Offer #2023-06973

CDD 24 months - Engineer Position F/H - Translating with large language models in low-resource scenarios and for unseen languages

Contract type: Fixed-term contract
Level of qualifications required: Graduate degree or equivalent
Fonction: Temporary scientific engineer

Context

Context of the project

Within the span of six short years (2017-2023), the field of Natural Language Processing (NLP) has been transformed by the advances of general-purpose neural architectures, which are both used to learn representations for linguistic units and to generate high-quality textual content. These architectures are nowadays ubiquitous in NLP applications; trained at scale, these “large language models” (LLMs) offer multiple services (summarisation, writing aids, translation) in one model through human-like conversations and prompting techniques.

The aim of the recently accepted TraLaLam ANR project is to analyse the new state of play from the perspective of machine translation (MT) and ask two main questions: (a) prompting techniques make it straightforward to inject various types of contextual information that could help an MT system to take context into specific account such as to adapt to a domain, a genre, a style, to a client's translation memory, to the readers' language proficiency, etc. Is prompting equally effective for all these situations, assuming good prompts can be generated, or is it hopeless to expect improvements without (instruction) fine-tuning? (b) as LLMs can be trained without any parallel data, they open the perspective of improved MT for multiple language pairs, domains and styles for which such resources are scarce if they exist at all. Can this promise be held, especially for low-resource dialects and regional languages? To address these two questions, TraLaLaM will also collect data for low-resource languages (and use them to extend existing LLMs) and develop new testing corpora and associated evaluation strategies.

Assignment

Activities of the position

The role of the engineer will be to work carry out experiments on point (b) outlined above (Context of the Project): using LLMs for low-resource scenarios on which they were not originally trained. In particular, this means implementing approaches for adapting LLMs to unseen languages by using embedding-based approaches (e.g. Kumar et al. 2021), multilingualism and transfer with similar (seen) languages (e.g. Song et al. 2023) and synthetic data creation (e.g. Tars et al. 2021).

The engineer will be co-supervised by Rachel Bawden and Benoît Sagot at Inria Paris (France) in the ALMAnaCH project-team in the context of the ANR project TraLaLam, in collaboration with SYSTRAN and Sorbonne Université (ISIR, CNRS).

Main activities

The main activities of the position will include:

- keeping up-to-date with related work on the topic and producing a report on the related work
- assisting research on the topic outlined above and carrying out experiments
- the presentation of work both internally to colleagues and externally in the form of conference/journal/workshop papers
- interacting and exchanging with colleagues and project partners on NLP topics

Skills

The position is a 2-year funded research engineer position of starting from the 1st March 2024 at the earliest. Candidates should have a Master 2 or equivalent (e.g. engineering school) in computer science (speciality artificial intelligence, machine learning or natural language processing). They should have a good level in programming (python), experience with neural networks and an interest in natural language processing. A good written and spoken level of English is required and a good level of French would be a plus.
Qualities sought:
We are looking for highly motivated candidates with a strong background in NLP and machine learning. Ideally, candidates should be able to show initiative, creativity and have a good eye for analysis of data and results.

Benefits package
- 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 teleworking
- DFlexible organization of working hours (after 12 months of employment)
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

General Information
- Theme/Domain: Language, Speech and Audio
- Town/city: Paris
- Inria Center: Centre Inria de Paris
- Starting date: 2023-03-01
- Duration of contract: 2 years
- Deadline to apply: 2024-01-21

Contacts
- Inria Team: ALMANACH
- Recruiter: Bawden Rachel / rachel.bawden@inria.fr

About Inria
Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

Warning: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply
In your application (which can be in English or in French), please include:
- CV
- Letter of motivation
- Letter(s) of recommendation
- Optionally an example of your previous written work (if possible related to NLP), for example a master’s thesis, research paper, etc.

Defence Security:
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy:
As part of its diversity policy, all Inria positions are accessible to people with disabilities.