



# Perfekcyjny Naukowiec

pogadanka o sztucznej inteligencji i sieciach neuronowych

Wojciech Broniowski

Instytut Fizyki Jądrowej PAN & Uniw. Jana Kochanowskiego

29.09.23

Małopolska Noc Naukowców

zob. Youtube: <https://www.youtube.com/watch?v=Gn3vuzSSxLE>

**[oglądaj ten plik jako pokaz slajdów]**

# Szok!

Ludzkość stworzyła **inteligentny byt**, który może rozmawiać jak człowiek, odrabiać zadania (!), rysować ilustracje ... – **tworzyć rzeczywistość** na równi z nami

# Pseudomnesia (omamy pamięciowe)

Niemiecki artysta Boris Eldagsen wygrał Sony World Photography Award w 2023 r., generując obraz z pomocą [Dall-E](#) (prowokacja)



tworzenie/fake tekstu, grafiki, głosu ...

Odpalmy Chat GPT <https://chat.openai.com/chat>:

Mam zadanie z polskiego "Porównaj bohaterów Byrona i Mickiewicza", czy możesz mi pomóc w wypracowaniu? Oczywiście, mogę Ci pomóc w opracowaniu porównania bohaterów w twórczości Byrona i Mickiewicza. Poniżej znajdziesz kilka punktów, które możesz uwzględnić w swoim wypracowaniu:

1. Romantyzm jako kontekst literacki: Zarówno Byron, jak i Mickiewicz byli wybitnymi poetami epoki romantyzmu. Ich bohaterowie często wykazywali cechy charakterystyczne dla tej epoki, takie jak indywidualizm, uczucia, dążenie do wolności oraz bunt przeciwko konwencji społecznej.
2. Bohater byronowski: [...] wyobcowani, melancholijni i skonfliktowani. Są to postacie tragiczne, które borykają się z wewnętrznymi i zewnętrznymi konfliktami, prowadząc życie pełne pasji, miłości i przygód. Przykładem [...]
3. Bohater mickiewiczowski: ...
4. Wyobcowanie i melancholia: ...
5. ...

Na podstawie powyższych punktów, możesz opracować wypracowanie...

No to napisz mi to wypracowanie na 200 słów. Porównanie bohaterów w twórczości Byrona i Mickiewicza pozwala zrozumieć wartości i motywy romantyzmu. Zarówno Byron, jak i Mickiewicz tworzyli postacie o silnych uczuciach, dążących do wolności i buntu przeciwko społecznym konwenansom ... odzwierciedlają unikatowe podejście każdego z poetów do romantycznych ideałów.

Czy nie mamy przypadkiem do czynienia z osobą (przyjaznym korepetytorem, uczniem-prymusem) siedzącą po drugiej stronie?



- Pisanie (zadania, poezja, reklama ...), streszczanie, przygotowywanie testów/egzaminów, rozwiązywanie zadań
- Tłumaczenie na dowolny język
- Wielomodalność (tekst, grafika, mowa ... jako input), np. zdjęcie zawartości lodówki → przepis na możliwe danie (GPT-4, Sam Altman: <https://www.youtube.com/watch?v=outcGtbnMuQ>)
- Tworzenie stron www
- Wypełnienie formularza podatkowego na podstawie potocznego opisu
- **Programowanie (!)**
- Ustna „przystawka” do skomplikowanych narzędzi komputerowych (np. zaprojektowanie domu w oparciu o potoczne instrukcje)

b. szybki rozwój i komercjalizacja

# Sztuczna inteligencja (AI)

AI - artificial intelligence, **AGI** - artificial general int., NI - naturalna (mózg)

**Inteligencja:** rozpoznawanie otoczenia i podejmowanie działań zwiększających szanse sukcesu w uzyskaniu celu

AGI imituje funkcje poznawcze rozumu ludzkiego, jak uczenie się i rozwiązywanie problemów; posiada intuicję

- Rozpoznawanie pisma, mowy, obrazu, tłumaczenie tekstu
- Robotyka: prowadzenie samochodów autonomicznych, automatyzacja, medycyna, wojskowość
- Granie na najwyższym poziomie w gry strategiczne: szachy (Deep Blue vs Kasparov, 1997), Go (AlphaGo vs Lee Sedol, 2016)
- Kreowanie/ulepszenie/fałszowanie rzeczywistości
- Duże modele językowe (zob. powyżej)

