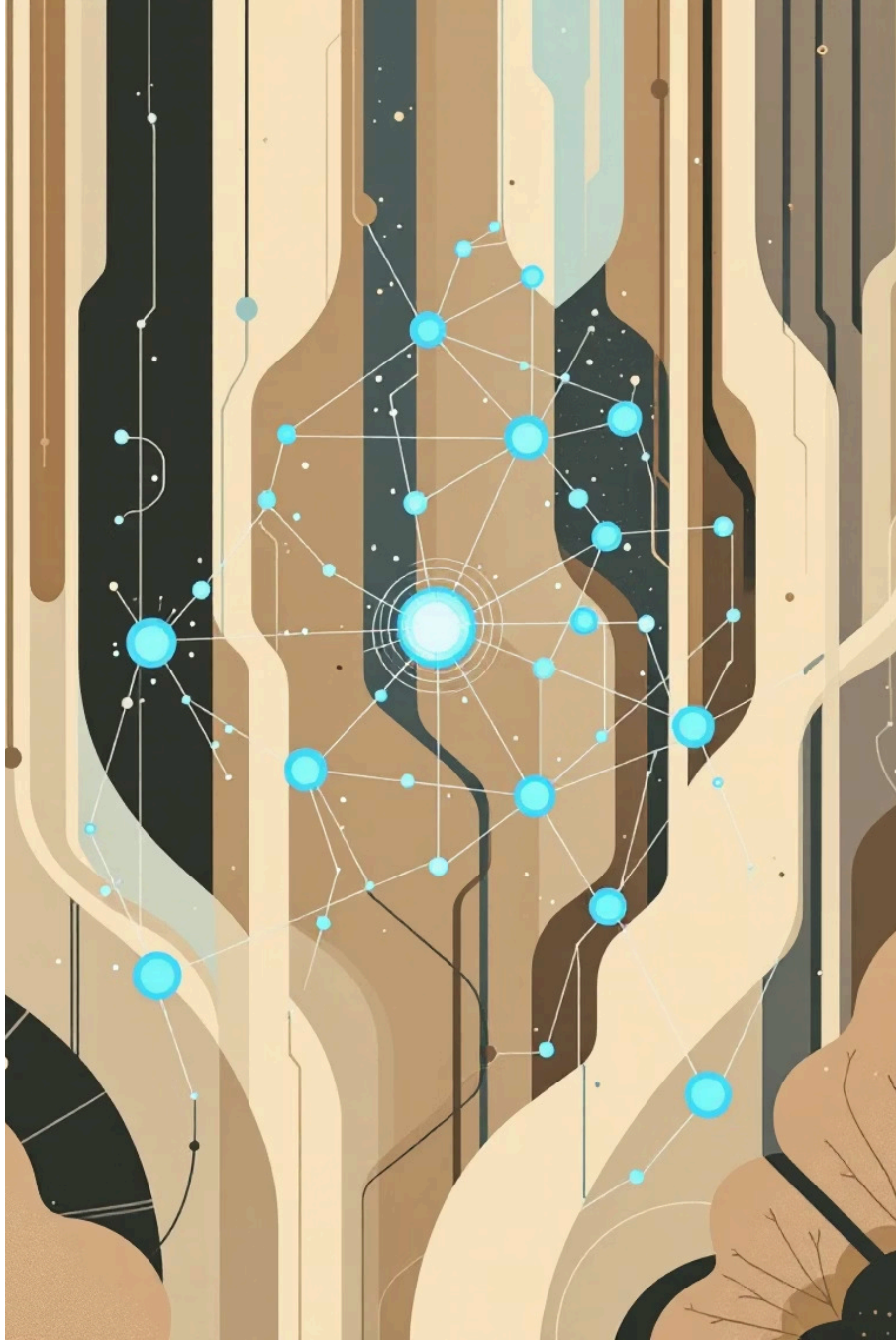
LLM: Duże Modele Językowe

Zaawansowana sztuczna inteligencja wytrenowana na ogromnych zbiorach tekstu do rozumienia i generowania języka naturalnego.

Dr hab. inż. Artur Wyrwa



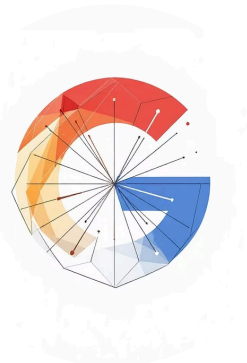
Popularne Modele LLM

GPT-4 / ChatGPT

Najpopularniejszy model konwersacyjny od OpenAI.

