

Renforcement de la robustesse d'un modèle hybride de pronostic de défaillances d'un moteur électrique par apprentissage progressif

M. Suhas¹, E. Abisset-Chavanne², P.-A. Rey²

¹ AMValor, Laboratoire I2M Bordeaux, morgane.suhas@ensam.eu

² Arts et Métiers, Laboratoire I2M Bordeaux, emmanuelle.abisset-chavanne@ensam.eu, pierre-andre.rey@ensam.eu

Résumé — Cet article présente une méthodologie innovante exploitant le contenu informationnel essentiel de vastes ensembles de données d'entrée pour améliorer la robustesse des modèles de régression. Cette approche repose sur l'apprentissage progressif, intégrant graduellement des données d'apprentissage de complexité croissante pour renforcer la robustesse du modèle. Cette méthode améliore la capacité prédictive par rapport à un apprentissage complet initial, sans écraser les poids du réseau préalablement entraîné. L'étude se concentre sur les moteurs électriques avec une généralisation visée vers d'autres systèmes industriels.

Mots clés — Curriculum Learning, Traitement données d'entrée, Modèle hybride, Réseau de neurones, Pronostic de défaillances.

1 Introduction

De nombreux processus industriels, en raison de leur complexité intrinsèque due à des relations non linéaires et aux nombreuses variables impliquées, nécessitent une optimisation constante. Les relations entre les entrées et les sorties de ces systèmes complexes peuvent être modélisées à l'aide de données industrielles, de principes fondamentaux ou d'une combinaison des deux, connue sous le nom de modélisation hybride. La modélisation hybride, qui s'appuie sur les avantages des modèles basés sur les principes fondamentaux et des modèles basés sur les données, offre un équilibre entre l'exactitude du processus et la qualité des données. Les principes fondamentaux et, plus généralement, les modèles basés sur la physique apportent un suivi analytique précis des phénomènes se produisant sur le système tandis que les données réelles, relevées aisément sur le système, retranscrivent qualitativement et quantitativement le comportement du système en temps réel. Plusieurs études approfondies [2, 6, 11] nous ont conforté dans le choix de procéder à ce type de modélisation pour résoudre notre problème.

Dans notre cas, il s'agit d'un problème de pronostic des défaillances tout en gérant la robustesse du modèle en présence de conditions d'usage extrapolées. Effectivement, le système peut subir de nouvelles conditions d'usage après avoir été modélisé dans son intégralité. En réutilisant l'implémentation réalisée initialement, on se propose de mettre en place une méthode ajustée aux données d'entrée du modèle physique en vue d'améliorer la robustesse du modèle complet. Après avoir validé numériquement le modèle sur un problème de classification [9], le principe est d'approfondir l'analyse sur un problème de régression en exploitant les paramètres présents dans les modèles de dégradation. L'augmentation de la robustesse d'un modèle de régression passe en premier lieu par l'étendue des bases de données d'entrée. Toutefois, lorsque la taille des données d'entrée est trop conséquente, il se peut que le modèle présente plus de difficultés pour détecter et généraliser les corrélations entre les variables. De plus, cela entrave la distinction des profils ayant pu conduire aux différentes valeurs de la variable de réponse. L'approche proposée repose principalement sur l'intégration progressive de lots de données d'apprentissage avec une complexité croissante durant l'entraînement.

L'objectif est de montrer que l'ajout gradué des connaissances ou des données améliore les capacités du modèle prédictif comparé à un apprentissage complet initialement. En pratique, cela permet d'enrichir les bases des profils d'usage ultérieurement, lorsque le système est soumis à de nouvelles utilisations, sans écraser les poids du réseau entraîné précédemment. Une stabilité du modèle est ainsi garantie. La méthodologie est développée autour des données d'entrée d'un cas d'usage bien qu'elle soit construite dans le but d'être généralisable à d'autres systèmes industriels.

Dans un premier temps, l'architecture du modèle de régression sera développée, de même pour le format des bases de données d'entrée. Le concept de l'apprentissage progressif, également dénommé le curriculum learning, est par la suite introduit. Son intérêt est démontré au travers de possibles applications dans le domaine de l'ingénierie [5, 7, 8], et plus précisément étudié sur le moteur à courant continu. Contrairement aux problèmes de reconnaissance de texte, ce concept est très peu répandu, voire inexistant, pour les régressions sur des séquences temporelles. Le challenge réside alors dans la mise en place d'une fonction d'évaluation des bases de données afin de déterminer la complexité de chaque base dans le cas des séries temporelles. Cette fonction de notation permet de trier les exemples d'apprentissage par difficulté et de déterminer l'ordre de présentation des exemples de données au réseau.

