



 Kiszámoló

**A kiszamolo.hu  
cikkeihez tartozó kommentek  
összefoglalása és témákhoz érkező  
kommentek polarizációjának  
vizsgálata**

# A [kiszamolo.hu](https://kiszamolo.hu) cikkeihez érkező kommentek összefoglalása

Miért hasznos a kommentek összefoglalása?

- Kiegészíti a cikk tartalmát
- Röviden kiemeli az olvasói nézőpontokat, reakciókat
- Más szemszögből mutathatja be a témát  
pl. behozhat vitás pontokat, személyes tapasztalatokat

Ugyanakkor:

- Jelentős részük nem tartalmaz érdemi információt
- Időigényes kiválogatni, átolvasni a relevánsakat

# A [kiszamolo.hu](https://kiszamolo.hu) cikkeihez érkező kommentek összefoglalása

## Felhasznált eszközök:

- Python, Pandas
- BeautifulSoup, requests: webscraping:
- OpenAI API: kommentek összefoglalása, szentiment analízis
- IPyWidgets: widgetek

## Miért kellett a widgetek?

- Kb. 3400 cikk
    - összes komment letöltése és feldolgozása költséges
    - use case?
    - Elmenteni CSV állományba későbbre? → outdated lesz
- interaktív vezérlők (beviteli mezők, gombok, csúszka, selector)

# Témákhoz érkező kommentek szentiment elemzése és a polarizáció vizsgálata



Miért hasznos a kommentek szentiment elemzése és a polarizáció vizsgálata?

- megmutatja, hogy mennyire megosztó egy téma
  - a sok nagyon negatív vagy nagyon pozitív vélemény erős érzelmi reakciót jelez
- Következtetni lehet (predikció)
  - melyik témáknak lesz nagyobb elérés a médiában
  - mennyire kell moderálásra készülni
  - egy marketing kampány mennyire lesz provokatív
- Trendet is lehet figyelni a polarizáció idősoros elemzésével

# Témákhoz érkező kommentek szentiment elemzése és a polarizáció vizsgálata

Hogyan működik? “Interactive Brokers” példán keresztül.

## Adatok gyűjtése

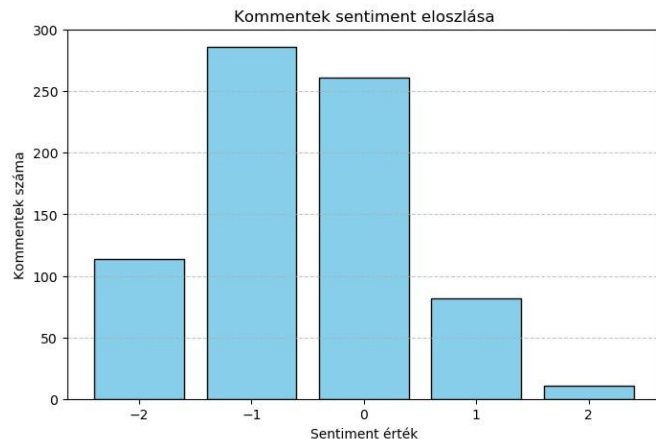
- Meg kell adni a témát (a blog saját keresőjét használom)
- Cikkek (11) + kommentek (754) → dataframe

## Adatok tisztítása

- Dátumok javítása
- Kommentek tisztítása (pl. HTML tagek, stop words, linkek)

## Szentiment analízis

- Kommentek szentiment analízise egyesével OpenAI segítségével
- 5 fokozatú skálán [-2 , +2]
- Kapott értékek ábrázolása charton



# Témákhoz érkező kommentek szentiment elemzése és a polarizáció vizsgálata

## Polarizáció

- Polarizáció számítás
- Idősoros polarizáció számítás
- Idősoros polarizáció ábrázolás

```
[16]: date
      2017-05    0.710161
      2017-06    0.447214
      2019-09    0.755929
      2019-10    0.679636
      2020-03    0.843505
      2020-04    0.000000
      2021-09    0.937164
      2021-10    0.516398
      2023-06    0.927961
      2023-07    0.931862
      2023-08    1.022620
      2023-10    0.844411
      2025-08    0.899252
```

```
[15]: polarization = comments_df["sentiment"].std()
      print("Polarizáció (szórás):", polarization)
```

Polarizáció (szórás): 0.9255526348283378