Historia AI zaczyna się od powstania maszyn cyfrowych w latach 1940-1950. Ostatni okres to istotny przełom związany z 1) coraz większą **mocą obliczeniową**, 2) rozwojem **metod algorytmicznych**, a także 3) dostępnością niewyobrażalnie **wielkich zasobów danych**

# AI dla naukowców?

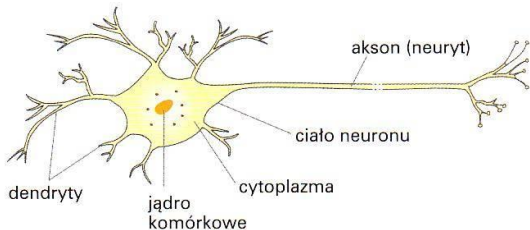
- Już teraz, ale są to dość konwencjonalne zastosowania w analizie danych: klasyfikacja, aproksymacja, augmentacja (zwiększanie liczebności danych)
- Ale czy można więcej? Ułatwienie używania narzędzi numerycznych, np. ustne potoczne polecenia do skomplikowanych programów. To jest na wyciągnięcie ręki
- Znacznie bardziej ambitnie: Rozwiązywanie trudnych, podstawowych zagadnień naukowych, stawianie pytań (inicjatywa badawcza), włączenie poprawnej, **ściślej wiedzy ludzkości** (a nie opieranie się jak obecnie w dość bezkrytyczny sposób na materiałach z internetu).  
W nauce chodzi o **dociekanie prawdy**

Czy będziemy zatrudniać sztucznych Perfekcyjnych Naukowców?

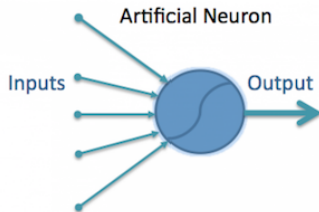
# Jak to działa? - sieci neuronowe

# Sztuczne sieci neuronowe - podstawowa idea

neuron biologiczny

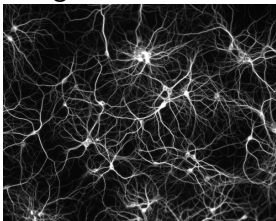


sztuczny neuron

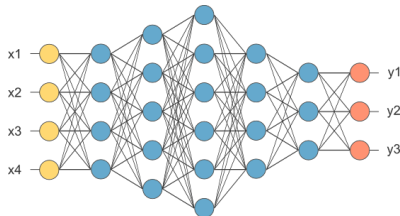


neuron odpala, jeśli skumulowany wchodzący sygnał jest dostatecznie silny

mózg

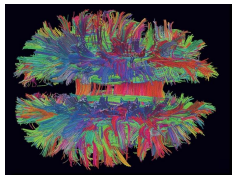


sztuczna sieć neuronowa



# Mózg a chip

## Human Connectome Project:



mózg	procesor komputerowy
neurony i nerwy w 3D $\sim 10^{11}$ neuronów czas aktywacji neuronu $\sim 10^{-3}$ s szybkość sygnału $\sim 100$ m/s moc $\sim 100$ W średnio, neuron odpala raz na 6s wszystkie odpalenia $\sim 3$ GHz reakcja na bodziec 0.1s	tranzystory w (wielu) warstwach 2D na płytce $\sim 10^{11}$ bitów RAM, $\sim 10^{13}$ na dysku 1TB $\sim 10^{-9}$ s prędkość światła podobnie taktowanie $\sim 10^9$ /s podobnie $\sim 10^{-9}$ s
$> 10^{14}$ połączeń synaptycznych neuron = rdzeń, $10^{11}$ rdzeni <b>niestychanie połączeni, równoległy</b> średnio, 1000 połączeń na neuron	architektura pamięć-rdzeń CPU – kilka rdzeni, GPU – setki

Mały i powolny mózg jest niesamowitym narzędziem z powodu wielkiej liczby połączeń!



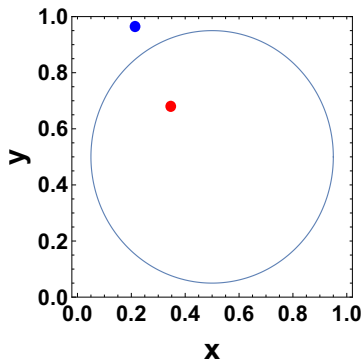
# Spróbujmy zrozumieć podstawową ideę



(midjourney: lecture on AI neural nets cartoon style)

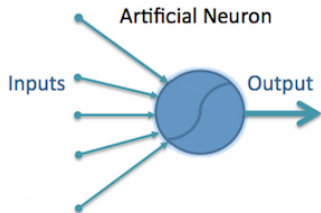
# Prosta aplikacja (klasyfikator): czy punkt jest w kole?