2 Modèle hybride global

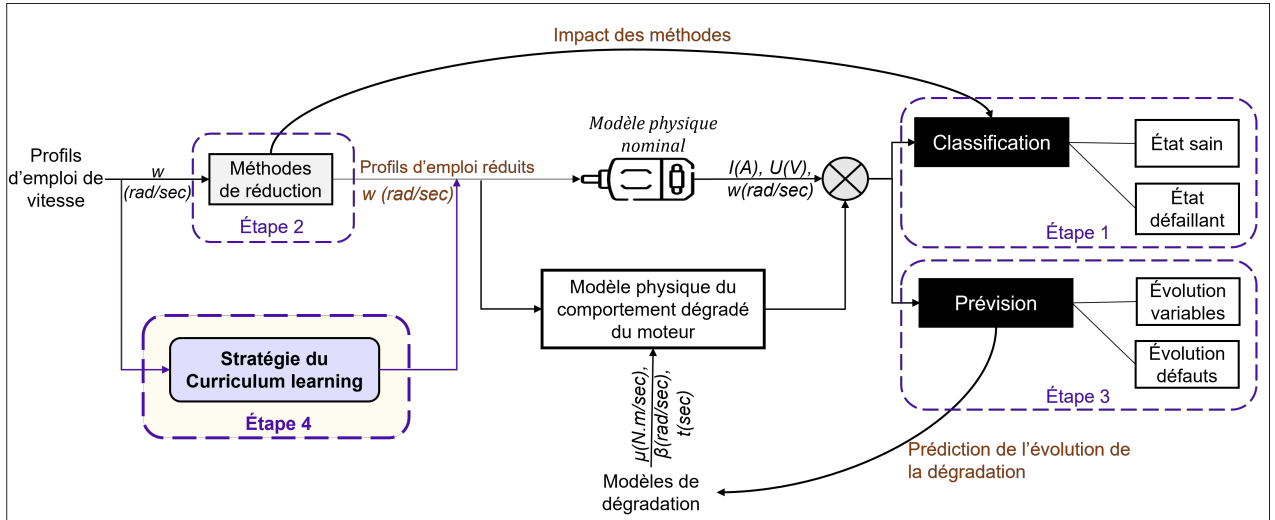


FIGURE 1 – Cadre général pour contextualisation de la stratégie d'apprentissage implémentée

Pour évaluer la durée de dégradation du moteur électrique, cet article propose une méthode de pronostic en se basant sur les modèles guidés par les données et sur les connaissances physiques du moteur [11]. Le modèle hybride développé dans un travail antérieur est représenté sur la figure 1. La contribution énoncée en détail dans ce travail est illustrée par l'étape 4 de la figure 1. Les équations caractérisant le modèle hybride sont telles que :

$$\Leftrightarrow \Theta_{sys}(\lambda) = \Theta_{\varphi}(\psi) + \varepsilon_{res}(\lambda) \quad (1)$$

$$\Theta_{sys}(\lambda) = \Theta_{\varphi}(\lambda) + \Theta_{RNN}(\lambda) \quad (2)$$

Θ_{sys} représente la modélisation du système en fonction du profil d'utilisation λ . Θ_{φ} est la modélisation physique calculée à partir des équations du moteur en comportant sain ψ . Le comportement ψ peut par exemple être déterminé grâce à plusieurs points de fonctionnement ou grâce à des équations de fonctionnement. ε_{res} (éq. 1), constituant l'ensemble des résidus, est la déviation entre le modèle nominal et le système réel. La modélisation finale du système est obtenue en exploitant les résidus en tant que données d'entrée du réseau de neurones Θ_{RNN} (éq. 2). Ce dernier est un modèle de pronostic de défaillances du moteur électrique estimant la durée de vie utile.

