



Generative AI: Large Language Models

Як працюють великі мовні моделі *

Serhii Zabolotnii
zabolotniua@gmail.com

* засновано на матеріалах статті "How Large Language Models Work: From zero to ChatGPT" by Andreas Stöffelbauer ([Medium](#))

1. Важливість розуміння LLMs:

- 1.1. Штучний інтелект та його роль у сучасному світі.
- 1.2. LLMs як ключова складова AI-революції.

2. Еволюція мовної взаємодії з машинами:

- 2.1. Від простих команд до складної бесіди.
- 2.2. LLMs перетворюють спілкування з машинами на природний процес.

3. Практичне застосування LLMs:

- 3.1. Вплив на освіту, бізнес, розваги та інші сфери.
- 3.2. Розширення можливостей професійної та особистої діяльності.

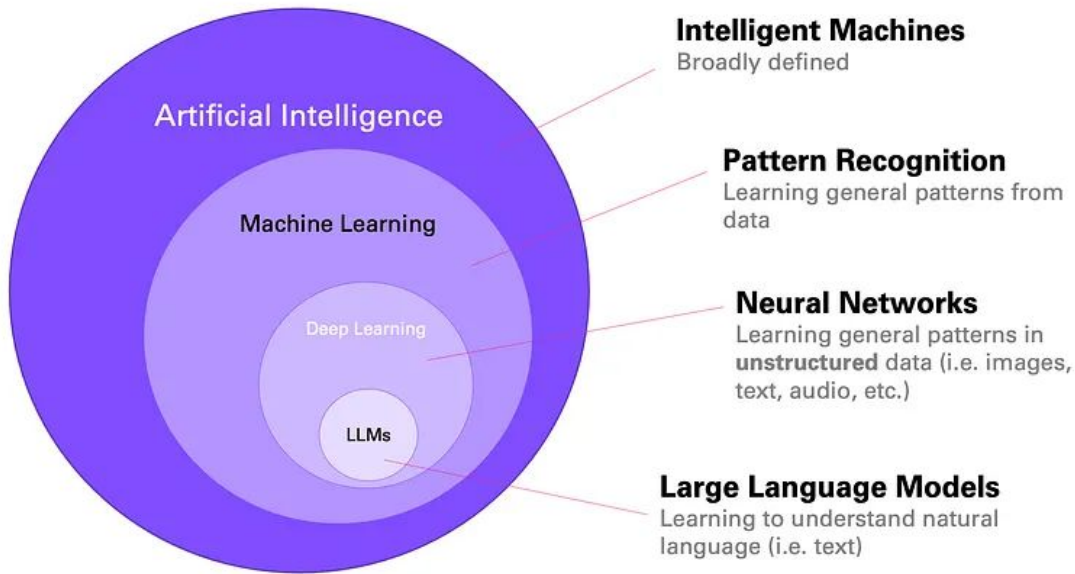
4. Мета доповіді:

- 4.1. Глибоке розуміння механізмів роботи LLMs.
- 4.2. Огляд тренування та використання LLMs у різних контекстах.



Ієрархія Штучного Інтелекту

- Штучний інтелект (AI) — це дуже широкий термін, але загалом він стосується інтелектуальних машин.
- Машинне навчання (ML) — це підгалузь штучного інтелекту, яка спеціально фокусується на розпізнаванні шаблонів у даних. Як ви можете собі уявити, як тільки ви розпізнаєте шаблон, ви можете застосувати його до нових спостережень.
- Глибинне навчання — це сфера в ML, яка зосереджена на неструктурованих даних, які включають текст і зображення. Він покладається на штучні нейронні мережі, метод, який (незначно) натхненний людським мозком.
- Великі мовні моделі (LLM) стосуються конкретно тексту, і це буде в центрі уваги цієї доповіді.