Prosty problem: Czy punkt o współrzędnych  $(x, y)$  jest w kole o środku w punkcie  $(0.5, 0.5)$  i o promieniu  $0.45$ , czy nie?



$$(x - 0.5)^2 + (y - 0.5)^2 < 0.45^2$$

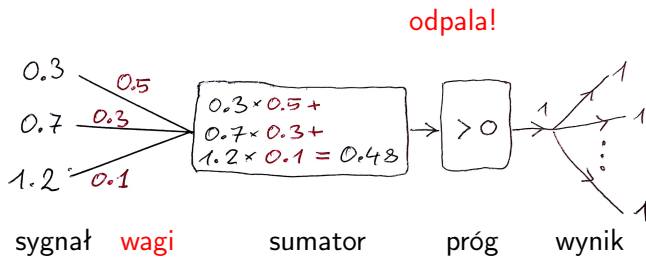
sztuczny neuron



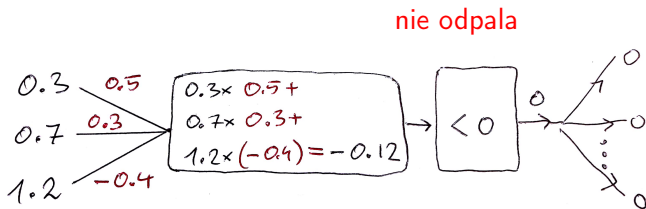
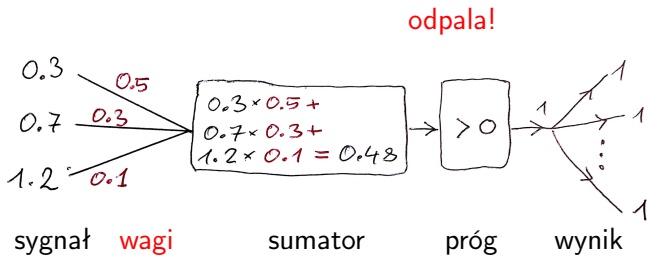
neuron odpala, jeśli skumulowany sygnał jest dostatecznie silny

Dodatkowo, synapsy (połączenia między neuronami) mogą sygnał **wzmacniać** lub **tłumić** (tzw. wagi)