Un modèle de dégradation exponentielle est appliqué pour produire des données synthétiques dégradées du système et ensuite pour prévoir le temps de défaillance [3]. Ses paramètres caractéristiques sont considérés en première partie comme des variables explicatives. Comme on peut le voir sur l'étape 3 de la figure 1, les paramètres du modèle de dégradation sont calculés dans une deuxième partie dans le but de prévoir l'évolution progressive de la dégradation au sein du moteur. Dans ces conditions, les paramètres deviennent la variable de réponse et les résidus appartiennent aux variables explicatives. Le modèle global a pour but de fournir une réponse de plus haute fidélité tout en réduisant la dimension de l'espace des données d'entrée. Il s'agit pour cela d'extraire l'information essentielle contenue dans les données d'entrée et de supprimer celles qui sont redondantes. Cette section expose tout d'abord l'implémentation du

modèle prédictif ainsi que ses particularités. Elle s'attarde ensuite sur la stratégie d'injection des données d'entrée pour insérer l'information nécessaire au moment opportun afin d'optimiser l'apprentissage.

2.1 Pré-traitement des données d'entrée

Les données d'entrée sur lesquelles nous allons porter notre attention sont celles du modèle physique. Elles représentent des séquences temporelles et forment une base de profils de vitesse. Les données de sortie sont les données physiques de notre cas d'usage comportant des défaillances mécaniques. Les profils de vitesse d'entrée sont répartis en 6 bases initiales dépendant de l'allure du signal et de la plage de fréquences à laquelle il appartient. Les signaux, possédant différentes fréquences et amplitudes, s'apparentent uniquement à des ondes sinusoïdales ou des ondes à impulsions. Les 6 bases comprennent chacune 60 signaux avec des fréquences allant de 1 Hertz à 100 Hertz. Les bases se différencient par la fréquence qui augmente progressivement au sein des signaux. Elles sont ordonnées en fonction de leurs spécificités physiques, à savoir leur fréquence et leur amplitude qui augmentent progressivement. Cet ordre provient de la manière dont elles ont été créées : le modèle a préalablement été testé sur des profils d'emploi classiques. A posteriori, sa capacité à s'adapter sur des profils d'emploi plus larges a été vérifiée. Une mesure de la complexité des bases, du point de vue du modèle, est par la suite décrite dans la section 3.2.

2.2 Architecture du réseau de neurones profond

Cette sous-section concerne l'étape 3 de la figure 1. Le réseau CNN (Convolutional Neural Network) et le réseau LSTM (Long-Short-Term-Memory) sont des modèles d'apprentissage automatique possédant chacun leur distinction et qui ont prouvé leurs performances et d'autant plus dans le domaine de l'analyse des séries temporelles. Pour gagner en efficacité, leurs caractéristiques respectives sont exploitées en élaborant un modèle de régression composé simultanément de couches CNN et LSTM. Effectivement, la couche LSTM décèle les dépendances à court et long terme contenues dans les séquences temporelles. Elle sera ainsi utilisée en tant que première couche du réseau. D'autre part, une couche CNN est greffée à la suite de la couche LSTM. Les couches CNN sont employées en raison de leur capacité à extraire les caractéristiques précieuses intégrées dans les données brutes et sont utiles pour filtrer les données de la série temporelle. Ce type de réseau appartient à la catégorie des modèles d'apprentissage profond hybrides (HDL) [10]. Pour l'instant, l'imbrication de cette manière a uniquement montré une précision concluante pour les tâches de classification de texte, comme dans les travaux de Zhang [13]. Par conséquent, un modèle hybride qui exploite les avantages du CNN et du LSTM devrait permettre d'améliorer la prévision de la durée de vie du système.

3 Combinaison du modèle hybride avec la stratégie du Curriculum Learning

3.1 Concept

D'après la définition originelle de Bengio [1], le curriculum est considéré comme une séquence C de critères $C_{i,i \in [1;T]}$ de T étapes de formation. Chaque critère de la séquence est associé à un ensemble différent de poids pour les données cibles d'apprentissage \mathcal{D} . Prenons $P(z)$ la distribution de l'exemple z de l'ensemble des données cibles et $0 \leq W_i(z) \leq 1$ le poids appliqué à l'exemple z à l'époque i . A l'époque i , la formalisation du curriculum est ainsi : $C_i(z) \propto W_i(z)P(z) \quad \forall z \in \mathcal{D}$.

