2019-02001 - Post-Doctoral Research Visit F/M Post-doc position Pattern mining for Neural Networks debugging: application to speech recognition

**Type de contrat :** CDD de la fonction publique  
**Niveau de diplôme exigé :** Thèse ou équivalent  
**Fonction :** Post-Doctorant  
**Niveau d’expérience souhaité :** jeune diplômé

**A propos du centre ou de la direction fonctionnelle**

Inria, the French national research institute for the digital sciences, promotes scientific excellence and technology transfer to maximise its impact. It employs 2,400 people. Its 200 agile project teams, generally with academic partners, involve more than 3,000 scientists in meeting the challenges of computer science and mathematics, often at the interface of other disciplines. Inria works with many companies and has assisted in the creation of over 160 startups. It strives to meet the challenges of the digital transformation of science, society and the economy.

**Contexte et atouts du poste**

**Context:** Understanding the inner working of deep neural networks (DNN) has attracted a lot of attention in the past years [1, 2] and most problems were detected and analyzed using visualization techniques [3, 4]. Those techniques help to understand what an individual neuron or a layer of neurons are computing. We would like to go beyond this by focusing on groups of neurons which are commonly highly activated when a network is making wrong predictions on a set of examples. In the same line as [1], where the authors theoretically link how a training example affects the predictions for a test example using the so called “influence functions”, we would like to design a tool to “debug” neural networks by identifying, using symbolic data mining methods, (connected) parts of the neural network architecture associated with erroneous or uncertain outputs.

In the context of **speech recognition**, this is especially important. A speech recognition system contains two main parts: an acoustic model and a language model. Nowadays models are trained with deep neural networks-based algorithms (DNN) and use very large learning corpora to train an important number of DNN hyperparameters. There are many ways to automatically tune these hyperparameters. However, this induces a huge computational cost, and does not empower the human designers. It would be much more efficient to provide human designers with understandable clues about the reasons for the bad performance of the system, in order to benefit from their creativity to quickly reach more promising regions of the hyperparameter search space.

**Mission confiée**

**Description of the position:** This position is funded in the context of the **HyAI and “Hybrid Approaches for Interpretable AI” Inria project lab** (https://www.inria.fr/en/research/researchteams/inria-project-labs/). With this position, we would like to go beyond the current common visualization techniques that help to understand what an individual neuron or a layer of neurons is computing, by focusing on groups of neurons that are commonly highly activated when a network is making wrong predictions on a set of examples. Tools such as activation maximization [8] can be used to identify such neurons. We propose to use discriminative pattern mining, and, to begin with, the DiffNorm algorithm [6] in conjunction with the LCM one [7] to identify the discriminative activation patterns among the identified neurons.

The data will be provided by the MULTISPICE team and will consist of two deep architectures as representatives of acoustic and language models [9, 10]. Furthermore, the training data will be provided, where the model parameters ultimately derive from. We will also extend our results by performing experiments with supervised and unsupervised learning to compare the features learned by these networks and to perform qualitative comparisons of the solutions learned by various deep architectures. Identifying “faulty” groups of neurons could lead to the decomposition of the DL network into “blocks” encompassing several layers. “Faulty” blocks may be the first to be modified in the search for a better design.

The recruited person will benefit from the expertise of the LACODAM team in pattern mining and deep learning (https://team.inria.fr/lacodam/) and of the expertise of the MULTISPICE team (https://team.inria.fr/multispice/) in speech analysis, language processing and deep learning.

**Bibliography:**


**Informations générales**

- **Thème/Domaine :** Représentation et traitement des données et des connaissances  
- **Ville :** Rennes  
- **Centre Inria :** CRI Rennes - Bretagne Atlantique  
- **Date de prise de fonction souhaitée :** 2019-12-01  
- **Durée de contrat :** 1 an  
- **Date limite pour postuler :** 2019-11-05

**Contacts**

- **Equipe Inria :** LACODAM  
- **Recruteur :** Fromont Elisa / elisa.fromont@irisa.fr

**A propos d’Inria**

Inria, the institute national de recherche dédié aux sciences du numérique, promeut l’excellence scientifique et le transfert pour avoir le plus grand impact. Il emploie 2400 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l’interface d’autres disciplines. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 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 would ideally like to recruit a 1 year (with possibly one additional year) post-doc with the following preferred skills:

- Some knowledge (interest) about speech recognition
- Knowledgeable in pattern mining (discriminative pattern mining is a plus)
- Knowledgeable in machine learning in general and deep learning particular
- Good programming skills in Python (for Keras and/or Tensor Flow)
- Very good English (understanding and writing)

The position will be funded by INRIA (https://www.inria.fr/en), See the INRIA web site for the post-doc and PhD wages.

**Consignes pour postuler**

The candidates should send a CV, 2 names of referees and a cover letter to the four researchers (firstname.lastname@inria.fr) mentioned above.

Please indicate if you are applying for the post-doc or the PhD position.

The selected candidates will be interviewed in October for an expected start in December 2019.

**Sécurité défense :**  
Ce poste est susceptible d’être affecté dans une zone à régime restrictif (ZRR), tel que défini dans le décret n°2011-
**Principales activités**

The recruited person will benefit from the expertise of the LACODAM team in pattern mining and deep learning (https://team.inria.fr/lacodam/) and of the expertise of the MULTISPEECH team (https://team.inria.fr/multispeech/) in speech analysis, language processing and deep learning.

**Compétences**

We would ideally like to recruit a 1 year (with possibly one additional year) post-doc with the following preferred skills:

- Some knowledge (interest) about speech recognition
- Knowledgeable in pattern mining (discriminative pattern mining is a plus)
- Knowledgeable in machine learning in general and deep learning particular
- Good programming skills in Python (for Keras and/or Tensor Flow)
- Very good English (understanding and writing)

The position will be funded by INRIA (https://www.inria.fr/en/). See the INRIA web site for the post-doc and PhD wages.

**Avantages**

- Subsidized meals
- Partial reimbursement of public transport costs

**Rémunération**

PhD student : monthly gross salary amounting to 1982 euros for the first and second years and 2085 euros for the third year

Post-doc : monthly gross salary amounting to 2653 euros

---

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.