



HAMLET August 19 - 21, 2024
Copenhagen, Denmark
How to Apply Machine Learning to
Experimental & Theoretical
PHYSICS

Contribution ID: 52

Type: **Parallel or poster**

Chatbots for astrophysicists

Tuesday, 20 August 2024 14:15 (20 minutes)

Large Language Models (LLMs) have become extremely popular over the past year and a half. Researchers have experimented with using LLM-powered chatbots (usually Chat-GPT) as an “assistant” to help them understand existing work and even discuss new research ideas. However, a pure LLM chatbot tends to make plausible sounding but unsupported statements, and cite non-existing papers, which severely limits its usefulness. Retrieval Augmented Generation (RAG) restricts an LLM chatbot to using only a particular body of literature to answer questions. The answer is extracted from retrieved text, and the sources (papers, sections, paragraphs) are listed. Therefore a RAG chatbot can greatly support literature search and research more generally.

The field of astrophysics is ideally suited for experimenting with this technique. The obstacles are limited, as scientific literature and technical documentation tend to be readily available, and the researchers have a high level of technical literacy and affinity. At the same time, the potential benefits are high, as researchers often need to familiarize themselves with new instruments and techniques.

In collaboration with the NLP team at ESA we are developing a RAG chatbot using Gaia manuals and science papers. Gaia data products are complex and will remain relevant for decades, making it important that helpful information remains easily accessible for astrophysicists. A working prototype can be accessed at <https://gaiachat.streamlit.app>. In parallel, we are working on a similar system for the upcoming PLATO mission (<https://platochat.streamlit.app>).

We are exploring the potential of supporting students, as well as academia - industry collaborations, by developing a smart portal that helps students in finding well-fitting internship opportunities. Here there will be a central sign-up for industry partners, and students can then search for opportunities that match their interests and skills in dialog with the system. We expect that this will encourage more students to do an internship, leading to positive experiences for both students and industry partners, and increased collaboration.

Primary authors: RECKMAN, Hilke (Veridicity); ALBRECHT, Simon (Aarhus University)

Presenter: ALBRECHT, Simon (Aarhus University)

Session Classification: Parallel