Comme expliqué précisément dans la revue de littérature de Wang [12], une appropriation de la stratégie du curriculum learning doit être réalisée en fonction du problème à résoudre. Selon la nature des données d'entrée, un critère de complexité est à fixer pour évaluer la facilité relative de chaque exemple dans le jeu de données. Les observations sont ensuite triées de façon croissante de la moins complexe à la plus complexe. Elles sont injectées dans un deuxième temps dans la fonction rythme. La fonction de rythme définit l'échantillonnage à appliquer à chaque époque pour découper l'ensemble de données d'apprentissage en sous-ensembles. Des fonctions de rythme empiriques, telle que celle nommée "One-Pass curriculum" [4], ont déjà été testées sur plusieurs types de problèmes. De ce fait, le curriculum

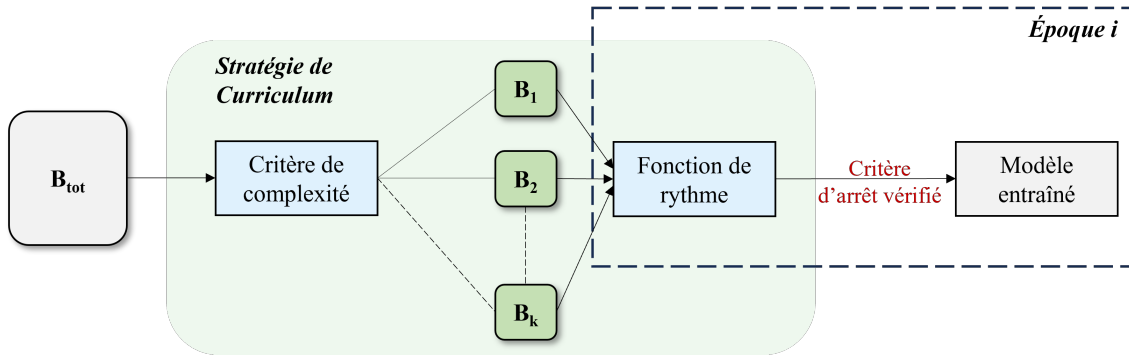


FIGURE 2 – Concept originel du curriculum learning illustré pour une époque. B_{tot} représente la base initiale des données d'apprentissage et $\{B_1, B_2, \dots, B_k\}_{k \in \mathbb{N}}$ est l'ensemble des sous-bases déterminées par le critère de complexité.

learning est caractérisé comme une technique d'apprentissage qui entraîne le modèle prédictif avec le principe du curriculum (voir figure 2) introduit ci-dessus.

3.2 Méthodologie proposée

Après avoir synthétisé les profils d'usage représentatifs du comportement du système dans des bases réduites, elles sont triées selon leur complexité. Les bases réduites sont issues des méthodes de réduction et constituent les modes de ces méthodes. Ce sont les données d'apprentissage initiales B_{tot} avant le passage par la mesure de complexité. Par souci de réduction des données d'entrée, le critère déterminant la complexité de la base pour notre étude repose sur l'importance de chaque mode. L'importance d'un mode peut être quantifié par différents termes selon la méthode de réduction sans changer la signification. Par exemple, pour la méthode de l'Analyse en Composantes Principales (ACP), on évoquera plutôt le "score" tandis que pour la méthode de décomposition en valeurs singulières (SVD) le terme "énergie" sera employé. Par conséquent, une base B_1 sera considérée plus complexe qu'une base B_2 si sa contribution pour la reconstruction des données d'entrée initiales est moindre.

La première étape consiste à créer et graduer les sous-bases selon leur difficulté. Les bases initiales, avant leur découpage conformément au critère de complexité, sont constituées des profils de vitesse décrits à la section 2.1. L'ensemble des bases initiales représentent la base B_{tot} . B_{tot} subit donc un premier processus d'analyse de contribution grâce à la méthode SVD. Une étude de l'énergie des modes dans le domaine fréquentiel a permis d'identifier que les 20 premiers modes suffisent à expliquer la variabilité de chacune des bases élaborées, et cela sans perdre d'informations dans le contenu fréquentiel. Les profils de vitesse d'entrée se retrouvent alors sous la forme de modes classés selon l'importance de l'énergie résultante de chacun des modes. Ce processus est retranscrit au travers de la fonction de notation, dépendant elle-même du critère de complexité. La séquence de critères $C_{i,i \in [1;T]}$, telle que introduite dans la sous-section 3.1, délimite le nombre de modes primaires à extraire de chaque base initiale de B_{tot} sans duplication. Lors de l'apparition d'une nouvelle base, les nouveaux modes primaires sont alors soustraits aux modes primaires trouvés dans les bases précédentes via la fonction de notation et la séquence de critères C . De cette façon, l'apprentissage complet n'est pas effectué dès qu'une nouvelle condition d'usage apparaît. La fonction de notation prédéfinit également la distance à respecter pour juger si le nouveau mode est différent de ceux des bases isolés dans les itérations précédentes.

