PhD offer – INSA de Lyon

Topic: Towards certification of vibration-based condition monitoring systems using explanatory AI

Description

Objectives. Reducing our carbon footprint is one of the major challenges of the coming decades. This PhD is concerned with two approaches. Firstly, the transition to new, cleaner sources of energy, such as wind power. Secondly, increasing the lifespan and energy efficiency of existing systems, particularly in air transport. In both cases, the safety, reliability and durability of the equipment must be guaranteed if the solutions are to be feasible. Predictive maintenance is used to control these risks. It is based on data collected during operation, typically vibration data, representative of the state of health of the system being monitored, and analysed by experts in search of symptoms of failure ('Health and Usage Monitoring System' in aeronautics) [1]. Machine learning has recently demonstrated its potential to automate these tasks with considerably reduced reaction times, while guaranteeing the same or even better performance than the expert [2]. Despite numerous publications demonstrating this potential, this approach has not yet been implemented in critical industrial sectors such as aeronautics or wind power, due to the difficulty of certifying it. The aim of this thesis is precisely to make the automatic learning methods used to monitor the health of mechanical systems certifiable, so that the results they propose can be validated against a benchmark [3][4]. In this respect, it deals with "explanatory artificial intelligence" in "vibration monitoring", thus covering two areas of expertise represented respectively by the LIRIS and LVA laboratories at INSA Lyon.

Scientific challenges: The AI models used in predictive maintenance and which constitute the object of study of the thesis are deep convolutional neural networks, known as 'end-to-end', which take the vibration signals measured at the input and output a binary classification ("healthy" versus 'faulty') or multiple classification (association with different types of fault) [5]. These models have demonstrated excellent performance, but are still difficult to explain. They operate like 'black boxes', which implies a lack of transparency in their decisions, which poses a problem in the event of a diagnostic error.

The aim of the research project is to answer a number of fundamental questions, which represent the current barriers in the field:

- How can we guarantee a level of confidence in the diagnosis based on the results of an AI model?
- How can the knowledge and experience of a human expert be taken into account when learning an AI model?

These general questions can be answered in more specific ways.

- What are the patterns discovered in vibratory signals by a deep convolutional neural network? The answer will enable us to check that these patterns are relevant from the point of view of the human expert, either because they corroborate the current state of knowledge, or because they provide new knowledge [10][11].

• How can the size of the data be reduced to a small number of latent variables with a high information content and potentially interpretable on the basis of physical principles? The human expert will expect to find here an image of vibratory excitations.

• How can digital learning and symbolic learning be combined? The extraction of logical rules will facilitate the interpretation of decisions made by an AI model and their validation by a human expert [13][14].

The construction of answers to these questions is the essential object of study in explanatory AI [12].

Keywords: explanatory AI, machine learning, predictive maintenance, engineering.

Context of the project: The doctoral thesis described below is part of a series of theses designed to build a multidisciplinary scientific approach to the societal challenge of a "responsible digital society", and more specifically, the specific theme of "Data and AI in a sustainable and responsible approach", identified as a priority issue by the 4 institutions of the Lyon Saint-Etienne Engineering College (Centrale Lyon, ENTPE, INSA Lyon, Mines Saint-Étienne) and by the Université Jean Monnet Saint-Étienne, which are providing financial support for the theses making up this 2025 package. In this context, the thesis provides an opportunity to exchange ideas with other doctoral students in related fields throughout the project.

Industrial partnership: The PhD project is supported by the SAFRAN TECH (Paris-Saclay) and ENGINE Green (Lyon) groups, with which regular exchanges are planned.

Benefits for the PhD student: The project will enable the PhD student to acquire high-level skills in AI and its applications in mechanical engineering, in a multidisciplinary research environment with links to industry. Training through and in research will enable him/her to join, for example and among many other possibilities, an R&D department at international level.

Start date and duration: 3 years from the planned start date between September and November 2025.

