Contribution ID: **32**   Type: **not specified**

Exercise: Understanding Large Language Models and their Applications in Code Generation

Tuesday 25 March 2025 15:15 (1 hour)

Lecture

The field of Deep Learning has experienced a significant turning point in the past years, driven by the emergence and rapid development of Large Language Models (LLMs). These advanced models have not only redefined standards in Natural Language Processing (NLP) but are increasingly being integrated into applications and services due to their natural language capabilities.

Interest in using LLMs for coding has grown rapidly, and some companies have sought to transform natural language processing into code generation. However, this trend has already exposed several unresolved challenges in applying LLMs to coding. Despite these difficulties, it has spurred the development of AI-powered code generation tools, such as GitHub Copilot.

In this lecture, we will introduce Large Language Models and explore the related terminology and core components. We will examine the process of creating an LLM from scratch and highlight the requirements for developing domain-specific LLMs, such as those for code generation. Additionally, we will review the initial footprints of these technologies in code generation and discuss their limitations in this domain. We will also explore strategies and architectural approaches that improve the performance of LLMs generally and for code generation specifically.

We will conclude by addressing the ethical concerns surrounding LLMs, such as the authorship and ownership of AI-generated code. Finally, we will explore other applications of LLMs in science, particularly within the High Energy Physics community and at CERN.

Hands-on

I think it could be interesting (but optional) to have a hands-on session, including:

1. Some exercises to interact with an LLM from Python (either from a small pre-trained model in PyTorch or through API calls).
2. Some exercises to explore domain-specific models for code generation.
3. Hands-on exploration and testing of the LLMs components.
4. Prompt Engineering strategies or small Fine-Tuning to improve the performance of a model.
5. Some simplified strategies such as Retrieval Augmented Generation could also be implemented.

Attended school

CSC 2023 (Tartu)

Number of exercise hours

1

Number of lecture hours

1

Author: VALENZUELA RAMIREZ, Andrea (CERN)

Presenter: VALENZUELA RAMIREZ, Andrea (CERN)