Les bases de modes primaires sans duplication résultantes sont appelées les sous-bases $\{B_1, B_2, \dots, B_k\}$ (voir figure 2). Les bases initiales ont été créées pour contenir premièrement les profils d'utilisation les plus conventionnels. En revanche, les dernières bases construites balayent des profils d'utilisation plus larges. Ainsi, plus B_{tot} s'enrichira de nouvelles bases initiales, moins les critères de complexité récupéreront de modes primaires finaux.

Pour connaître à présent le planificateur d'injection des bases au cours de l'apprentissage et selon la définition du curriculum learning, une fonction de rythme a été mise en place. Son rôle permet la gestion du critère d'arrêt d'apprentissage entre chaque base et dès qu'il est atteint, la sous-base suivante

est fournie en tant que données d'apprentissage supplémentaires. Le critère d'arrêt choisi pour notre étude est la convergence du modèle de deep learning présenté dans la section ci-dessus. La convergence du modèle est testée pendant la phase d'apprentissage pour chaque sous-base injectée par la fonction de rythme. Pour illustrer l'algorithme avec la première itération, ce critère d'arrêt est vérifié pour la sous-base B_1 , la séquence de poids \mathcal{W}_1 , obtenue en fin d'apprentissage de l'époque 1, est transférée en tant que poids initiaux pour l'époque suivante avec la sous-base B_2 . La nouvelle séquence de poids \mathcal{W}_2 est à son tour transmise pour l'époque suivante. L'apprentissage se termine une fois que toutes les sous-bases ont été traitées et que la perte finale du modèle a convergé.

Les performances du modèle sont comparées avant et après implémentation de la stratégie de cet apprentissage progressif. Nous avons déjà montré sur un problème de classification que ce type d'apprentissage dynamique réduisait considérablement le temps de convergence durant l'apprentissage et qu'il améliorait naturellement les performances de classification (voir figure 3). La fonction de notation n'intégrait pas le retrait des signaux "similaires" à chaque nouvelle base se référant aux modes. Elle réalisait une fusion de chaque base dès la convergence de la base précédente. La méthode impliquait un tri préliminaire des bases pour s'assurer d'une complexité croissante. Dès qu'un nouveau profil d'utilisation est identifié, une vérification est effectuée pour déterminer à quelle base il correspondait le mieux. En effet, pour respecter la difficulté croissante des données, il faut s'assurer que le nouveau profil soit plus complexe que ceux de la dernière base ayant servi à l'apprentissage. Les résultats concluants sur le modèle de classification ont encouragé l'affinement de la méthode en personnalisant une fonction de notation et une fonction de rythme.

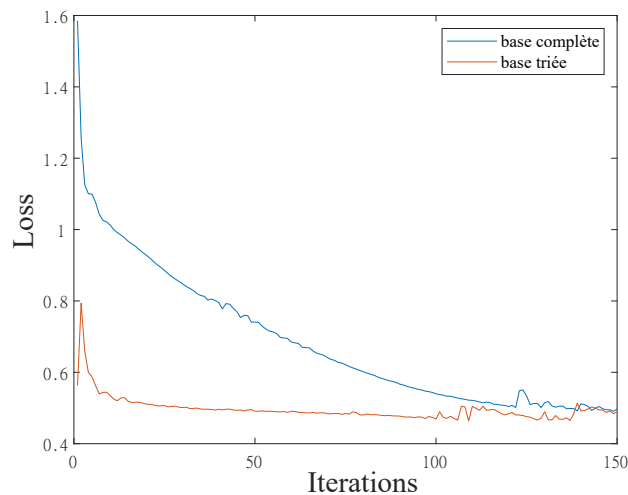


FIGURE 3 – Performances du curriculum learning combiné à un modèle de classification

Dans ce travail, une analyse plus fine est proposée puisque les nouveaux profils sont immédiatement traités grâce à leur mode. L'un des principaux atouts de cette méthode est par conséquent la gestion des nouvelles conditions d'usage tout en fixant les performances du modèle entraîné antérieurement. L'évolution de la perte finale du modèle sera affichée à chaque nouvelle injection de base et une amélioration progressive pourra être constatée.

