

Wykorzystanie infrastruktury HPC w uczeniu maszynowym Dużych Modeli Językowych

Wednesday, 13 March 2024 15:00 (4h 30m)

Termin:

środa 13 marzec 2023, 15:00-19:30,
z przerwą kawową po ok. 2 godzinach od rozpoczęcia.

Miejsce:

Bel-Ami, ul. Goszczyńskiego 24, Zakopane

Prowadzący:

- * Szymon Mazurek (ACK Cyfronet AGH)
- * Krzysztof Wróbel (UJ, Enelpol)
- * Aleksander Smywiński-Pohl (AGH, Enelpol)
- * Klemens Noga (ACK Cyfronet AGH)

Zagadnienia

Szkolenie jest przeznaczone dla osób planujących efektywne wykorzystanie zasobów Komputerów Dużej Mocy zainstalowanych w Cyfronecie do badań wykorzystujących Duże Modele Językowe (ang. Large Language Models) w uczeniu maszynowym. Jest skierowane do użytkowników, którzy chcieliby jak najszybciej zdobyć praktyczną wiedzę pozwalającą na samodzielną pracę z własnymi zbiorami, przy wykorzystaniu najpopularniejszych dzisiaj narzędzi.

W ramach szkolenia uczestnicy zostaną wprowadzeni w zagadnienia wykorzystania Dużych Modeli Językowych przy użyciu języka Python. Szkolenie składać się będzie z krótkiego teoretycznego wstępu opisującego zagadnienia związane z Dużymi Modelami Językowymi oraz ćwiczeń wykonywanych na zasobach obliczeniowych udostępnianych przez Cyfronet. W trakcie szkolenia zostaną przedstawione praktyczne wskazówki jak dostosować strategie uczenia maszynowego do zasobów HPC, w tym sposoby efektywnego ładowania danych, doboru parametrów treningu oraz rozpraszania uczenia na wiele kart GPU.

Szkolenie będzie składać się z modułów:

1. Wprowadzenie teoretyczne, w którym krótko zostaną przedstawione podstawowe zagadnienia.
2. Warsztaty z podstawowego wykorzystania LLM:
 - a. ekosystem HuggingFace
 - b. kwantyzacja i inne metody ułatwiające wykorzystanie LLM
 - c. rozpraszanie obliczeń na wiele kart GPU dla dużych LLM
 - d. benchmarking LLM
3. Dostosowanie modeli LLM do konkretnych zastosowań oraz zbiorów danych
 - a. Pre-trening, fine-tuning i prompting modeli językowych
 - b. Wymagania pamięciowe treningu LLM
 - c. Przegląd technik do fine-tuningu
4. Mini projekt polegający na samodzielnym treningu oraz benchmarku wybranego modelu

Presenters: SMYWIŃSKI-POHL, Aleksander (AGH / Enelpol); NOGA, Klemens (ACK Cyfronet AGH); WRÓBEL, Krzysztof (AGH / Enelpol); MAZUREK, Szymon (ACK Cyfronet AGH)

Session Classification: Training - Wednesday