Основи машинного навчання

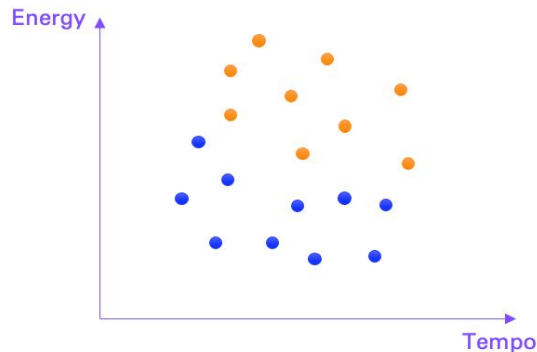
Steps



How it looks in practice

Classification Example: Predicting Music Genre

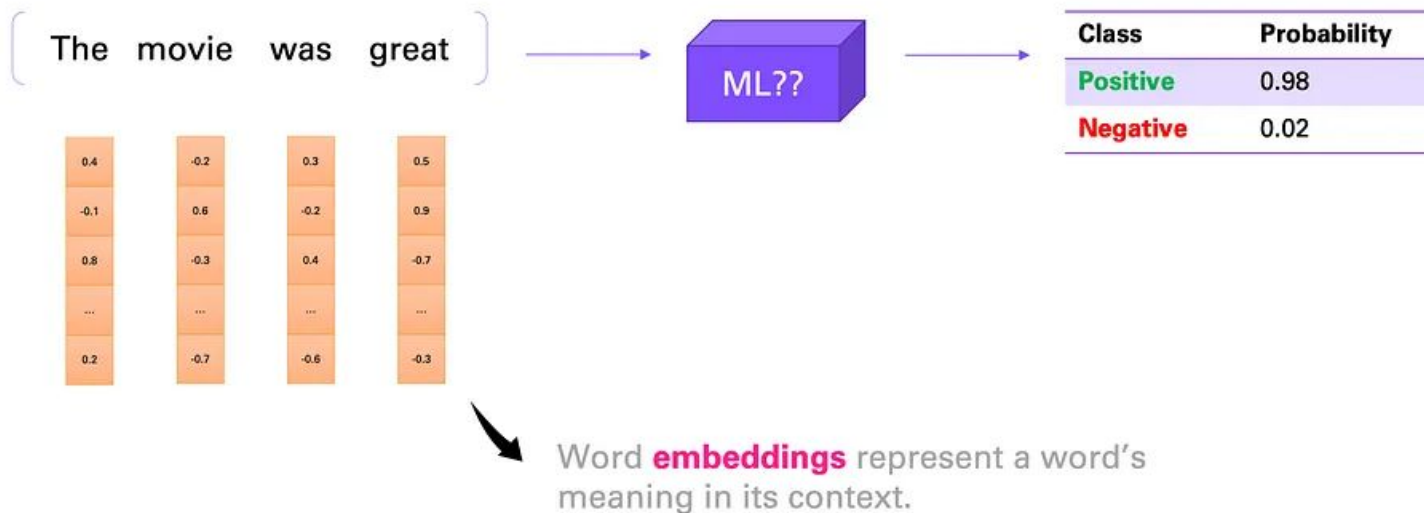
● R&B Songs ● Reggaeton Songs ● New Observation



Приклад Машинного навчання: Задача визначення приналежності (класифікації) пісні до одного із двох музичних жанрів (reggaeton чи R&B) основі аналізу Темпу і Енергійності аудіозапису

Основа обробки мови (NLP): векторне представлення

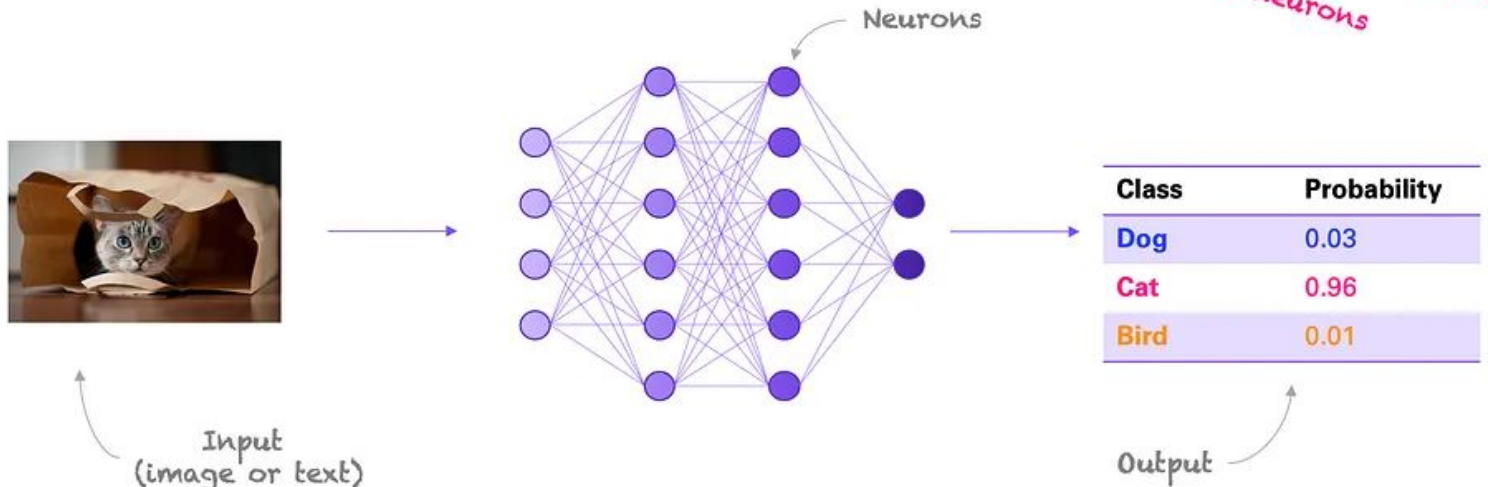
Or what if the input is text?



