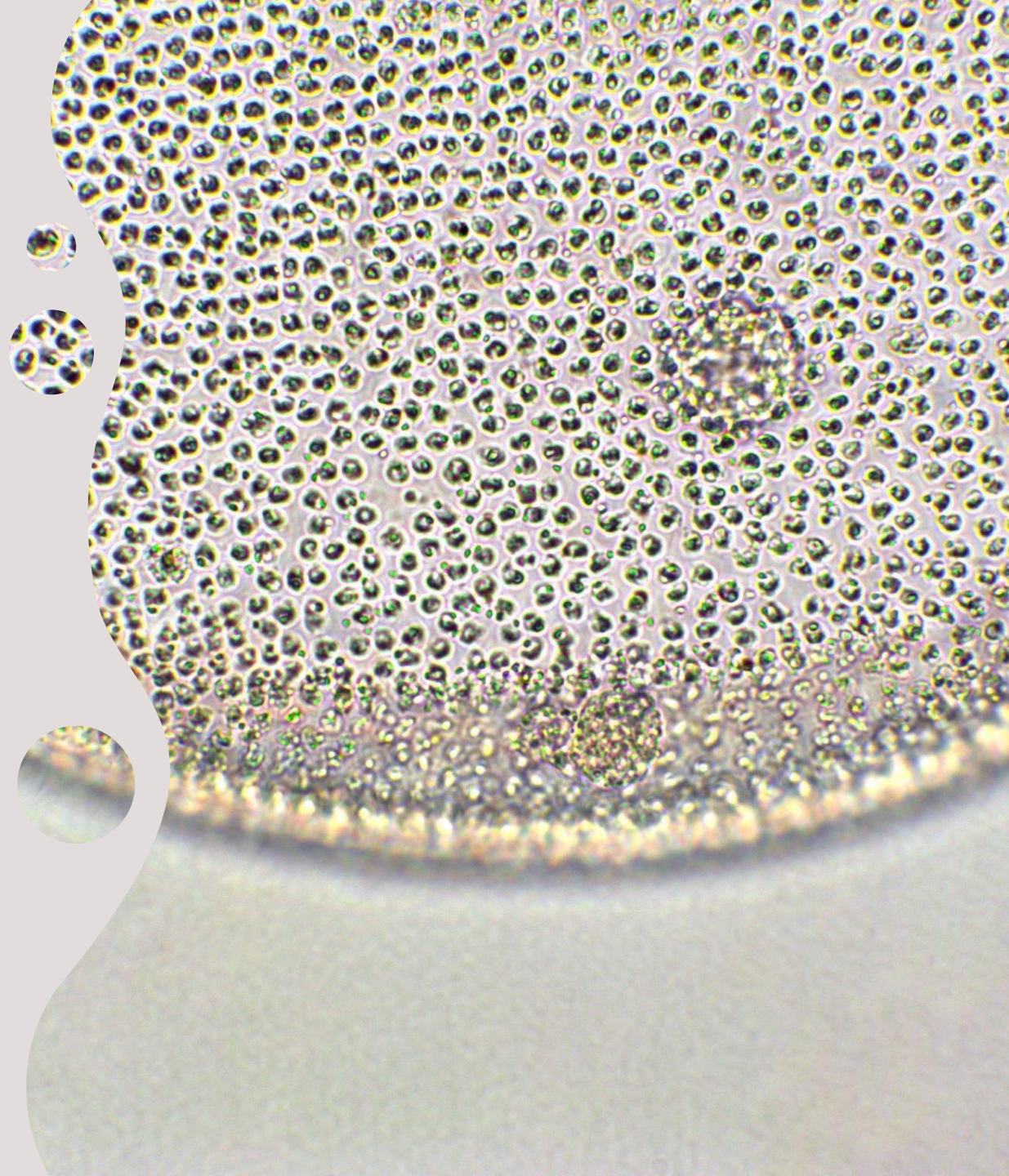


SimQuery

An abstract graphic consisting of several overlapping, flowing, ribbon-like shapes in shades of teal and green. The shapes have a glossy, metallic texture and are set against a light gray gradient background. They flow from the left side of the image towards the right, creating a sense of movement and depth.

by team Query Minds:
Alexandrescu Stefan, Andrei
Cosmin, Baci Florin, Bacter
Cristian, Bara George, Florea
Razvan, Iasinovschi Filip

V-ati gandit vreodata cate
dintre informatile aflate
de noi zilnic sunt false?



Problema
noastra?



Solutia:

În
termeni
simpli,
aplicatia:

Analizează texte și determină
probabilitatea ca acestea să fie
corecte

Oferă un scor de adevăr pe baza unui
set de referință

Ce poate face utilizatorul in aplicație?

- **Cont de utilizator:** Utilizatorii pot crea un cont personalizat, ceea ce le permite să salveze și să acceseze istoricul interacțiunilor.
- **Verificarea emailului:** Toți utilizatorii trebuie să își verifice emailul pentru a asigura securitatea contului.
- **Istoricul conversațiilor:** Aplicația stochează istoricul conversațiilor cu AI-ul, pentru ca utilizatorii să poată urmări evoluția analizelor de-a lungul timpului.

Cum sunt gestionate datele și ce protecție oferă aplicația?

Verificare email pentru securitate: Procesul de autentificare este protejat printr-o verificare a emailului, care ajută la prevenirea accesului neautorizat.

Protecția datelor: Aplicația utilizează criptarea pentru a asigura confidențialitatea datelor utilizatorilor, atât în tranzit (pentru datele transmise) cât și în repaus (pentru datele stocate).

