

# NAGY NYELVI MODELLEK HASZNÁLATA (CHATGPT...)

1. forduló



A kategória támogatója: Cambridge Mobile  
Telematics

## Ismertető a feladatlaphoz

Kérjük, hogy a feladatlap indítása előtt mindenképp olvasd el az alábbi útmutatót:

MINDEN kérdésre van helyes válasz.

Olyan kérdés NINCS, amire az összes válasz helyes, ha mégis az összes választ bejelölöd, arra a feladatra automatikusan 0 pont jár.

Több válaszlehetőség esetén a helytelen válasz megjelölése mínusz pontot ér.

A radio button-os kérdésekre egy helyes válasz van.

Ha lejár a feladatlap ideje, a rendszer AUTOMATIKUSAN beküldi azt az addig megjelölt válaszokkal.

Azokat a feladatlapokat, amelyekhez csatolmány tartozik, javasoljuk NEM mobilon elindítani, erre az érintett feladatlapok előtt külön felhívjuk a figyelmet.

Az adatbekérős feladatokra NEM jár részpontszám, csak a feleletválasztósakra.

Helyezéseket a 4. forduló után mutatunk, százalékos formában: adott kategóriában a TOP 20-40-60%-hoz tartozol.

Ha egyszerre több böngészőből, több ablakban vagy több eszközről megnyitod ugyanazt a feladatlapot, nem tudjuk vállalni az adatmentéssel kapcsolatban esetlegesen felmerülő anomáliákért a felelősséget!

A LLM-ek használata nemhogy nem tiltott ezúttal, de melegen ajánlott!

A feltűnően rövid idő alatt megoldott feladatlapok kizárást vonnak maguk után, bármilyen más gyanús esetben fenntartjuk a jogot a forduló érvénytelenítésére!

Jó versenyzést kívánunk!



## Üdvözöllek a Nagy Nyelvi Modellek kategória 1. fordulóján! Ezt a fordulót egyértelműen bemelegítésnek szánjuk!

A ChatGPT emberhez hasonló szöveg és kód generálására való képessége számos felhasználáshoz vezetett, ami aggodalmat keltett az ilyen technológiával való lehetséges visszaélésekkel kapcsolatban (pl. álhírek generálása, ártalmas tartalom terjesztése, stb). Ezen aggályok ellenére a ChatGPT-hez hasonló nyelvi modellek fejlődése új lehetőségeket nyitott sokak számára, és forradalmasíthatja a számítógépekkel való interakcióinkat. Akár félünk ettől az új technológiától akár nem, az most már látszik, hogy egyes munkák esetében a mindennapi rutin részeként fogjuk ezeket az eszközöket használni.

A kategória célja, hogy a kitöltő jobban megismerje a nagy nyelvi modelleket, valamint a mögöttük lévő technológiát az itt megtalálható feladatokon keresztül. A verseny 7 fordulója során szó lesz a ChatGPT-ről, Bard-ról, OpenAssistant-ról, BingChat-ről illetve a HuggingFace-en elérhető modellekről is. Jelen forduló teljesítéséhez nem szükséges egyik eszköz használata sem, azonban biztatom a Versenyzőt, hogy már most regisztráljon ezen platformokra és kezdjen el ismerkedni az eszközökkel!

**ChatGPT:** az ingyenes verzióval dolgozunk, elérhető a [chat.openai.com](https://chat.openai.com) címen.

**Bard:** belépés után elérhető a [bard.google.com](https://bard.google.com) címen

**BingChat:** Microsoft Edge böngésző használata során elérhető az oldalsávban

**OpenAssistant:** regisztráció után az [open-assistant.io](https://open-assistant.io) címen elérhető

**HuggingFace:** modellek gyűjteménye, adott feladathoz linkelve lesz a modell\interfész amit használni kell

Jó versenyzést és sok sikert!

## 1. feladat 1 pont

Jeleld meg az alábbiak közül a Nagy Nyelvi Modelleket!

### Válaszok

☐ MidJourney

☐ LLaMA

- ☐ GPT-3.5
- ☐ ChatGPT
- ☐ DALL-E

## 2. feladat 1 pont

A mesterséges intelligencia (MI) szakirodalma rendkívül gazdag és folyamatosan bővülő terület, mely számos tudományágat és iparágat érint. Jelöld meg az alábbiak közül melyek kötődnek szorosan a nagy nyelvi modellekhez:

### Válaszok

- ☐ Gépi tanulás (Machine Learning)
- ☐ Számítógépes látás (Computer Vision)
- ☐ Szakértői rendszer (Expert System)
- ☐ Természetes nyelvfeldolgozás (Natural Language Processing - NLP)
- ☐ Objektumdetektálás (Object Recognition)

## 3. feladat 1 pont

Az alábbiak közül hol dolgoztak a Transzformer architektúra publikálói?