Ключова технологія NLP - Word Embedding (векторне представлення слів/речень/документів)

Нейронні мережі і глибоке навчання (Deep Learning)

We need something way more powerful... **Neural Networks**



Нейронні мережі — це потужні моделі машинного навчання, які дозволяють моделювати довільно складні зв'язки. Вони є двигуном, який дозволяє вивчати такі складні відносини у великому масштабі.

Мовне моделювання (Великі Мовні Моделі - LLM)

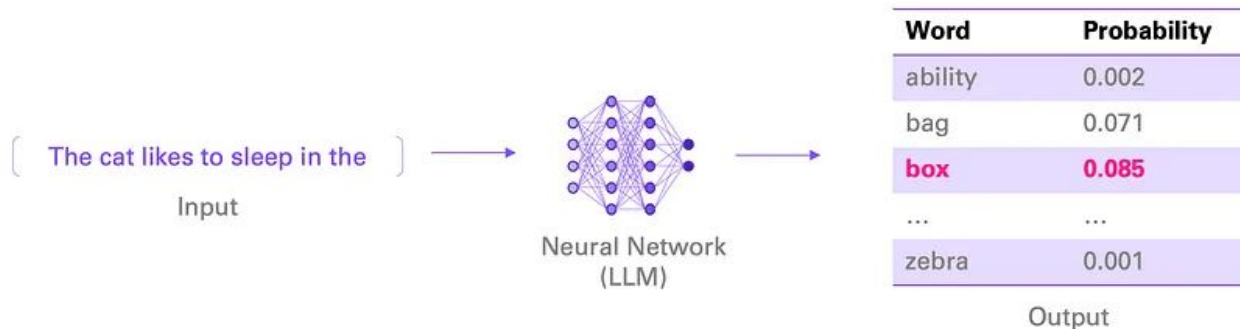
Language modeling

Imagine the following task: Predict the next word in a sequence

[The cat likes to sleep in the _____] → What **word** comes next?

Can we frame this as a ML problem? Yes, it's a classification task.

Now we have (say)
~50,000 classes (i.e.
words)



"Велика Мовна Модель" (LLM, від англ. Large Language Model) - це складний (більше 1 млрд. зв'язків) тип нейронних мережу. Її особливість - у здатності глибоко розуміти мову, що дозволяє ефективно генерувати текст та відповідати на запити.

Навчання мовних моделей

Massive training data



We can create **vast amounts of sequences** for training a language model

● Context ● Next Word ● Ignored

[The cat likes to sleep in the]
[The cat likes to sleep in the]
[The cat likes to sleep in the]
[The cat likes to sleep in the]
[The cat likes to sleep in the]

We do the same with much **longer sequences**. For example:

A language model is a probability distribution over sequences of words. [...] Given any sequence of words, the model predicts the **next** ...

Or also with **code**:

```
def square(number):  
    """Calculates the square of a number."""  
    return number ** 2
```

And as a result - the model becomes **incredibly good at predicting the next word** in any sequence.

"Основне завдання - навчити нейронну мережу (LLM) **передбачати наступне слово**. Для цього використовується велика кількість текстових даних з Інтернету, книг, досліджень тощо. Методика самоконтрольованого навчання дозволяє створювати набори даних без позначення, використовуючи наступне слово як мітку. Робимо це для різноманітних послідовностей, коротких і довгих. Навчання LLM передбачати наступне слово в будь-якій мові і контексті, чи то твіт, вірш чи код, покращує його здатність робити відповідні вибори слова, хоч і не завжди ідеально."