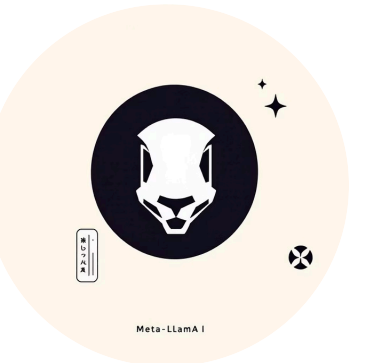
Claude

Zaawansowany asystent AI od Anthropic.



Google Gemini

Multimodalny model od Google.



LLaMA

Open-source model od Meta.

Jak Działa LLM?

Uczenie się wzorców

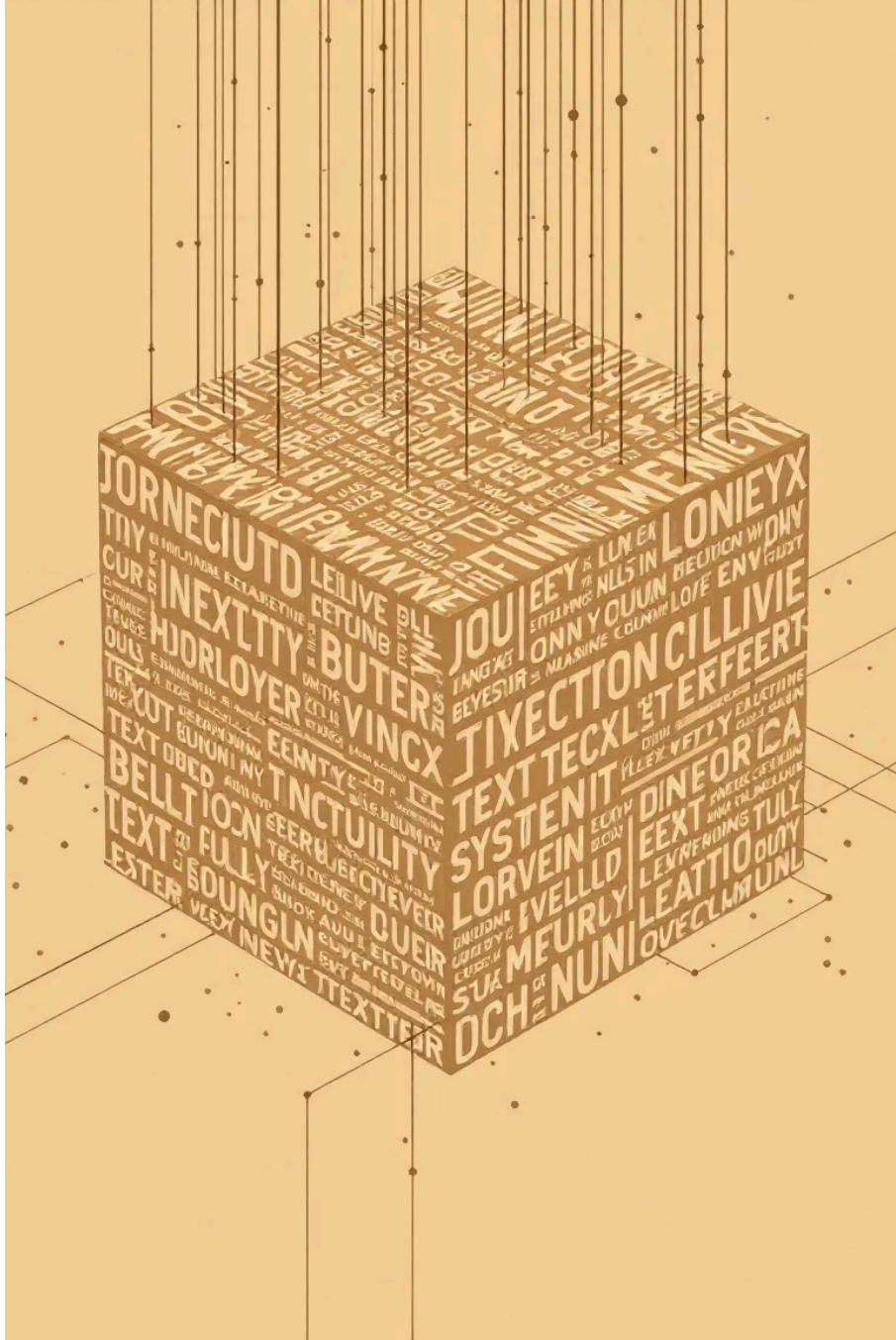
Model analizuje miliardy zdań, aby zrozumieć strukturę języka.