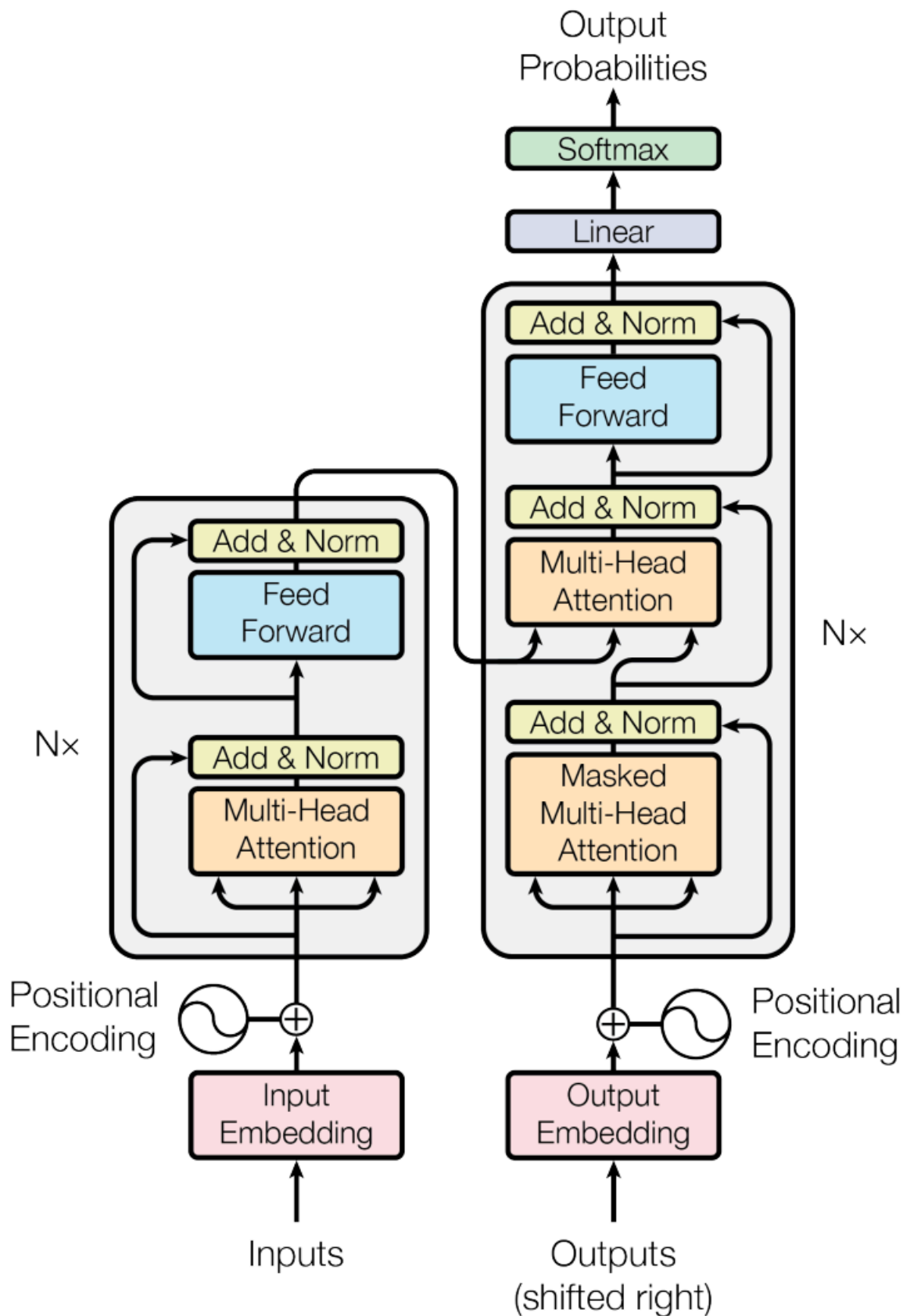
### Válasz

- ☐ Meta (Facebook)
- ☐ Alphabet (Google)
- ☐ Microsoft
- ☐ MIT

## 4. feladat 1 pont

Számos nagy nyelvi modell a Transzformer architektúrát alkalmazza, amely egy speciális formájú neurális hálózat. Mivel több architektúra is megjelent az elmúlt pár hónapban/évben, az egyértelműség kedvéért leszögezzük, hogy az alábbiakban a következőről beszélünk:

### Attention Is All You Need



Az alábbi kijelentések közül mi **NEM** jellemző erre az architektúrára?

