ainsi, l'élaboration d'un indicateur de dégradation

SUJETS DE THESE 2015
- promotion en Français sur le site ABG -

équilibrant qualité et économie est un enjeu majeur.

D'un point de vue scientifique, cet enjeu implique trois problématiques principales.

1. Développer une méthode de réduction de la dimension des données adaptée à la fois aux contextes supervisé et non supervisé.
2. Spécifier, parmi les capteurs prévus, l'emplacement des capteurs permettant des données significatives pour la construction des indicateurs de dégradation.
3. Élaborer des indicateurs de dégradation à partir des données retenues et les valider.

L'objectif de cette thèse est de développer un nouvel outil pour résoudre les trois problématiques ci-dessus.

[Faire ALT+Entrée pour aller à la ligne](#)

Profil du candidat)

(maximum 4000 caractères espaces compris)

Comptage des caractères contenus dans la cellule ci-dessous :

3906

La réalisation de ce projet de thèse nécessite d'avoir des connaissances à l'interface des domaines des probabilités/statistiques, de l'intelligence artificielle, de la maintenance prédictive. Plus précisément, le projet peut se déployer selon les deux axes suivants.

1. Détection d'anomalies et sélection de variables : La détection d'anomalies est un enjeu majeur dans l'industrie actuelle, et l'afflux de données issues d'une multitude d'objets connectés impose la recherche de méthodes multidimensionnelles, intelligentes et adaptées à la complexité des données récoltées. Nous proposons d'explorer et compléter la théorie des matrices aléatoires combinée aux méthodes du deep learning afin de proposer de nouvelles méthodes de détection d'anomalies bien adaptées au contexte industriel complexe. Après récolte et nettoyage des données, nous les stockerons dans une matrice, qui sera de grande dimension, et nous traduisons la complexité de la donnée par un facteur aléatoire. Nous nous proposons donc de traiter cette grande matrice aléatoire de données à l'aide de réseaux de neurones de types auto-encodeurs convolutifs (AEC) afin de détecter d'éventuelles anomalies.

Dans un premier temps, nous mobilisons les connaissances en matrices aléatoires afin de définir une stratégie AEC adaptée de détection. Nous proposons de définir un modèle à variances isolées adapté, qui modélise bien les données contenant des outliers ou des informations rares ou soudaines. Par le biais d'une étude spectrale, les événements rares peuvent être mis en évidence en étudiant le comportement global des valeurs et vecteurs propres dans un premier temps et ensuite l'étude des valeurs propres extrêmes via un AEC ainsi que les vecteurs propres associés. Cette étude se fera dans un contexte asymptotique, cas où les dimensions de la matrice des données sont très grandes, mais aussi dans un cadre fini, dans le cas où les données significatives récoltées sont de petites tailles. Dans un second temps, nous nous intéresserons à la sélection des capteurs les plus pertinents. Nous pensons traduire la matrice de données en terme de matrice d'adjacence d'un certain graphe aléatoire et développer des méthodes exploitant les connaissances en graphes aléatoires afin de trancher quant au choix des features.

2. Définition d'un indicateur de dégradation : Une maintenance prévisionnelle nécessite aussi bien l'évaluation de l'état de dégradation du système que le pronostic sur l'apparition de futures panne. Un indicateur de dégradation permet de se rendre compte de l'état de dégradation du système. Une définition efficace de cet indicateur est donc indispensable. Nous partons d'une définition d'un indicateur de dégradation en terme de métrique entre l'instant de défaillance et tous les autres instants. Nous nous intéresserons aux méthodes à noyaux appliquées dans un cadre aléatoire afin de définir/tester plusieurs types de métriques (distance euclidienne, distance exponentielle, etc.) et d'en choisir les plus prometteuses.

D'un point de vue pratique, les méthodes développées peuvent être testées et validées sur plusieurs jeux de données mis à disposition par la communauté PHM. Nous citons à titre d'exemple le jeu de données construit par un centre de recherche de la NASA à partir des simulations des turboréacteurs d'avion avec le Modèle C-MAPSS.

Nous sommes à la recherche un(e) candidat(e) qui possède une ou plusieurs compétences suivantes:

1. Probabilité et statistiques
2. Intelligence artificielle, machine learning.
3. Un bon niveau en programmation (Python, Matlab, ou Julia) et une première expérience en recherche sont appréciés.

Pour candidater, merci d'envoyer votre CV, lettre de motivation, les relevés des notes de Master et d'ingénieur et d'autres documents (2 lettres de recommandation, publications, ...) à malika.kharouf@utt.fr et tuan.huynh@utt.fr avant le 20 mai 2024.