Funding: Doctoral contract (https://www.enseignementsup-recherche.gouv.fr/fr/le-financement-doctoral-46472)

Location: Campus of INSA Lyon

- Laboratoire Vibrations Acoustique (LVA) UR677 https://lva.insa-lyon.fr/, 25 bis avenue Capelle, 69621 Villeurbanne Cedex
- Laboratoire d'Informatique en Image et Systèmes d'Information (LIRIS) UMR 5205 https://liris.cnrs.fr/, Bâtiment Blaise Pascal, Campus de la Doua, 7 avenue Jean Capelle, 69622 Villeurbanne

Hosting university: INSA Lyon -- **Doctoral school:** Ecole Doctoral MEGA (ED 162)

Supervisors & contacts: Codirection par Jérôme Antoni (LVA, jerome.antoni@insa-lyon.fr) et Céline Robardet (LIRIS, celine.robardet@insa-lyon.fr)

Deadline for application: 15 May 2025

Candidate's profile: 5 years' higher education or more and initial research experience (Master project or laboratory internship), with a solid theoretical background enabling to tackle problems in statistics, data processing and AI.

Application: Send a CV, transcripts for the last 5 years and a cover letter to the supervisors.

References

[1] R.B. Randall and J. Antoni, “Rolling Element Bearing Diagnostics – A Tutorial”, Mechanical Systems and Signal Processing, 25(2), 2011, pp.485-520, https://doi.org/10.1016/j.ymssp.2010.07.017.

[2] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, A. K. Nandi, Applications of machine learning to machine fault diagnosis: A review and roadmap, Mechanical Systems and Signal Processing, 138, 2020, 106587, https://doi.org/10.1016/j.ymssp.2019.106587.

[3] B. Fresz et al.. The Contribution of XAI for the Safe Development and Certification of AI: An Expert-Based Analysis. 2024, https://arxiv.org/abs/2408.02379.

[4] L. Cummins et al., Explainable Predictive Maintenance: A Survey of Current Methods, Challenges and Opportunities, 2024, arXiv: arXiv:2401.07871. 10.48550/arXiv.2401.07871.

[5] F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mechanical Systems and Signal Processing, 72–73, 2016, 303-315, https://doi.org/10.1016/j.ymssp.2015.10.025.

[6] P. Borghesani, N. Herwig, J. Antoni, W. Wang, A Fourier-based explanation of 1D-CNNs for machine condition monitoring applications, Mechanical Systems and Signal Processing, 205, 2023, 110865, https://doi.org/10.1016/j.ymssp.2023.110865.

[7] S. Gawde, S. Patil, S. Kumar, P. Kamat, K. Kotecha and S. Alfarhood, "Explainable Predictive Maintenance of Rotating Machines Using LIME, SHAP, PDP, ICE," in IEEE Access, 12, 29345-29361, 2024, doi: 10.1109/ACCESS.2024.3367110.

[8] T. Decker, M. Lebacher, and V. Tresp, “Does Your Model Think Like an Engineer? Explainable AI for Bearing Fault Detection with Deep Learning,” arXiv.org. Accessed: Sep. 26, 2024. https://arxiv.org/abs/2310.12967v1.

[9] A. Shrikumar, P. Greenside, and A. Kundaje, Learning Important Features Through Propagating Activation Differences, 2019, arXiv: arXiv:1704.02685. doi: 10.48550/arXiv.1704.02685.

[10] U. Schlegel and D. A. Keim, Time Series Model Attribution Visualizations as Explanations, 2021, arXiv: arXiv:2109.12935

[11] F. Mujkanovic, V. Doskoč, M. Schirneck, P. Schäfer, and T. Friedrich, “timeXplain -- A Framework for Explaining the Predictions of Time Series Classifiers,” 2023, arXiv:2007.07606.

[12] F. Bodria, F. Giannotti, R. Guidotti, F. Naretto, D. Pedreschi, S. Rinzivillo: Benchmarking and survey of explanation methods for black box models. Data Min. Knowl. Discov. 37(5): 1719-1778 (2023).

[13] L. Veyrin-Forrer, A. Kamal, S. Duffner, M. Plantevit, C. Robardet: On GNN explainability with activation rules. Data Min. Knowl. Discov. 38(5): 3227-3261 (2024).

[14] A. Ragno, M. Plantevit, C. Robardet, R. Capobianco: Transparent Explainable Logic Layers. ECAI 2024: 914-921.