Przewidywanie słów

Na podstawie kontekstu generuje najbardziej prawdopodobne odpowiedzi.

Bez świadomości

Nie ma emocji, ale świetnie przetwarza informacje językowe.



Od zwykłej sieci neuronowej do Large Language Model

Wcześniej poznaliśmy prosty model: **wejście → neurony → wyjście**. Sieć nauczyła się wykrywać wzorce w liczbach i podejmować decyzje na ich podstawie.

LLM to także sieć neuronowa, jednak dużo większa i specjalnie zaprojektowana do pracy z tekstem. Przyjrzyjmy się, jak to działa.

Jak komputer „widzi" tekst?

Problem

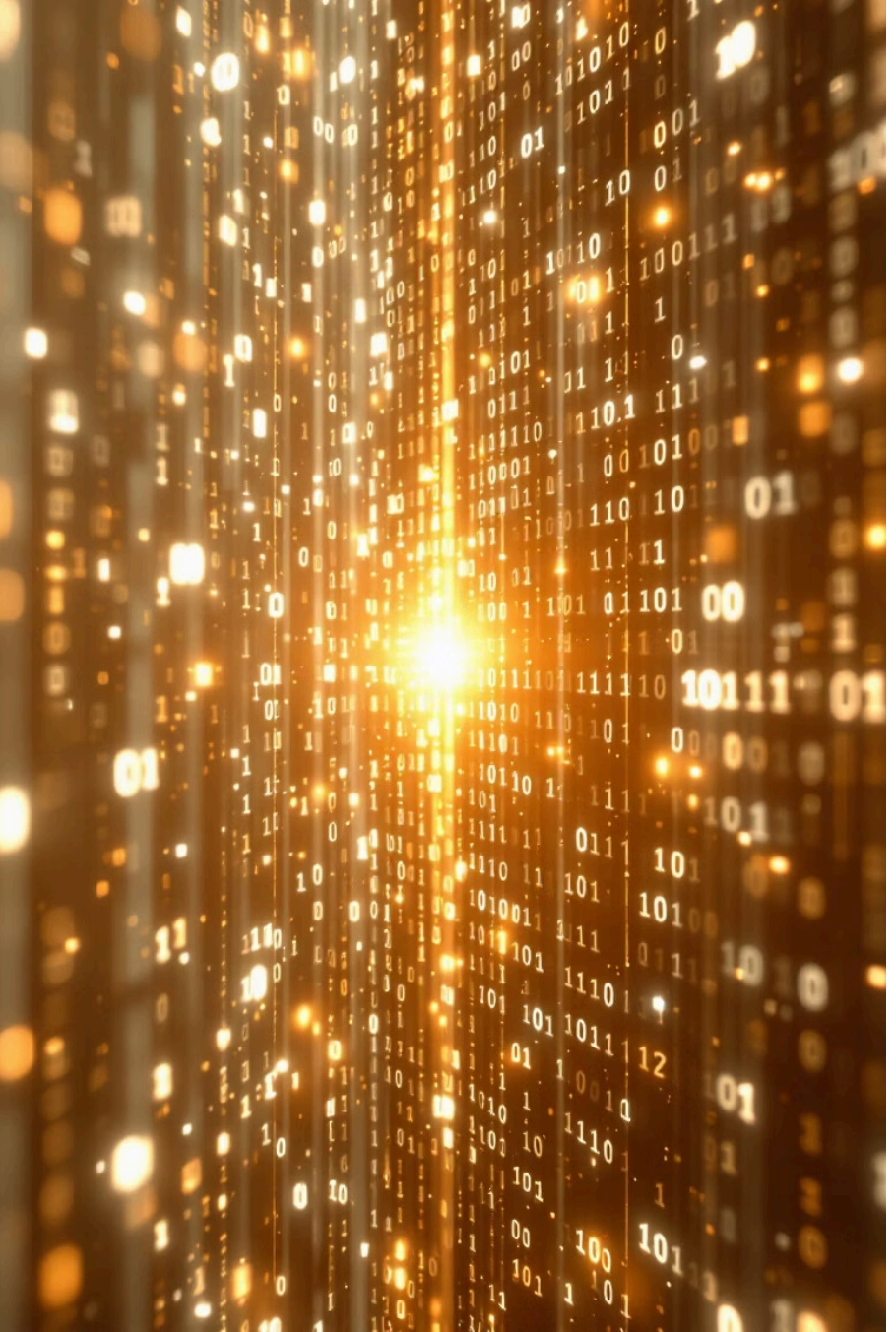
Komputer nie rozumie liter, słów ani zdań w taki sposób jak my. Potrzebuje liczb, aby mógł przetwarzać informacje.