Algorithm & Dataset principal

Pre-fine-tuning

AG News Classification Dataset

Fine-tuning

News Category Dataset

Input

Userul introduce o informație și pe baza ground truth-ului, aplicația va afișa gradul de similaritate

Grad corectitudine

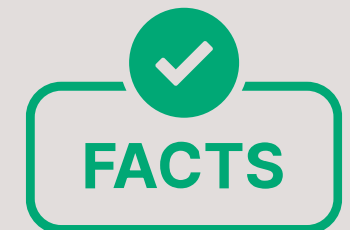
60%

Algorithm

SBERT

Categorii articole

- Politics
- Wellness
- Entertainment
- Travel
- Style & beauty
- Parenting
- Healthy living
- Queer voices
- Food & drink
- Business
- Comedy
- Sports
- Black voices
- Home & living
- Parents



SBERT

Un model avansat de procesare a limbajului natural (NLP).

Preprocesare text

Tokenizarea informației introdusă de user.

Procesare text

BERT procesează informația bidirecțional, generând reprezentări contextuale pentru fiecare token (embeddings)

Normalizarea embeddings-urilor

Vectorii sunt normalizați folosind norma L2 pentru a facilita compararea.

Indexare folosind FAISS

Ground truth-ul sunt adăugat într-un index pentru a permite căutări eficiente.

Calculul similarității

Se calculează produsul scalar între embedding-ul textului introdus și embedding-urile ground truth-ul.

Performance Comparison Between Closely Approached Models

Metric	Fine-tuned GPT-2	Original SBERT	Fine-Tuned SBERT
Pearson Correlation	0.601	0.8796	0.8801
Spearman Correlation	0.480	0.8816	0.8816

	Performance Metrics			
	Accuracy	Precision	Recall	F1-Score
T5 Pre-trained	40%	67%	29%	40%
T5 Fine-tuned	30%	50%	29%	36%
RoBERTa Fine-tuned	63%	58%	86%	69%

Comparison of Performance Metrics Between Previous T5 and RoBERTa Models

Hardware requirements

CPU: Intel(R) Core(TM) i7-8086K CPU @ 4.01 GHz

RAM Memory: 32GB

GPU: Nvidia GeForce GTX 1080 Ti 11GB VRAM, 3584 CUDA Cores

Exemplu

Ground Truth	Prompt	Highest Similarity
<p>TikTok CEO Shou Chew has been invited to and is expected to attend President-elect Donald Trump's inauguration Monday, according to three sources familiar with the matter, two of whom are Trump transition officials ^(0.54). Chew was invited by the president-elect's team and will sit on the dais in front of the Capitol where Trump will be sworn in as president. Chew is expected to sit alongside the CEOs and leaders of other Big Tech platforms, including Tesla and SpaceX CEO Elon Musk, Amazon founder Jeff Bezos and Meta CEO Mark Zuckerberg, according to one of the officials ^(0.48). Google CEO Sundar Pichai is also planning to attend and sit with the other executives, one Trump transition official said ^(0.70). NBC News has reached out to representatives of TikTok and Google for comment. TikTok declined to comment and Google did not immediately respond. The New York Times was the first outlet to report about Chew's inauguration attendance. Chew's attendance comes as the popular TikTok app is preparing to possibly shut down Sunday, the day a U. S. ban is set to go into effect if the Supreme Court doesn't overrule it. The nine justices on the conservative-majority court heard oral arguments last week about whether to implement the ban that was passed by Congress and signed into law by President Joe Biden last year. Remarks and questions by the justices during the arguments suggested it was likely they would uphold the law that would effectively ban the platform.</p>	<p>CEO of Google Sundar Pichai is planning to attend the inauguration.</p>	<p>0.70</p>

Exemplu

```
Server URL: http://localhost:8000/get-similarity
```

```
Enter a prompt to calculate similarity or type 'exit' to quit.
```

```
Prompt: Donald Trump
```

```
Similarity score: 0.43
```

```
Prompt: Donald Trump is coming to his inauguration
```

```
Similarity score: 0.65
```

```
Prompt: The sky is blue
```

```
Similarity score: 0.20
```

```
Prompt:
```

Oportunități de piață

- **Creșterea fenomenului fake news:** Conform studiilor recente, fake news-ul a devenit o problemă majoră în multe domenii, iar soluțiile de combatere a acestuia sunt căutate tot mai mult. Cu peste 4.5 miliarde de utilizatori globali, cererea unor instrumente de încredere pentru verificarea informațiilor de zi cu zi este într-o creștere continuă.
- **Tendențele de consum:** Utilizatorii devin din ce în ce mai conștienți de importanța informării corecte, ceea ce face ca soluțiile precum a noastră să fie extrem de atractive.

Profilul clientului

Companii de media și jurnaliști care doresc să verifice veridicitatea informațiilor înainte de a le publica.



Utilizatori individuali care vor să fie mai siguri de sursele lor de informare.

Valoare propusă

- Oferim un instrument accesibil și rapid pentru verificarea veridicității informațiilor.
- Aplicația poate salva timp prețios utilizatorilor și poate contribui la combaterea dezinformării.

Modelul de business

Abonamente premium: Utilizatorii vor putea plăti pentru funcționalități avansate, cum ar fi analiza în profunzime a surselor și integrarea cu platforme de știri.

Parteneriate cu instituții media: Vom colabora cu agenții de știri pentru a implementa soluția în fluxurile lor de muncă.

Alertă în timp real: Utilizatorii vor putea primi alerte atunci când un articol sau o știre este clasificată ca posibil falsă.

Integrarea cu rețelele sociale: Verificarea rapidă a postărilor de pe unele platforme sociale.



Demo:



Va multumim!