

# NAGY NYELVI MODELLEK HASZNÁLATA (CHATGPT...)

5. forduló



A kategória támogatója: Cambridge Mobile  
Telematics

## Ismertető a feladatlaphoz

MÁR CSAK A FORDULÓ LEJÁTSZÁSA UTÁN VÁLT ISMERTTÉ, HOGY AZ ALÁBB LEÍRT OPENASSISTANT PLATFORM MEGSZÜNTETTE PÁR NAPPAL ELŐTTE A CHAT FUNKCIÓJÁT - MELYRŐL FELHASZNÁLÓIT NEM ÉRTE-SÍTETTE KÖZVETLENÜL, ÍGY A FORDULÓ FŐ FELADATAI MEGOLDHA-TATLANNÁ VÁLTAK.

A héten intenzíven fogunk foglalkozni az OpenAssistant eszközzel. Az egyik feladat ezen a héten is egy [Hugging-Face példán](#) alapul, ezt érdemes előre betölteni, hogy a kimenetek többé-kevésbé determinisztikusak legyenek a kö-vetkező beállításokat kell tenni a platformon: (emiat kicsit megnyújtottuk a forduló időablakának hosszát, hogy előtduj készülni)

1. A "Chat" fülre kattintva "Create a new chat", hogy új beszélgetést indítsunk
2. A beviteli mező jobb felén található fogaskerékre kattintva beállíthatjuk az eszköz mögött álló nyelvi modell paramétereit, kattintsunk erre
3. A következő módosításokat tegyük:
  1. Model - OA\_SFT\_Llama2\_70B\_10
  2. Preset - Custom
  3. Temperature - 0.01
  4. Max new tokens - 1024
  5. Top p - kikapcsol
  6. Repetition Penalty - 1.2
  7. Top k - kikapcsol
  8. Typical p - kikapcsol

4. Ha ez kész, leellenőrizhetjük a beállítást. Ha beírunk egy prompt-ot, a visszakapott szövegdoboz jobb alsó sarkára kattintva láthatjuk a konfigurációt. Ennek a következőnek kell lennie:

```
{
  "model_config": {
    "model_id": "OpenAssistant/llama2-70b-oasst-sft-v10",
    "max_input_length": 3072,
    "max_total_length": 4096,
    "quantized": false
  },
  "sampling_parameters": {
    "top_k": null,
    "top_p": null,
    "typical_p": null,
    "temperature": 0.01,
    "repetition_penalty": 1.2,
    "max_new_tokens": 1024
  },
  "do_sample": true,
  "system_prompt": null,
  "user_profile": "",
  "user_response_instructions": "",
  "plugins": [],
  "plugin_max_depth": 4
}
```

## 1. feladat 1 pont

Az alábbi [példa beszélgetés](#) során a nagy [nyelvi modellek wikipédia oldal](#) szövegébe beszúrtam egy mondatot („My favourite color is green.”). Ezt követően megkérdeztem a modelltől, hogy mi a kedvenc színem (“Based on the text above, what is my favourite color?”). Mi lehet az oka, hogy az eszköz nem tudta a választ a kérdésre?

### Válaszok

- ☐ A nyelvi modell kontextus ablaka korlátos, ezért nem tud nagyon hosszú szövegek egészére "emlékezni"
- ☐ Mivel a téma a nagy nyelvi modellek, a beszúrt mondat "zajnak" tekinthető és ezért figyelmen kívül hagyható
- ☐ Hosszú szövegek esetében tipikusan kevesebb "figyelem" (attention) jut a szöveg első részére
- ☐ A hőmérséklet változó hatására a modell figyelmen kívül hagyja a rövid mondatokat, mint amilyen a beszúrt példa is