Rozwiązanie

Tekst jest dzielony na małe części zwane **tokenami** — najczęściej są to całe słowa lub ich fragmenty.

Przykład tokenizacji

„Koty lubią mleko" → ["Koty", "lubią", "mleko"]



Tokeny zamieniamy na liczby

Każdy token otrzymuje swoją unikalną **reprezentację numeryczną** — dokładnie jak wejście w zwykłej sieci neuronowej. Te liczby to „adresy" w ogromnym słowniku modelu.

1

„Koty" → 825

Pierwszy token zamieniany na liczbę w słowniku

2

„lubią" → 412

Drugi token otrzymuje inny identyfikator

3

„mleko" → 1901

Każde słowo ma unikalny kod numeryczny

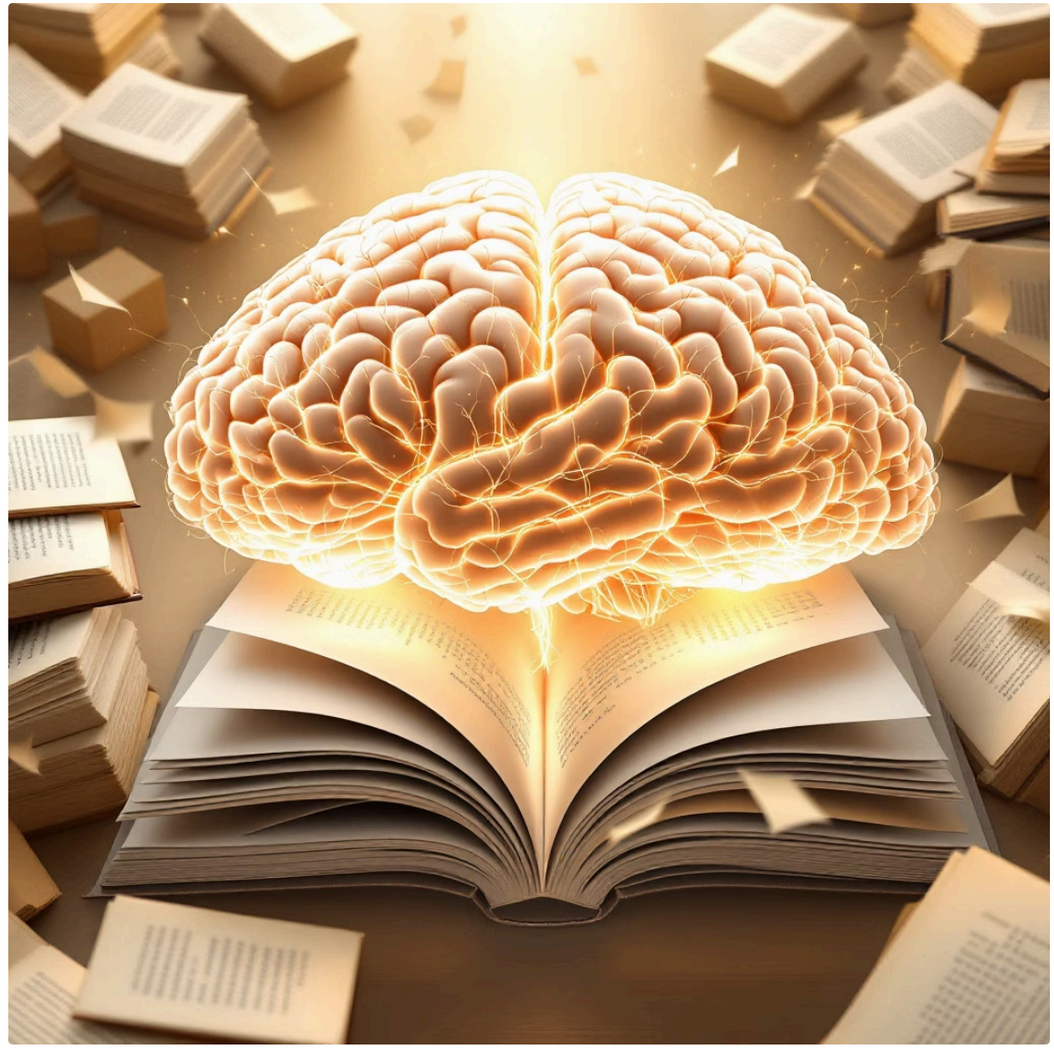
Jak LLM uczy się języka?