### Válasz

- ☐ Segítségével hatékonyabb a szekvenciális adatok feldolgozása

- ☐ Két fő komponensből áll: decoder és encoder
- ☐ Az adat szekvenciálisan folyik az egész hálóban egyik rétegből a másikba
- ☐ A figyelemmechanizmus (attention) lehetővé teszi, hogy a modell súlyozza a bemeneti szekvencia különböző részeit

## 5. feladat 2 pont

Bár alapvetően természetes nyelvi szöveg generálására hozták létre a nyelvi modelleket, tanításuk során forráskódot is láttak. Ennek köszönhetően számos programozási nyelvet is ismernek és képesek forráskódot generálni. Az alábbi példában arra kértük a ChatGPT-t, hogy írjon egy JavaScript függvényt, ami megoldja a Hanoi tornyai problémát. Az alábbi válaszlehetőségek közül jelöld meg azt ami igaz a generált kódra.

Write a function in JavaScript that solves the Hanoi tower problem!

```
function hanoiTower(diskCount, sourcePeg, destinationPeg, auxiliaryPeg) {  
  if (diskCount === 1) {  
    console.log(`Move disk 1 from ${sourcePeg} to ${destinationPeg}`);  
    return;  
  }  
  
  hanoiTower(diskCount - 1, sourcePeg, auxiliaryPeg, destinationPeg);  
  console.log(`Move disk ${diskCount} from ${sourcePeg} to ${destinationPeg}`);  
  hanoiTower(diskCount - 1, auxiliaryPeg, destinationPeg, sourcePeg);  
}  
  
// Example usage:  
const numberOfDisks = 5;  
const sourcePeg = 'A';  
const destinationPeg = 'C';  
const auxiliaryPeg = 'B';  
  
hanoiTower(numberOfDisks, sourcePeg, destinationPeg, auxiliaryPeg);
```

## Válaszok

- ☐ A kód szemantikailag helytelen, nem a hanoi tornyai problémát oldja meg
- ☐ A kód szintaktikailag helyes
- ☐ A kód szemantikailag helyes, a hanoi tornyai problémát oldja meg

☐ A generált kód nem JavaScript

## 6. feladat 3 pont

A ChatGPT tehát egyszerűbb algoritmikus feladatok megoldásában segíthet, mi a helyzet az adatfeldolgozással? Tegyük fel, hogy a következő struktúrájú csv fájlt szeretnénk feldolgozni Python-ban:

```
rank;line_num;MLP-FL;buggy
1;org.mockito.internal.invocation.InvocationMatcher#98;0.5497653268;False
2;org.mockito.internal.invocation.InvocationMatcher#145;0.6487953124;True
3;org.mockito.internal.invocation.ArgumentsProcessor#32;0.5164975863;False
```

Programozóként a lehető legkevesebb munkával szeretnénk megúszni a dolgot, ezért arra kértük a ChatGPT-t oldja meg helyettünk a feladatot. Kettő alternatívát hoztunk létre, amelyek elérhetők [itt](#) ('pandas' megoldás) és [itt](#) ('csv' megoldás). Mit tapasztalsz a megosztott példákkal kapcsolatban, melyik fut le hiba nélkül a fenti példa csv-re és mi okozhatja a működésbeli különbséget annak ellenére, hogy ugyanazt a prompt-ot használtuk? Jelöld meg a következő állítások közül a helyeseket!

### Válaszok

- ☐ Mindkét generált szkript helyes.
- ☐ A '[pandas](#)' hiba nélkül lefut a példára.
- ☐ A '[csv](#)' hiba nélkül lefut a példára.
- ☐ A '[csv](#)' megoldás "Custom Instruction"-t használ.
- ☐ A '[pandas](#)' megoldás "Custom Instruction"-t használ.
- ☐ A használt "Custom Instruction" **hatással van** a kimenetre.
- ☐ A használt "Custom Instruction" **nincs hatással** a kimenetre.

Megoldások beküldése