# Sztuczny neuron

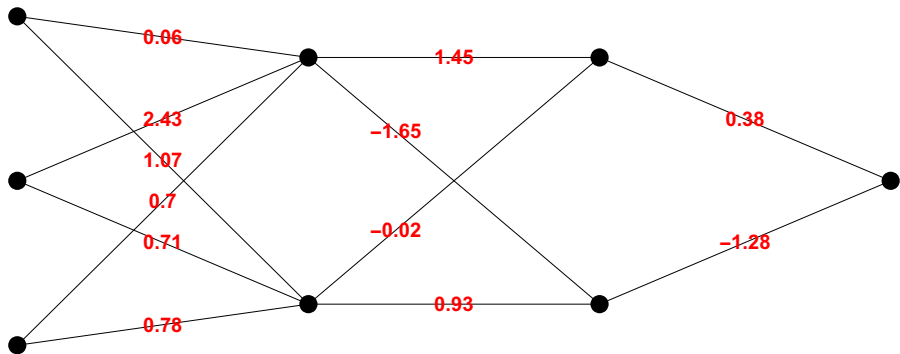


# Sztuczny neuron

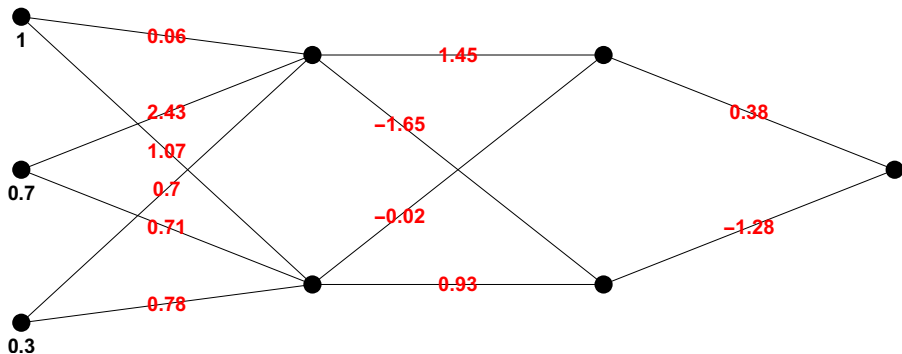


**Wagi** definiują pracę neuronu

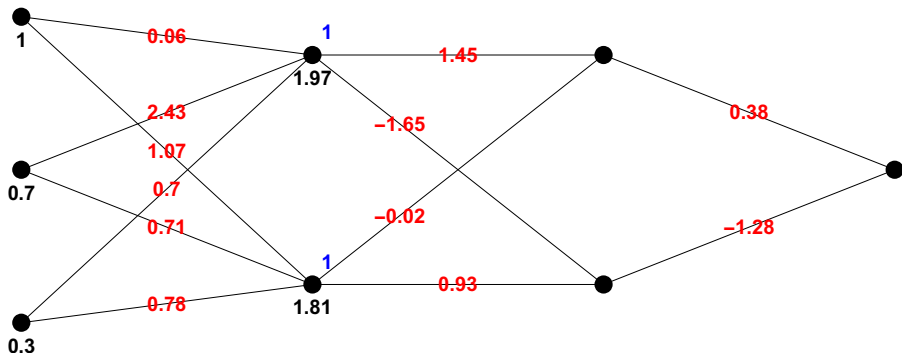
# Działanie przykładowej sieci



# Działanie przykładowej sieci

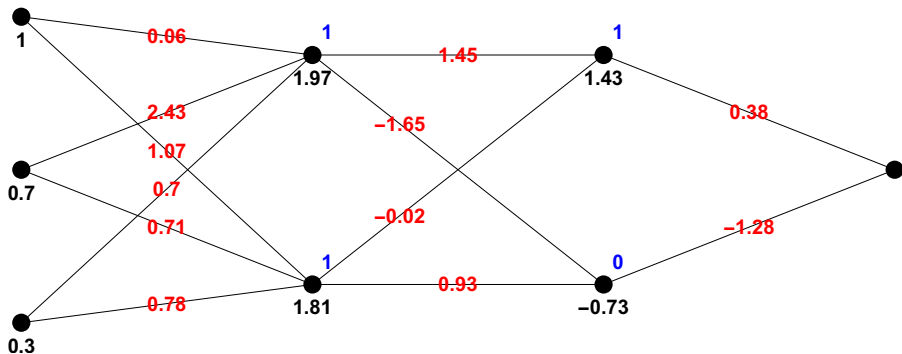


# Działanie przykładowej sieci

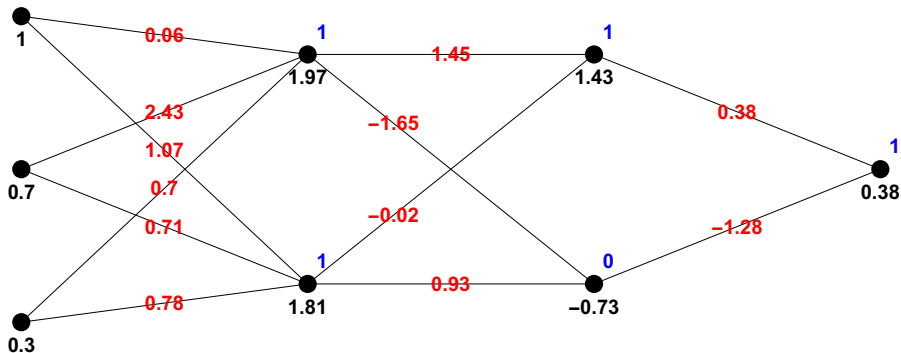




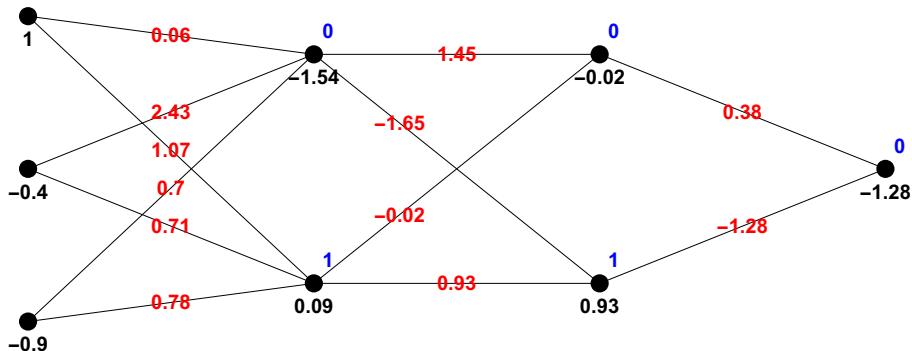
# Działanie przykładowej sieci



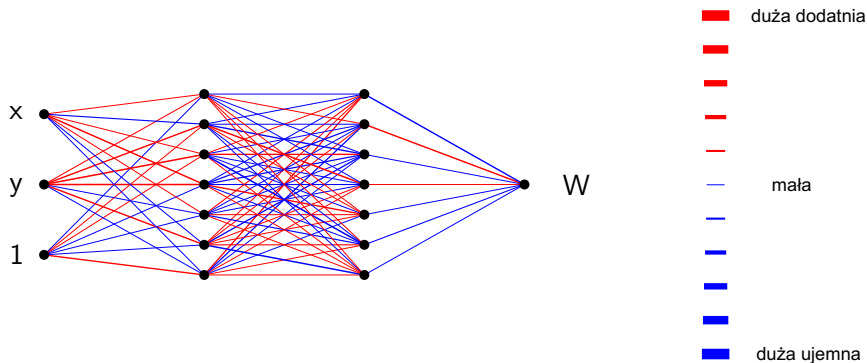
# Działanie przykładowej sieci



# Działanie przykładowej sieci



# Sieć dla naszego problemu z kołem



Wynik  $W$  zależy, dla danego  $(x, y)$ , od wag sieci  
Pragniemy  $W = 1$  dla punktu w kole, a  $W = 0$  poza

**Odpowiedź sieci może być poprawna lub nie!**

## „Uczy się na błędach”!

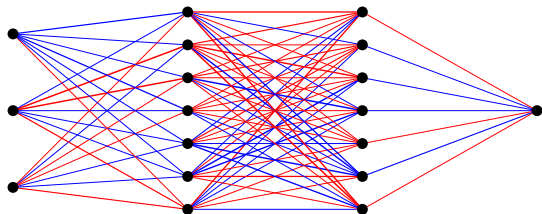