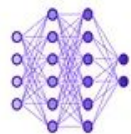
Генеративні Мовні Моделі (Generative AI)

Natural language generation

After training: We can **generate text** by predicting **one word at a time**

A trained language model can

Input



LLM

Word	Probability
speak	0.065
generate	0.072
politics	0.001
...	...
walk	0.003

Output at step 1

Word	Probability
ability	0.002
text	0.084
coherent	0.085
...	...
ideas	0.041

Output at step 2

LLMs are an example of what's called "Generative AI"

Завдяки навчанню LLM передбачати слова, ми можемо генерувати тексти, передаючи послідовності назад у модель для передбачення наступних слів. Таким чином, LLM є прикладом **Generative AI**. Важливо, що LLM не обмежується передбаченням найбільш ймовірного слова, але може вибирати з кількох варіантів, забезпечуючи креативність у відповідях. Така стратегія використовується в ChatGPT, де відповіді не завжди однакові. ChatGPT не називається ChatLLM, оскільки мовне моделювання - лише частина його функціоналу, а GPT відображає його здатність до генерування тексту.

GPT - Генеративний попередньо навчений трансформер.

What does **Generative Pre-trained Transformer (GPT)** mean

Generative

Means "next word prediction."

As just described.

Pre-trained

The LLM is pretrained on massive amounts of text from the internet and other sources.

See next slide.

Transformer

The neural network architecture used (introduced in 2017).

Won't go into more details here.

'G' у GPT означає 'generative', що вказує на здатність моделі генерувати мову. 'P' означає 'попереднє навчання', що відображає етаповий процес навчання моделі. 'T' означає 'трансформер', тип архітектури нейронної мережі, що зосереджує увагу на найважливіших частинах вхідної послідовності, схоже на людське сприйняття.

Фази навчання LLM

Phases of training LLMs (GPT-3 & 4)

1. Pretraining

Massive amounts of data from the internet + books + etc.

Question: What is the problem with that?

Answer: We get a model that can babble on about anything, but it's probably not **aligned** with what we want it to do.

2. Instruction Fine-tuning

Teaching the model to respond to instructions.

Model learns to respond to instructions.

→ Helps alignment

"Alignment" is a hugely important research topic

3. Reinforcement Learning from Human Feedback

Similar purpose to instruction tuning.

Helps produce output that is closer to what humans want or like.

- (1) Попередня підготовка,
- (2) Точне налаштування інструкцій,
- (3) Підкріплення за допомогою зворотного зв'язку людини (RLHF).

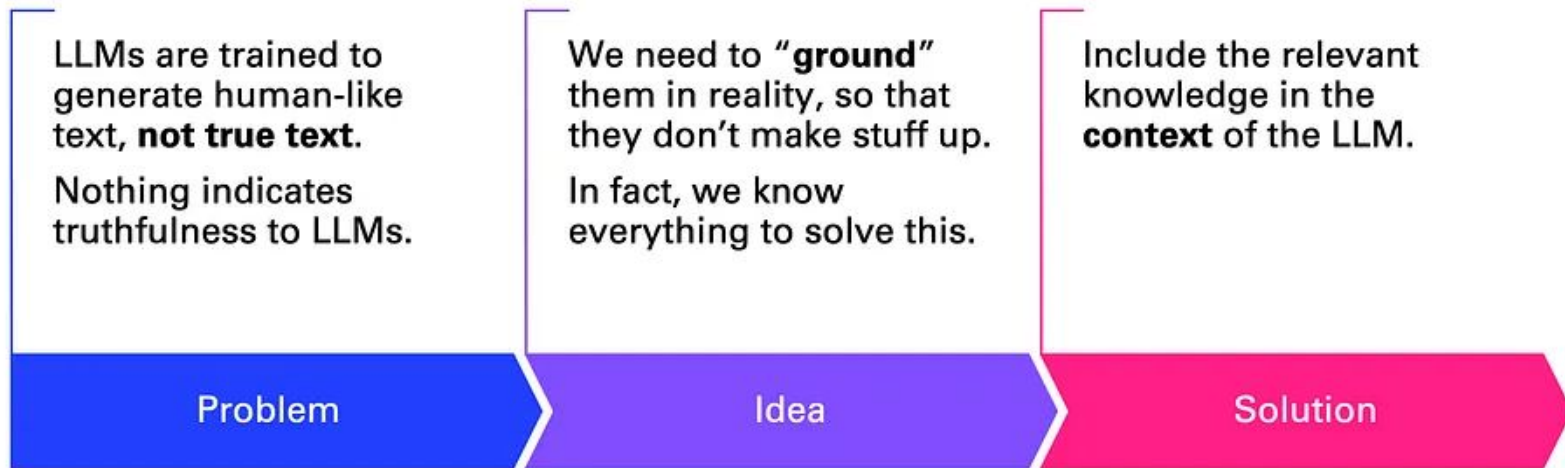
Приклади для перевірки нашого розуміння LLM

