2023-06550 - Machine Learning Engineer: Enhancement of ReservoirPy - the Python Reservoir Computing Library

Type de contrat : CDD
Niveau de diplôme exigé : Bac + 5 ou équivalent
Fonction : Ingénieur scientifique contractuel

Contexte et buts du poste
ReservoirPy is a highly user-friendly library based on Python's scientific modules. With approximately 250 stars on Github, it stands as the most widely used and regularly updated Python library in the domain of Reservoir Computing.

Reservoir Computing is a well-established method that has been proven to outperform other techniques in complex tasks, such as chaotic time series prediction [3, 5]. Since its inception, it has found application in a multitude of domains [4]. Despite its low computational cost and less dependence on vast training data compared to LSTMs [2], it is currently underutilized, overshadowed by the prevailing deep learning trend. To address this, a new Python library, ReservoirPy [7], was recently developed. This cutting-edge tool allows for the easy development of complex models [8], and already integrates several advanced features drawn from the literature.

ReservoirPy offers a flexible interface to efficiently implement Reservoir Computing (RC) [4] architectures, with a specific emphasis on Echo State Networks (ESN) [3]. The library's advanced features significantly enhance computation time efficiency, even on a basic laptop, in comparison to standard Python implementation. Its suite of features includes offline and online training, parallel implementation, sparse matrix computation, rapid spectral initialization, advanced learning rules (such as Intrinsic Plasticity), and more. Furthermore, it facilitates the creation of complex architectures with multiple reservoirs (e.g., deep reservoirs), readouts, and intricate feedback loops. For enhanced user experience, the library is equipped with graphical tools to explore hyperparameters, aided by the hyperopt library. Comprehensive tutorials explicating exotic architectures, as well as examples of scientific paper reproductions, are included for deeper understanding and practical application.

Github: https://github.com/reservoirpy/reservoirpy
Documentation: https://reservoirpy.readthedocs.io/
Related projects: https://github.com/reservoirpy

Mission confiée
The engineer will be working on developing the following new features of ReservoirPy (see "Main activities").

The first weeks will be dedicated to taking charge of the ReservoirPy library and creating new tutorials on ReservoirPy. Then the recruited engineer will work on the various features that will be created, while interacting with other people working with ReservoirPy in the team and outside the team.

We have been extensively developing ReservoirPy during the past 3 years. The library is now mature to be used in complex scenarios with the high flexibility provided by the library. The recruited engineer will be required to maintain the library by answering issues on the Github repository, continue to enhance the current documentation, tests, create new tutorials, etc.

This job is highly collaborative: it will happen within the Mnemosyne Inria team. There will be weekly interactions with the team members using ReservoirPy (PhD students, ..) and also with other research teams using ReservoirPy (e.g. with application on health).

Informations générales

- Thème/Domaine : Optimisation, apprentissage et méthodes statistiques
- Ingénierie logicielle (BAP E)
- Ville : Bordeaux
- Centre Inria : Centre Inria de l'Université de Bordeaux
- Date de prise de fonction souhaitée : 2023-10-01
- Durée de contrat : 2 ans
- Date limite pour postuler : 2023-08-03

Contacts

- Équipe Inria : MNEMOSYNE
- Recruteur : Hinaut Xavier / xavier.hinaut@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2 600 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3 500 scientifiques pour relever les défis du numérique, souvent à l’interface d’autres disciplines. L’institut met à disposition de nombreux talents dans plus d’une quarantaine de métiers différents. 900 personnels d’appui à la recherche et à l’innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde.

Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 180 start-up. L’institut s’efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l’économie.

L’essentiel pour réussir

We are looking for a skilled and passionate Software Programmer to help us extend the capabilities of our ReservoirPy library, through rigorous analysis and innovative method development. If you have a strong proficiency in the Python scientific software stack, predominantly numpy and scipi, and possess a keen understanding of GitHub, we would like to interview you.

Attention : Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.
The main new features that will be developed are the following:
- Creation of backend management system (NumPy/Cupy or PyTorch)
- Parallelisation of computations for any kind of model
- Creation of spiking models of reservoirs
- Integrate tools to analyse reservoir dynamics and better control reservoir topologies

It will be performed along with other “smaller” features:
- Creation of modifier of behavior of RNNs (bidirectionality, ...) and creation of Conceptors (Jaeger 2014)
- Update the interfacing between ReservoirPy and Scikit-Learn
- Improvement of the hyper-parameter search tool and implementation of new optimization methods
- Integration of other recent features from the Reservoir Computing literature

This will be done jointly with the implementation of the development infrastructure (Travis, Coverage, ReadTheDocs) along with transversal tasks:
- Development + testing + documentation
- Help in the organisation of Hackathons on Reservoir Computing
- Writing of scientific articles in English (conferences and journal papers)
- Management of external pull-requests + correction of possible bugs found by the community
- Weekly reporting of advances and problems.

Compétences
Ideal Candidate Profile:
- Holds an engineering or scientific degree and/or a PhD in digital sciences (computer science, automation, signal processing).
- Has a first professional experience or internship (6 months to 5 years) in one or more of the following areas: Machine Learning, Python Development, and GPU/Parallel Programming
- Possesses strong expertise in the scientific Python software and scientific stack (numppy/scipy).
- Demonstrates a solid grasp of linear algebra concepts.
- Proficiency in technical English is crucial, as it enables efficient collaboration with our international partners and effective presentations at conferences.
- Has proven experience with version management, familiarity with Git, and proficiency in using the GitHub platform.

We are also looking for someone with the following general abilities:
- Ability to work in a multidisciplinary team.
- Ability to adapt to the project context.
- Knows how to establish a relationship of trust with their interlocutors.
- Has good written and oral communication skills in French.
- Masters technical and scientific English.

Avantages
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of partial teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training

Rémunération
remuneration according to qualifications and professional experiences: between 2692 euros and 3509 euros gross per month