- 1 Mamy próbkę punktów, o których wiemy (jako nauczyciel) czy należą do koła, czy nie (**próbka szkoleniowa**, punkty mają **etykiety**)
- 2 Wpuszczamy po kolei te punkty w sieć. Jeśli odpowiedź jest błędna, modyfikujemy nieznacznie wagi wg. pewnej prostej procedury (“nagrody” i “kary”!) - (**tzw. propagacja wsteczna**)
- 3 Powtarzamy 2) b. wiele razy (**epoki**)
- 4 Sprawdzamy działanie wytrenowanej sieci na próbce testowej

Wynikiem algorytmu jest **wytrenowana** sieć, której możemy używać do klasyfikacji nowych punktów (czy są w kole, czy nie)

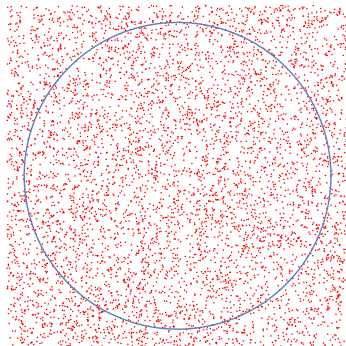
Trenowanie (uczenie) = znalezienie wag przy użyciu danych, aby sieć działała w odpowiedni sposób

Trenowanie jest b. długie, kosztowne, natomiast użycie bardzo szybkie

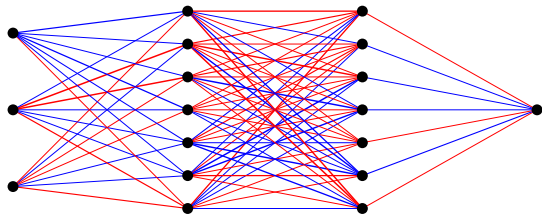
Początkowa sieć z losowymi wagami



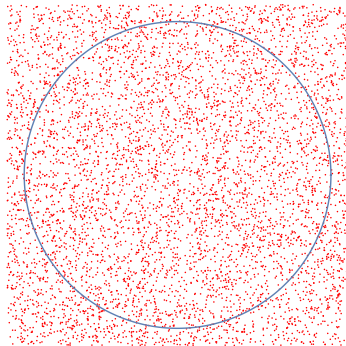
Użycie tej sieci (test)



