Post-Doctoral position

Knowledge integration in Machine learning for time series prediction of ground water level

Profile: PhD in machine learning (computer science or applied mathematics)

Duration: 1 year contract renewable

Affiliation: LIFO (Laboratoire d'Informatique Fondamentale d'Orléans) - Constraints and Machine learning (CA) team.

Gross salary: around 2600€/month

Supervisors: Thi-Bich-Hanh Dao, Vincent Nguyen, Christel Vrain

Skills

- Good experience in data analysis and machine learning is required.
- Experiences/knowledge in time series prediction and environmental science is welcome.
- Curiosity and ability to communicate (in English or in French) and to work in collaboration with scientists in environmental science.
- Ability to propose and validate new solutions and to publish the results.
- Autonomy and good organizational skills.

How to candidate

Candidates are invited to send a pdf file that contains:

- A CV, with descriptions on your thesis and experiences in machine learning or deep learning (including projects you were involved in)
- A motivation letter
- References from academics

by email to thi-bich-hanh.dao[at]univ-orleans.fr and vincent.nguyen[at]univ-orleans.fr as soon as possible and no later than 15th of June 2024.

Context

The JUNON project is granted from the Centre-Val de Loire region through an ARD program (Ambition Recherche Développement). The project is driven by BRGM and involves BRGM, University of Orléans (LIFO), University of Tours (LIFAT), CNRS, INRAE, ATOS and ANTEA companies. The main goal of JUNON is to develop digital twins to improve the monitoring, understanding and prediction of environmental resources evolution and phenomena, for a better management of natural resources. Digital twins will allow us to virtually reproduce natural processes and phenomena using combinations of AI and environmental tools. They will rely on geological and meteorological data (time series) and knowledge, as well as physical-based models.

JUNON project is organized into 5 work packages (WP):

1. User's needs and geological knowledge for ground water
2. User's needs and biological/chemical knowledge about pollutants and greenhouse gases
3. Data management and data mining
4. Times series predictions
5. Aggregation and realization of digital twins

The postdoc program will be supervised by LIFO-CA and will be in WP4, focusing on the prediction of quantity/level of ground waters. There will be strong interactions inside WP4 with other postdocs and PhD in LIFO or LIFAT, with WP1 and WP3 (BRGM) with engineers.

The CA team is a dynamic team with 8-10 PhD. We work on Machine Learning, Data Mining and Deep Learning and have been interested in knowledge integration in ML/DM methods.

Objectives

In ground water level predictions, physical-based models or classic AI tools have achieved good performance in short term predictions, for instance up to 3 months. The performance, however, worsens for a more long-term prediction, such as for instance up to 1 year or more. Recently, several works have shown the interest of hybrid models, that combine both physical and AI models, in environmental science.

The goal of this work is to study how expert knowledge could be integrated to improve predictors. Expert knowledge can come from different sources. It may be information such as seasonal cycles, soil or subsoil natures that may impact on the prediction. Some physic-based models have already been developed, either global or distributed, these models encapsulate some expert knowledge that could be used to guide AI models. The aim of the postdoctoral program is to build new prediction models that take advantage of both physical-based and AI models and to study the integration of expert knowledge.

We have developed methods integrating prior knowledge into deep learning models in clustering tasks or in image classification tasks. We are interested in either pursuing this approach or considering physics-informed neural networks which is a hot topic.

Bibliography

- Tao Bai a, Pejman Tahmasebi, Graph neural network for groundwater level forecasting, Journal of Hydrology 616, 2023
- Wunsch, A., Liesch, T., & Broda, S. (2021). Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrology and Earth System Sciences*, *25*(3), 1671-1687, 2021.
- Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Computing Surveys*, *55*(4), 1-37, 2022.
- Muther, T., Dahaghi, A. K., Syed, F. I., & Van Pham, V. (2023). Physical laws meet machine intelligence: current developments and future directions. *Artificial Intelligence Review*, *56*(7), 6947-7013, 2023.

- Nguyen-Viet-Dung Nghiem, Christel Vrain, Thi-Bich-Hanh Dao: Knowledge Integration in Deep Clustering. ECML/PKDD (1) 2022: 174-190
- Franck-Anael Mbiaya, Christel Vrain, Frédéric Ros, Thi-Bich-Hanh Dao, Yves Lucas: Knowledge graph-based image classification, *Data & Knowledge Engineering,* Volume 151, 102285, 2024.