Proces uczenia

Sieć otrzymuje miliardy przykładowych zdań z książek, artykułów i stron internetowych.

Analizuje konteksty i uczy się:

- Które słowa często występują razem
- W jakich sytuacjach są używane
- Jakie mają znaczenia w różnych kontekstach



To dokładnie tak, jak **my uczymy się języka** — z ogromnej liczby przykładów i doświadczeń, tylko że model robi to w znacznie większej skali.

Główne zadanie: przewidzieć kolejne słowo

01

Analiza kontekstu

Model czyta dotychczasowy tekst i analizuje wzorce

02

Przewidywanie

„Jakie słowo powinno pojawić się następne?"

03

Generowanie

Wybiera najbardziej prawdopodobne słowo

04

Powtarzanie

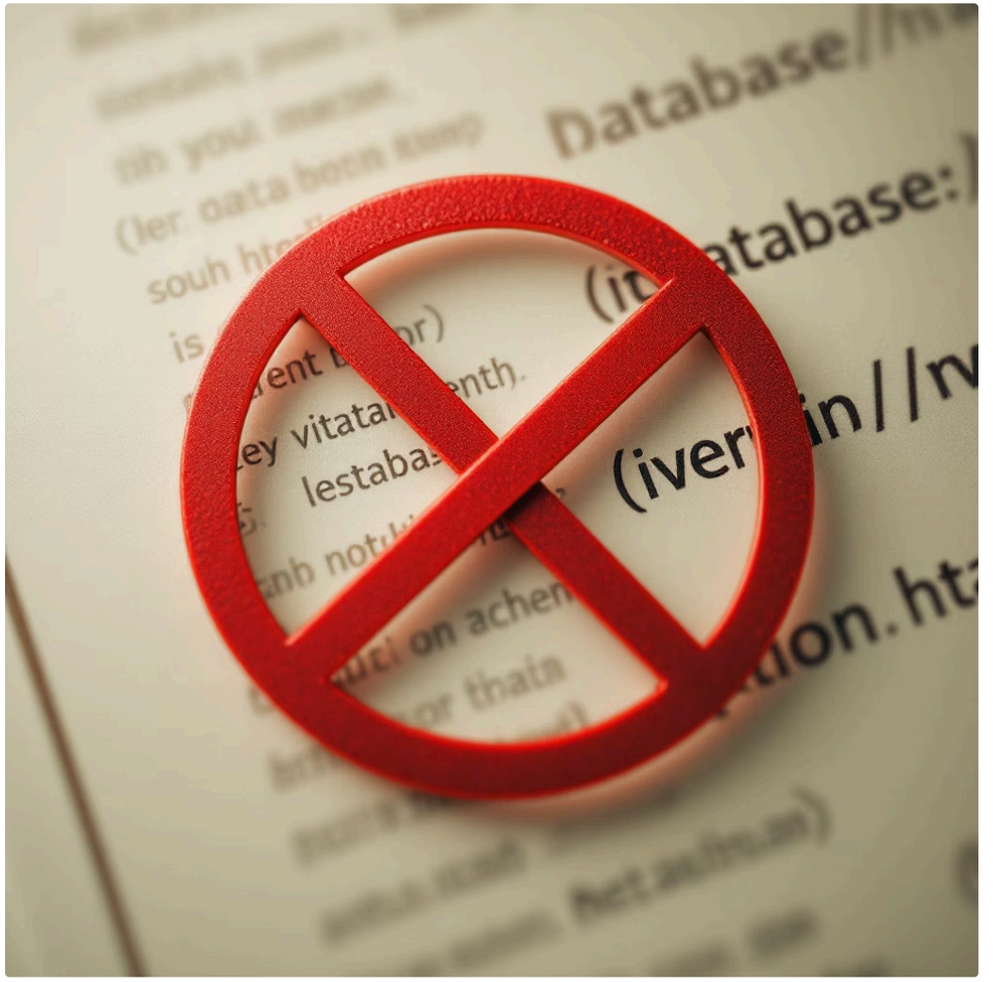
Proces kontynuowany jest słowo po słowie, aż powstanie pełny tekst