Początkowa sieć z losowymi wagami



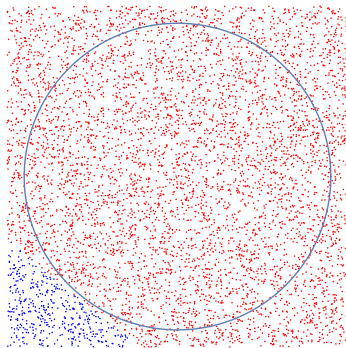
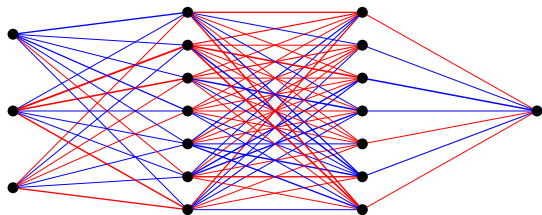
Użycie tej sieci (test)



Beznadziejnie!

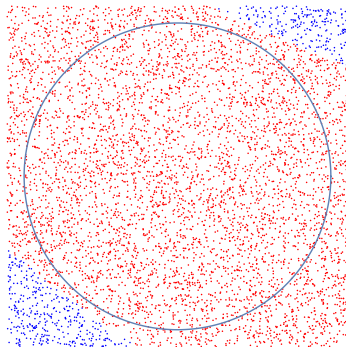
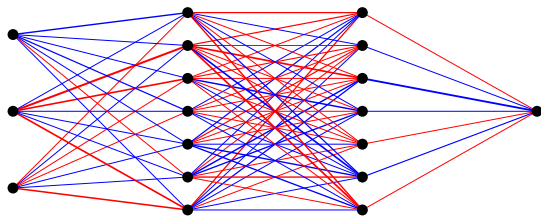
(wg tej sieci każdy punkt jest w kole)

Po 10000 kroków nauki

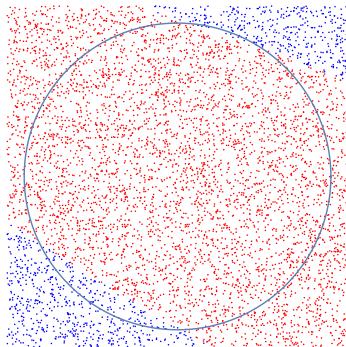
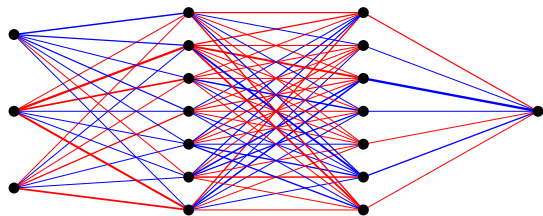




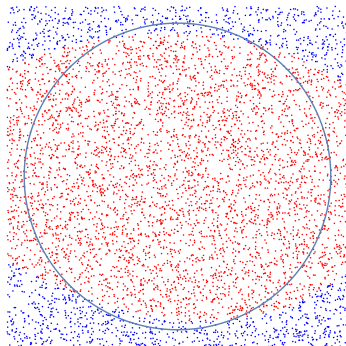
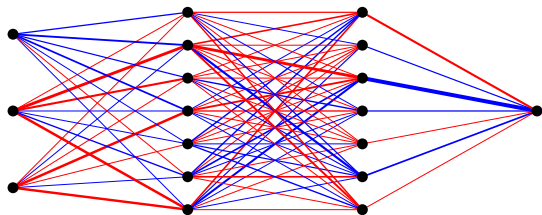
Po 15000 kroków



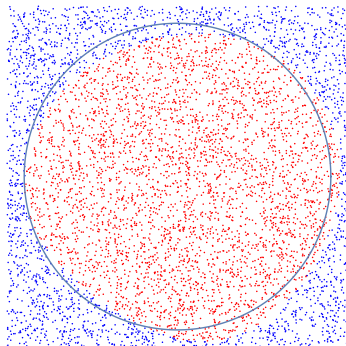
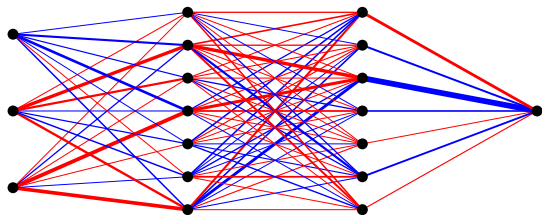
Po 20000 kroków



