

# Explainability of time series prediction models for water resources management

Regional ARD JUNON Project on Digital Twins\*

The **LIFAT UR 6300** of the University of Tours is offering an 18-month post-doctoral position in explainability as part of the JUNON project which aims at producing digital twins with **BRGM** in the context of Earth Science applications for the management of surface and sub-surface resources with a view to sustainable development (in French, “JUNON : des jumeaux numériques au service des ressources naturelles”).

## 1 Context

The JUNON project aims at building AI tools for making predictions about environmental objects. The first issue and main scientific challenge of the *Prediction* project within JUNON are to propose new predictive AI methods adapted to the specific features of environmental problems such as multi-source and heterogeneous data, non-stationarity, adaptation for digital twins transfer, and reusability and evaluation of the proposed model.

This post-doc takes place in the *Prediction* and more precisely in the evaluation process of AI models for digital twins with the development of new algorithms for the explainability of AI/ML models, also termed as eXplainable Artificial Intelligence (or XAI for short).

Current XAI approaches suffer from two main limitations [5]. The first [3] is related to the *complexity of the explanation process itself*, which involves the particular characteristics of “the training data, the precise shape of the decision surface, and the selection of one explanatory algorithm over another”. As a result, there is a risk of accepting plausible explanations that only reflect spurious correlations between internal layers of DL with input features. [3] emphasizes the need for explanations methods “that cast doubt on certain features of AI systems”. These results call for (i) a more thorough consideration of inner relationships in the data and how models use this information, (ii) methods to assist users in selecting explanation methods based on objective metrics. The second limitation is related to the lack “of XAI approaches tackling real-world machine learning issues” that would “help to clarify what is currently feasible and what is not feasible when employing XAI techniques” [5].

---

\*<https://www.brgm.fr/fr/programme/junon-jumeaux-numeriques-au-service-ressources-naturelles>

## 2 Objectives and tasks

The JUNON post-doc aims to address fundamental issues related to the **quality** and the **applicability** of explanations produced for DL models driven by the recent vision of **actionable** explainable AI (aXAI) [1]. In particular, we focus on **more expressive forms of explanations** that can answer not only *why* questions but also action-guiding explanations such as *how-to* and *what-if* as illustrated in Figure 1.

*why*: why do we obtain a specific prediction, given the features of input observations?  
*how-to*: what are the necessary actions to change the prediction of a specific input observation?  
*what-if*: what are the necessary and minimal sets of actions on input observations required to obtain an alternative prediction?

Figure 1: Example of user questions directed at an XAI system.

During the postdoc we envision several research questions attached to the aforementioned objectives of quality and applicability of XAI approaches:

1. Benchmark existing quality metrics [8] for the task of explanations exploration.  
The recruited post-doc will contribute with a preliminary literature survey on explanation quality metrics for feature influence, counterfactual, or other causal explanation methods. An emphasis will be made on coherence with the predictive model and plausibility as an alignment with user knowledge [4] (see next item), but also accuracy to ground-truth [6] or alternative solutions with diversity [7]. The post-doc may produce a reference implementation library for these metrics.
2. Model prior knowledge from BRGM experts and derive new quality metrics for explanations based on their covering, novelty, or interestingness. Modeling of prior knowledge will likely be based on causal models and knowledge graphs. Novelty and interestingness metrics will take inspiration from what is done in Exploratory Data Analysis (EDA) [2] that have a long record of research works to guide the user towards interesting patterns or insights hidden in very large databases.
3. Build a user-oriented protocol and carry out user studies with BRGM experts.  
The goal is to qualify how well quality measures match the expert's expectations in terms of quality, and whether it is possible to learn better

explanation quality measures (i.e. depending on the context of the analysis and previous analysis as done in exploratory data analysis [9]). Different use cases will be considered, such as debugging the ML model for digital twins, understanding predicted parameters for a simulation model like Gardenia, and observing common and distinctive important factors from one predictive problem to the next (e.g. analysis of similarities between prediction of underground water resources in different places).

In conjunction with the ARD JUNON project team, the successful candidate will be responsible for implementing the research program outlined in the preceding three points in consultation with their supervisor and the project partners. The successful candidate will also be involved in the animation of research groups interested in XAI such as (but not limited to) Explain'AI (GT EGC), Help (GdR Madics), Explicon (GdR RADIA).

### 3 Collaboration

The recruited candidate will be supervised by Nicolas Labroche (LIFAT, University of Tours), in partnership with BRGM (the French Geological Survey) and the PhD student and post-doc already at work in the *Prediction* project to develop new ML models.

Other external collaboration will be conducted with experts in XAI from other labs such as IRIT (Toulouse), ETIS and LIP6 (Paris).

### 4 Skills / Qualifications

PhD in Computer Science, specializing in artificial intelligence (explainability, possibly deep learning), experience in processing temporal data (multivariate time series and multivariate event sequences). Experience working with libraries offering implementations of XAI and deep learning models. Experience in setting user protocol would be appreciated.

### 5 General information

- **Work location (in-office):** the hired candidate will work in Tours, Faculté des Sciences et Techniques, avenue Monge 37200 Tours. This is a fully in-office position, although one day a week of remote work may be allowed by the supervisor.
- **Duration:** 18 months, starting from January 2025
- **Gross salary:** 2800 € /month

- **To apply:** Please, send CV + covering letter; list of publications; references with their e-mail addresses, affiliations and the contexts in which the candidate has worked with them, to be sent exclusively by e-mail. The deadline to apply is **1st december 2024, 12H00**. Applications will be accepted until the position is filled.

Contact Pr. Nicolas Labroche to obtain further information: [nicolas.labroche](mailto:nicolas.labroche@univ-tours.fr) [at] [univ-tours.fr](http://univ-tours.fr), Dr. Alexandre Chanson [alexandre.chanson](mailto:alexandre.chanson@univ-tours.fr) [at] [univ-tours.fr](http://univ-tours.fr)

## References

- [1] *xxAI - Beyond Explainable AI - International Workshop, Held in Conjunction with ICML 2020, July 18, 2020, Vienna, Austria, Revised and Extended Papers*, volume 13200 of *Lecture Notes in Computer Science*, 2022.
- [2] Sihem Amer-Yahia, Patrick Marcel, and Verónica Peralta. Data narration for the people: Challenges and opportunities. In *EDBT 2023, Ioannina, Greece, March 28-31, 2023*, pages 855–858, 2023.
- [3] Sebastian Bordt, Michèle Finck, Eric Raidl, et al. Post-hoc explanations fail to achieve their purpose in adversarial contexts. In *2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT '22*, June 2022.
- [4] Alon Jacovi and Yoav Goldberg. Aligning faithful interpretations with their social attribution. *Trans. Assoc. Comput. Linguistics*, 9:294–310, 2021.
- [5] Alessio Malizia and Fabio Paternò. Why is the current XAI not meeting the expectations? *Commun. ACM*, 66(12):20–23, 2023.
- [6] Christoph Molnar. *Interpretable Machine Learning*. 2 edition, 2022.
- [7] Ramaravind Kommiya Mothilal, Amit Sharma, and Chenhao Tan. Explaining machine learning classifiers through diverse counterfactual explanations. In *FAT\*’20: Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, January 27-30, 2020*, pages 607–617, 2020.
- [8] Meike Nauta, Jan Trienes, Shreyasi Pathak, et al. From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. *ACM Computing Surveys*, 55(13s):1–42, July 2023.
- [9] Junjie Xing, Xinyu Wang, and H.V. Jagadish. Data-driven insight synthesis for multi-dimensional data. *Proceedings of the VLDB Endowment*, 17, 2024.