A remplir UNIQUEMENT si vous souhaitez que l'E.D. diffuse cette offre en Anglais

Intitulé anglais du sujet de thèse (limité à 250 caractères espaces compris)

Comptage des caractères contenus dans la cellule ci-dessous : 91

Anomaly Detection and Degradation Indicator through Machine Learning:Random Matrix approach

Keywords

Predictive Maintenance, Anomaly Detection, Artificial Intelligence, Deep Learning, Random Matrix theory

Scientific fields

(à choisir OBLIGATOIREMENT parmi la liste "Champs scientifiques" en feuille 3 de ce fichier)

Data science, Statistics, Engineering sciences

Description du sujet de thèse en Anglais ((détaillant le poste à pourvoir et le profil du candidat))
(maximum 4000 caractères espaces compris)

Comptage des caractères contenus dans la cellule ci-dessous : 2666

Optimizing maintenance for industrial production systems is a major concern for maintenance managers who aim to implement the most relevant policies from both technical and economic perspectives. Existing age-based preventive maintenance policies on the market are starting to fall short of meeting the personalized and eco-responsible needs by efficiently utilizing resources. In this context, the transition from conventional maintenance to 'smart' maintenance has become an important topic. Predictive maintenance emerges as an effective solution because it not only anticipates failures in advance through system monitoring and health state prediction but also optimizes human and material resources to reduce maintenance costs and the severity of consequences resulting from a breakdown.

With the widespread adoption of digital technologies such as connected objects, the Internet of Things (IoT), big data, artificial intelligence, and data science, predictive maintenance is increasingly being embraced and adapted, becoming a major strategic focus within companies. For example, Air France - KLM utilized flight histories of A-380s and landings in Paris to establish a Big Data program capable of detecting a potential failure in less than an hour and diagnosing it within five minutes. A Big Data project by SNCF Transiliens aims to automate fault diagnosis using

real-time data from connected trains, providing a more precise and comprehensive view of equipment status.

Despite its promising expectations, implementing predictive maintenance remains a significant challenge due to the lack of real-time knowledge

[Faire ALT+Entrée pour aller à la ligne](#)

Profil du candidat)

(maximum 4000 caractères espaces compris)

Comptage des caractères contenus dans la cellule ci-dessous :

220

The completion of this thesis project requires knowledge at the interface of

Anomaly Detection and Variable Selection: Anomaly detection is a r

Initially, we will utilize knowledge in random matrices to define a suitable C

Definition of a Degradation Indicator: Predictive maintenance require

From a practical standpoint, the developed methods can be tested and va

We are looking for a candidate who possesses one or more of the followin

Probability and statistics

Artificial intelligence, machine learning

Proficiency in programming (Python, Matlab, or JuliaLang) and some

To apply, please send your CV, cover letter, Master's and Bachelor's trans

Faire ALT+Entrée pour aller à la ligne

of probability/statistics, artificial intelligence, and predictive maintenance are a major challenge in current industry, and the influx of data from numerous sensors has led to a data-driven, data-driven detection strategy. We propose to define an adapted isolation method for both evaluating the system's degradation state and forecasting future failures, validated on various datasets provided by the PHM community. As an engineering skills:

Research experience are appreciated.

Scripts, and other documents (2 recommendation letters, publications,

re. Specifically, the project can unfold along the following two axes.
The use of connected devices necessitates multidimensional, intelligent methods such as the robust variance model, which effectively models data containing outliers and sensor failures. A degradation indicator allows insight into the system's degradation. For example, we mention the dataset constructed by a NASA research center.

etc.) to malika.kharouf@utt.fr and tuan.huynh@utt.fr before April 30,

ods tailored to the complexity of collected data. We propose to explore
or rare/sudden information. Through spectral analysis, rare events ca
gradation state. An effective definition of this indicator is therefore ess
iter from simulations of aircraft turbojets using the C-MAPSS model.

2024.

and enhance the theory of random matrices combined with deep learning can be highlighted by studying the overall behavior of eigenvalues and eigenvectors. We start with a definition of a degradation indicator in terms of

arning methods to offer new anomaly detection methods well-suited to
eigenvectors initially, and then extreme eigenvalues via a CAE and as
the metric between the failure time and all other times. We will explor

the complex industrial context. After data collection and cleaning, we associated eigenvectors. This study will be conducted in both asymptotic and finite kernel methods applied in a random framework to define/test various

will store them in a high-dimensional matrix, translating the data com
c and finite contexts, where the dimensions of the data matrix are very
is types of metrics (Euclidean distance, exponential distance, etc.) and

plexity into a random factor. Thus, we aim to treat this large random data set as a large and where the collected significant data are of small sizes, researchers should choose the most promising ones.

lata matrix using convolutional autoencoder neural networks (CAEs) t
pectively. In a second stage, we will focus on selecting the most relev

o detect potential anomalies.
ant sensors. We plan to translate the data matrix into the adjacency n

matrix of a certain random graph and develop methods leveraging kno

wledge in random graphs to decide on feature selection.

Agronomie, agroalimentaire
Architecture, arts (appliqués, plastiques, du spectacle), musique
Biochimie
Biologie
Biotechnologie
Chimie
Communication, information, media, culture
Droit, science politique, géopolitique
Ecologie, environnement
Economie et gestion
Electronique
Energie
Génie civil, BTP
Génie des procédés
Géographie
Histoire, histoire de l'art et civilisations
Indifférent
Informatique
Littérature, langues, philosophie
Matériaux
Mathématiques
Numérique
Physique
Physiologie, neurosciences
Robotique
Santé, médecine humaine, vétérinaire
Science de la donnée (stockage, sécurité, mesure, analyse)
Sciences de l'ingénieur
Sociologies, anthropologie, sciences de l'éducation
Télécommunications
Terres, univers, espace