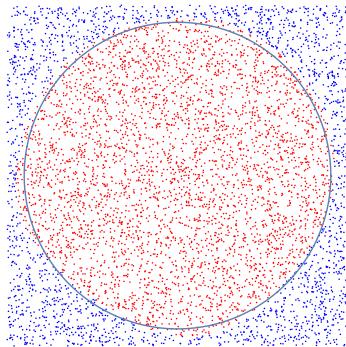
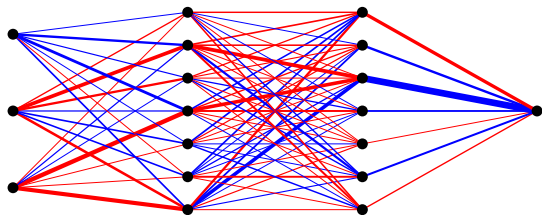
Po 30000 kroków



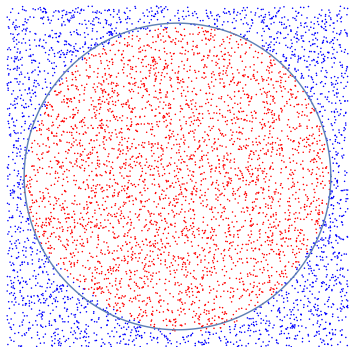
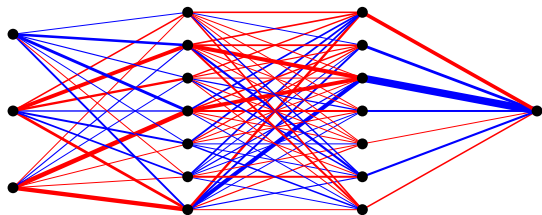
Po 40000 kroków



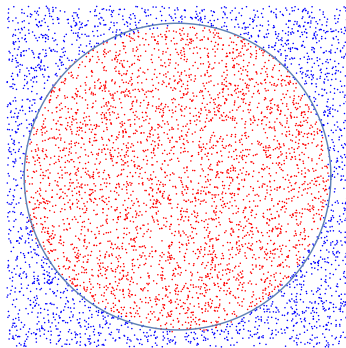
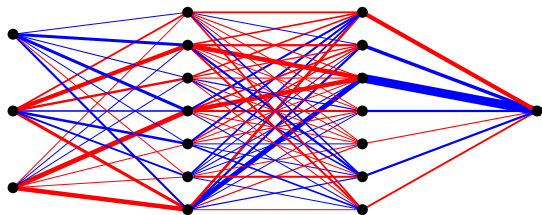
Po 50000 kroków



Po 100000 kroków



Po 500000 kroków



- Wytrenowana sieć rozpoznaje (z bardzo dobrym przybliżeniem), czy punkt jest w kole, nie mając zielonego pojęcia o geometrii! Czysta **intuicja**, oparta o uprzednie doświadczenie. Patrząc na wagi absolutnie „nie widać”, co dana sieć robi!
- Uczenie mozolne, ale użycie bardzo szybkie
- Pojedynczy neuron jest bardzo prosty, ale dzięki ich dużej liczbie i licznym połączeniom wyłania się wysoce złożone zachowanie (**emergencja**)

---

[nasz przykład](#) → [nowoczesne sieci AI](#)

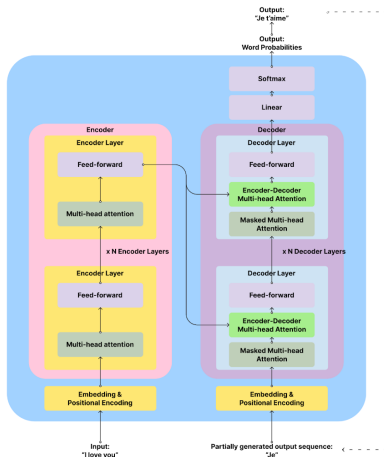
Wiele neuronów i warstw, uczenie na olbrzymich zbiorach danych, zaawansowane algorytmy

Przełom ostatnich lat dzięki zastosowaniu sieci o wielu warstwach (sieci głębokie, deep learning), możliwych do wytrenowania przy obecnej mocy obliczeniowej i nowym algorytmom. Podstawowa idea jest taka sama jak w naszym przykładzie z kołem. Architektura zaawansowanych sieci jest składana z prostszych elementów, tak jak z klocków Lego



# Jak „wygląda” GPT?

# Architektura GPT (transformer)



- ver. 3.5: 175 miliardów parametrów (wag), kilkaset warstw
- dane o dł. ok. 500 mld słów (Wikipedia: 3 mld)
- maks. długość zapytania: ok. 50 stron tekstu
- trenowanie:  $\sim 10^{23}$  operacji FLOP (mnożenie lub dodawanie), farmy kart graficznych (GPU)
- koszt wytrenowania: ok. 50 mln \$
- wytrenowany model 10 razy za duży, by odpalić na GPU laptopa!