4 Résultats

Cette approche met en lumière une amélioration des performances de régression en intégrant la technique du curriculum learning. Elle est approfondie afin d'alléger le processus d'apprentissage lorsque de nouvelles données entrent en compte. L'étude apporte, entre autres, des éléments de réponse sur l'interopérabilité des modèles prédictifs lorsqu'il s'agit de les intégrer dans un nouveau système ou à d'autres problématiques. En redéfinissant les poids du réseau de neurones de manière progressive, l'algorithme propose ainsi une amélioration continue de la robustesse du modèle hybride. Il prévoit également un enrichissement permanent des bases de données d'entrée dès lors que la stratégie de curriculum learning

devient inefficace sur de nouvelles conditions d'usage.

5 Conclusions

Les travaux contribuent à l'élaboration d'une méthodologie générique en utilisant des fonctions spécifiques pour répondre tout de même au problème posé. Une analyse plus poussée sera alors nécessaire pour déterminer ces nouvelles fonctions dans le cas de calculs de structures. Ce papier s'appuie effectivement sur la définition générale du curriculum learning pour mettre en place une méthode adaptée spécifiquement à ce problème de régression. Un nouveau critère de complexité a divisé la base complète en sous-bases triées selon ce critère. Elles sont par la suite incorporées une à une dans le processus d'apprentissage au travers de la fonction de rythme créée. La méthodologie complète mise en place vise en premier lieu à intégrer un traitement supplémentaire sur les données d'entrée pour améliorer les performances d'apprentissage. De par sa conception, le modèle hybride est capable de s'enrichir continuellement dès que de nouvelles conditions d'usage se présentent. La combinaison du curriculum learning avec les méthodes traditionnelles de réduction de dimension d'entrée offre en outre un allègement du coût de l'implémentation tout en accélérant la convergence du modèle. Bien que les performances de précision du modèle soient calculées pour une tâche de pronostic des défaillances d'un système industriel, le concept est généralisable à d'autres tâches de calcul. Notamment, cette stratégie pourrait être pensée en construisant des bases de données d'entrée au fur et à mesure que de nouvelles connaissances sur la structure du matériau soient établies.

Références

- [1] Y. BENGIO et al. "Curriculum learning". In : *Proceedings of the 26th annual international conference on machine learning*. 2009, p. 41-48.
- [2] N. BHUTANI, G. P. RANGAIAH et A. K. RAY. "First-principles, data-based, and hybrid modeling and optimization of an industrial hydrocracking unit". In : *Industrial and Engineering Chemistry Research* 45 (23 nov. 2006), p. 7807-7816. ISSN : 08885885. DOI : 10.1021/ie060247q.
- [3] S. CHAKRABORTY et al. "Residual-life estimation for components with non-symmetric priors". In : *Iie Transactions* 41.4 (2009), p. 372-387.
- [4] V. CIRIK, E. HOVY et L.-P. MORENCY. "Visualizing and Understanding Curriculum Learning for Long Short-Term Memory Networks". In : (nov. 2016). URL : <http://arxiv.org/abs/1611.06204>.
- [5] C. FLORENSA et al. "Reverse curriculum generation for reinforcement learning". In : *Conference on robot learning*. PMLR. 2017, p. 482-495.
- [6] H. HANACHI et al. "Hybrid data-driven physics-based model fusion framework for tool wear prediction". In : *International Journal of Advanced Manufacturing Technology* 101 (9-12 avr. 2019). référence et biblio, p. 2861-2872. ISSN : 14333015. DOI : 10.1007/s00170-018-3157-5.
- [7] L. JIANG et al. "Mentornet : Learning data-driven curriculum for very deep neural networks on corrupted labels". In : *International conference on machine learning*. PMLR. 2018, p. 2304-2313.
- [8] A. I. SCHEIN et L. H. UNGAR. "Active learning for logistic regression : an evaluation". In : *Machine Learning* 68 (2007), p. 235-265.
- [9] M. SUHAS et al. *Performance of hybrid modeling on a failure process in industrial systems*. 2022. URL : <https://hal.science/hal-03966562>.
- [10] C. SZEGEDY et al. *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*. 2015. URL : www.aaii.org.
- [11] J. WANG et al. "Hybrid physics-based and data-driven models for smart manufacturing : Modelling, simulation, and explainability". In : *Journal of Manufacturing Systems* 63 (avr. 2022), p. 381-391. ISSN : 02786125. DOI : 10.1016/j.jmsy.2022.04.004.

- [12] X. WANG, Y. CHEN et W. ZHU. “A Survey on Curriculum Learning”. In : (oct. 2020). URL : <http://arxiv.org/abs/2010.13166>.
- [13] J. ZHANG et al. “LSTM-CNN hybrid model for text classification”. In : *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. IEEE. 2018, p. 1675-1680.