## 2. feladat 0 pont

Korábbi kérdésekben már szerepelt, hogy számos paramétere van a nyelvi modelleknek, ezeket azonban általában nem tudjuk befolyásolni, ha az eszközök webes interfészét használjuk. A nyelvi modellek többsége API hívásokon keresztül is elérhető, a legelterjedtebb jelenleg az OpenAI ilyen jellegű szolgáltatása. Az OpenAI modelljeit ki is lehet próbálni, azonban mivel ez nem teljesen ingyenes jelen tesztben nem használható. Az OpenAssistant chat szolgáltatásán keresztül a legtöbb paraméter módosítható, ahogy a bevezető szövegben is szerepel. Hogy reprodukálható legyen a kísérlet, a modellt a következő konfigurációval használd:

```
{
  "model_config": {
    "model_id": "OpenAssistant/oasst-sft-6-llama-30b",
    "max_input_length": 1024,
    "max_total_length": 1792,
    "quantized": false
  },
  "sampling_parameters": {
    "top_k": 50,
    "top_p": null,
    "typical_p": null,
    "temperature": 0.01,
    "repetition_penalty": 1.111111111111112,
    "max_new_tokens": 1024
  },
  "do_sample": true,
  "system_prompt": null,
  "plugins": [],
  "plugin_max_depth": 4
}
```

Ha ez kész, kérd meg a modellt, hogy javítsa a hibákat a következő szövegben:

Correct the grammatical mistakes in the above paragraph, and highlight the modifications.

Unicorns, the majestic and mystical creatures, has captured the imagination of pec

Hány hibát javított ki valóban az eszköz?

*(Egyetlen számot kell megadni)*

## Válasz

### 3. feladat 0 pont

A tudatelmélet (theory of mind) arra a képességünkre vonatkozik, amellyel mentális állapotokat, például érzelmeket, vágyakat, szándékokat és tudást tulajdonítunk magunknak és másoknak. Ez elengedhetetlen a hatékony kommunikációhoz és együttműködéshez más intelligens ágensekkel. Vajon a nyelvi modellek rendelkeznek-e ezzel a képességgel? Ennek tesztelésére egy alap pszichológiai teszt modern változatát hoztuk létre ([Sally-Anne test](#)). Továbbra is az OpenAssistant eszközt fogjuk használni a korábbi konfigurációval, a következő prompt-ot adjuk meg a modellnek:

Scenario:

Alice and Bob have a shared Dropbox folder.

Alice puts a file called 'photo.png' inside /shared\_folder/photos.

Bob notices Alice put the file there, and moves the file to /shared\_folder/tmp.

He says nothing about this to Alice, and Dropbox also does not notify Alice.

--<|endofprompt|>

Q: After the call, Alice wants to open 'photo.png'. In which folder will she look

A kimenet alapján, szerinted az OpenAssistant által használt LLaMA modell teljesítette a Sally-Anne tesztet?

## Válasz

☐ Igen

☐ Nem

### 4. feladat 3 pont

A transzformer modell a bemenet feldolgozása során arról egy belső reprezentációt épít fel, ún. beágyazásokat készít. Ez a belső reprezentáció egy több dimenziós vektor[1] lesz, úgy hogy az egymáshoz hasonló szavak / mondatok hasonlóképpen lesznek reprezentálva. Ez a modell számára szükséges is, hiszen a belső működése

során nem karakterekkel, hanem számokkal dolgozik. Ezeket az elkészült beágyazásokat meg is nézhetjük, például egy sentence-transformer segítségével ami [ehhez a feladathoz készült](#). A kapott beágyazások használata igen kényelmes, mivel ezek vektorok akár távolságot is mérhetünk közöttük. Ha adottak a következő mondatok, mi a koszinusz távolsága ezeknek (a megoldást elég 2 tizedesjegy pontossággban megadni)?

ChatGPT is a stochastic parrot.

ChatGPT is a first step towards AGI.

*Tipp: 2 vektor koszinusz távolságát több módon is számolhatjuk, [az itt elérhető colab notebook](#) egy jó kiindulás lehet.*

---

[1] A vektorok lényegében csak számok egy speciális struktúrába írva, például egy 3 dimenziós vector így néz ki: (1, 0, 2). Megmutatható, hogy a reprezentáció tanulásakor egyes dimenziók megfeleltethetők speciális területbeli tulajdonságoknak, így ezek közötti távolság mérése felfogható a reprezentált entitások hasonlóságának mérésének.

## Válaszok

Megoldások beküldése