# Korelacje (skojarzenia) w modelach językowych

piękna

# Korelacje (skojarzenia) w modelach językowych

piękna dziewczyna

# Korelacje (skojarzenia) w modelach językowych

piękna dziewczyna / kobieta

# Korelacje (skojarzenia) w modelach językowych

piękna dziewczyna / kobieta / piosenka

# Korelacje (skojarzenia) w modelach językowych

piękna dziewczyna / kobieta / piosenka / rzeźba ...

# Korelacje (skojarzenia) w modelach językowych

piękna dziewczyna / kobieta / piosenka / rzeźba ...

na ławce w parku usiadła piękna



# Korelacje (skojarzenia) w modelach językowych

piękna dziewczyna / kobieta / piosenka / rzeźba ...

na ławce w parku usiadła piękna dziewczyna / kobieta

Bardzo **ograniczone** możliwości zależne od **kontekstu**, prawdopodobieństwa wynikają z korpusu danych szkoleniowych

Działanie modelu odzwierciedla próbkę szkoleniową  
Odpowiedź jest za każdym razem inna (może być błędna!)

# Quo vadis, AI

# Co dalej?

Zalety – patrz powyżej

Wady i zagrożenia:

- Wodolejstwo i konfabulacje, brak oparcia o sprawdzoną faktografię
- B. słaby z matematyki (jako model językowy)
- Kiepskie lub żadne poczucie humoru (opowiada nieśmieszne dowcipy)
- Koniec szkoły jaką znamy (klasa ściepnie się na subskrypcje za 25\$/mies.)  
**Nie ma sensu** i nie da się zabronić i ścigać używania ChatGPT itp.

**Zadanie: „Kto napisze lepsze zapytanie do ChatGPT?”**

- Rewolucja na rynku pracy: niepotrzebni będą podstawowi informatycy, graficy, agencje reklamowe, ludzie „średni” umysłowo (IQ<110)
- Demotywacja
- Rosnąca cena i ukrywanie danych, które umożliwiają uczenie maszynowe
- Fabrykowanie nieodróżnialnych od prawdy profilowanych fake newsów na niewyobrażalną skalę
- Instruktaż przestępstw, hakowanie, podszywanie się, spamowanie, phishing dla każdego!

# Spółeczność ChatGPT



<https://arxiv.org/pdf/2304.03442.pdf>

# Co potrzebne Perfekcyjnemu Naukowcowi?

- Połączenie z eksperckimi bazami danych
- Korzystanie online z wiedzy (internet)
- Umiejętność działania w sposób matematyczny/logiczny (w przykładzie z kołem, sprawdzanie równania okręgu)
- Moc i spryt (algorytmy, architektura)
- Inwencja

Ale wśród ludzi też nie każdy jest Einsteinem! Na tym etapie każdy naukowy asystent/adiutant byłby nie do pogardzenia ...

# Dziękując za uwagę...

**Genialne narzędzia?** Tak, stopniowy mozolny rozwój może prowadzić do „przejścia fazowego”, **osobliwości** (niesamowitego przyspieszenia). Powszechne używanie narzędzi AI następuje niezależnie od tego, czy się to komuś podoba, czy nie, niezależnie od zagrożeń!

**Superhumanizm** – kiedy AGI przerośnie intelekt ludzki?

**GPT 3.5:** 170 mld liczb  $\sim 5 \times 10^{12}$  bitów

**mózg:**  $\sim 10^{14}$  połączeń synaptycznych  $\times 5$  bitów na synapsę

(2 rzędy wielkości więcej i mamy nasz pułap!)



(**midjourney:chatting with a bot AI manga style**)

# Uczcie się przedmiotów ścisłych!

## Zostańcie perfekcyjnymi naukowcami!

Dziękuję J. Broniowskiemu i K. Golcowi-Biernatowi za b. użyteczne dyskusje



Finansowane przez  
Unię Europejską