Internship in ESA's Advanced Concepts Team
On
Neural information retrieval

European Space Research and Technology Centre
ESA ESTEC

Candidates interested are encouraged to visit the ESA website: www.esa.int/gsp/ACT/

Topic description

An emerging trend in information retrieval is the application of deep and shallow neural networks to support typical tasks, like the ranking of documents in correspondence to an ad-hoc query. With an increasing interest on this topic on all major AI conferences, neural information retrieval promises to deliver results of higher relevance than traditional approaches. The potential to distill large and unstructured data into meaningful representations provides a unique opportunity to accelerate research in general.

Since advances in the space sector are fuelled by interconnection of engineering with multiple scientific disciplines, the ESA and specifically its Advanced Concepts and Studies Office with the ACT has a rich history of multi-disciplinary studies. In 2019, it opened its Open Space Innovation Portal (OSIP, ideas.esa.int) as a new entry point for external ideas to the early stages of the ESA R&D pipeline. The goal of this internship is to explore the potential of neural information retrieval related to the missions and activities of ESA with a particular focus on detecting novelty levels of new ideas as well as the identification and highlighting of similar ideas or ESA projects.

Candidate's tasks

In this internship the selected candidate will look into recently proposed methods [1] and implement prototypes. This could include one or more of the following:

• Application of Unsupervised Learning of text representations using shallow neural networks
• Supervised Learning for Document ranking tasks (deploying either deep or shallow networks)
• Supervised Training of Deep neural networks on raw and unstructured representations of document and query data
• Word/Paragraph/Document/Concept Embeddings
• Deep Semantic Hashing
• Question-Answering Systems or Conversational Agents

The ideal candidate

Mandatory:

• Excellent programming skills in Python, including PyTorch or Tensorflow
• Knowledgeable in (applied) Machine Learning
• Experience in conducting traditional information retrieval tasks (e.g. text mining, topic

modelling, text clustering, categorization, etc.)

Desirable:

• Background in Natural Language Processing (NLP) and related frameworks (preferably NLTK)
• Experience with Word Vectorizers (e.g. word2vec or GloVe)
• Experience with Recurrent Neural Networks (RNNs) and/or Siamese Networks
• Knowledge in (REST-) API usage and web scrapping techniques, e.g. Python, BeautifulSoup

References

[1] Mitra, B., & Craswell, N. (2018). An introduction to neural information retrieval. Foundations and Trends® in Information Retrieval, 13(1), 1-126.