Internship in ESA's Advanced Concepts Team

on

Neural information retrieval

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 distil 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 Advanced Concepts Team (ACT) of the European Space Agency (ESA) has a rich history of multi-disciplinary studies. 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 the innovative projects and outcomes generated by the ACT and the application of this to the analysis of the novelty level of new ideas.

Candidate's tasks

In this internship the selected candidate will look into recently proposed methods [1] and implement a 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 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.

References

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