Three examples to test our understanding

Ability	Explanation
Why can an LLM perform Text Summarization ?	Ability probably learned during pre-training
Why can an LLM perform Question Answering ?	Knowledge acquired in pre-training, responds nicely due to fine-tuning
Why does a LLM sometimes answer wrong or even make stuff up?	Let's discuss this next...

*Чому LLM може виконувати резюмування довшого фрагмента тексту?
Чому LLM може відповісти на загальновідомі запитання ?
Чому LLM іноді відповідає неправильно і навіть вигадує щось?*

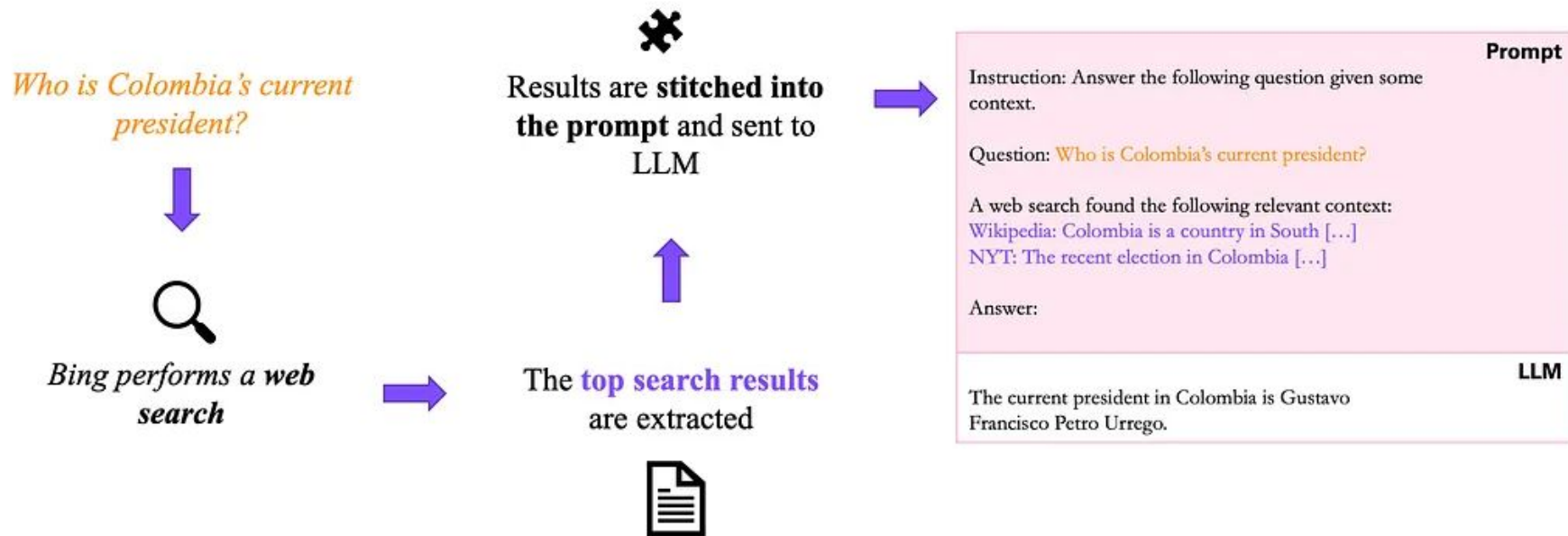
Truthfulness



LLM страждають від галюцинацій, але їх можна пом’якшити, надавши додатковий контекст (додаткові дані).