LLM działa jak **bardzo rozbudowana wersja sieci neuronowej**, którą już poznaliście — tylko na znacznie większą skalę.

Wzorce statystyczne, nie baza danych

Czym NIE jest LLM

Model **nie przeszukuje** gotowej bazy zdań ani nie cytuje tekstów z pamięci jak wyszukiwarka.



Jak działa naprawdę

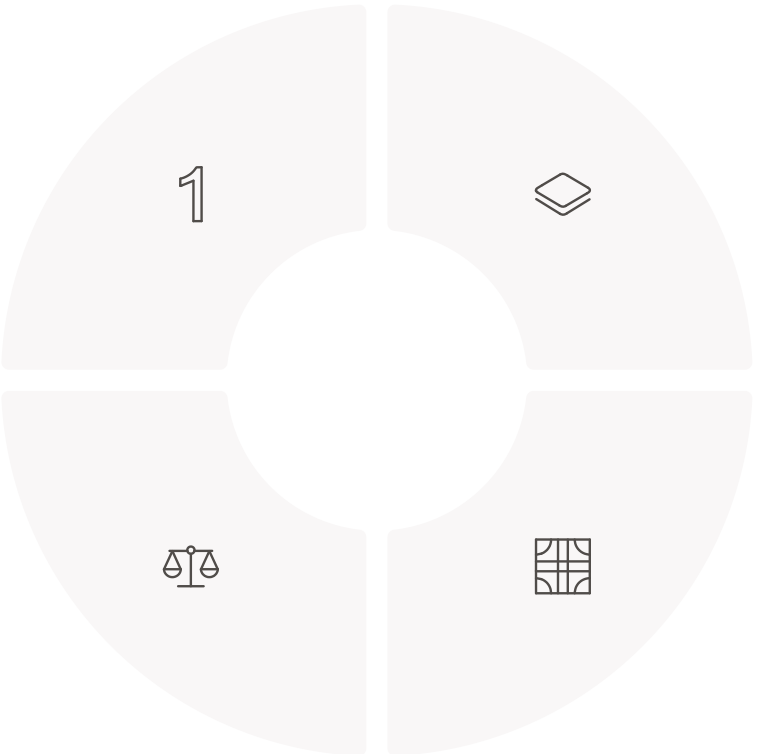
Trzyma w **wagach sieci wzorce statystyczne** wyuczone z danych:

- Co zwykle występuje po czym
- Jakie słowa są ze sobą związane
- Jaka struktura zdań jest poprawna

Podobieństwo do prostej sieci ReLU

Wejście: liczby

W obu przypadkach dane wejściowe to reprezentacje numeryczne



Warstwy neuronów

Informacje przechodzą przez kolejne warstwy przetwarzania

Skala

Główna różnica: LLM ma setki miliardów parametrów zamiast kilku tysięcy

Wykrywanie wzorców

Neurony uczą się rozpoznawać coraz bardziej złożone zależności

Podsumowanie: od tekstu do inteligentnych odpowiedzi

Tekst

Zaczynamy od zwykłego tekstu napisanego przez człowieka

Tokeny

Dzielimy go na małe części — tokeny

Liczby

Każdy token zamieniamy na reprezentację numeryczną

Sieć neuronowa

Przetwarzamy przez warstwy neuronów

Przewidywanie słów

Model generuje kolejne słowa na podstawie wzorców

> **LLM to ta sama idea, którą już poznaliście** — tylko w gigantycznej skali. Dzięki temu może generować teksty, tłumaczyć języki, podpowiadać rozwiązania i odpowiadać na pytania.