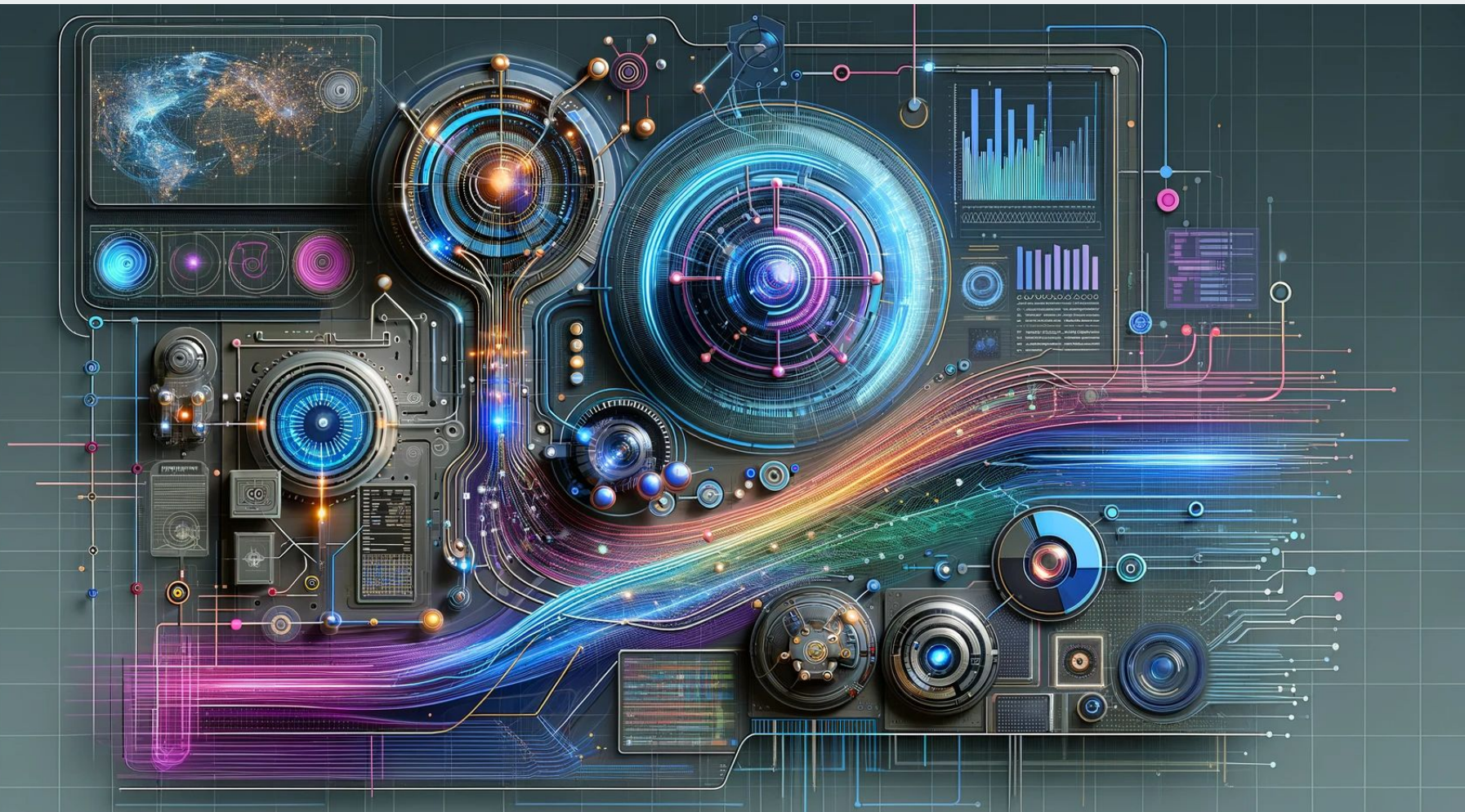
Приклад: робота LLM з елементами пошуку

Example: Bing Chat



При запитанні до LLM, наприклад, про нинішнього президента Колумбії, можливі помилки через галюцинації або застарілі дані. Рішенням є включення актуального контексту, як статті Вікіпедії, щоб збільшити шанси на правильну відповідь. Це 'заземлення' LLM у реальному світі. Bing Chat та інші програми використовують цей підхід, витягуючи контекст з Інтернету і передаючи його разом із запитанням.

Магія Штучного Інтелекту





Thank you