

Sztuczna Inteligencja Głupota Naturalna

...czyli słów kilka o tym, dlaczego nie mamy samych fajnych rzeczy.

dr, inż. Filip Turoboś
Zakład Kwantowych Systemów Informatyki
Instytut Informatyki Teoretycznej i Stosowanej
Polska Akademia Nauk.
10 marca 2026

- Czym jest inteligencja (i czym jest jej brak);

Plan ramowy

- Czym jest inteligencja (i czym jest jej brak);
- Sieci neuronowe, sieci rekurencyjne i duże modele językowe;

Plan ramowy

- Czym jest inteligencja (i czym jest jej brak);
- Sieci neuronowe, sieci rekurencyjne i duże modele językowe;
- Halucynacje – skąd się biorą i dlaczego są problemem;

Plan ramowy

- Czym jest inteligencja (i czym jest jej brak);
- Sieci neuronowe, sieci rekurencyjne i duże modele językowe;
- Halucynacje – skąd się biorą i dlaczego są problemem;
- Inne problemy z AI i (chwilowy brak) remedium;

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...
 - ...część badaczy uwzględniła nawet rośliny.

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...
 - ...część badaczy uwzględniła nawet rośliny.

Problemy z definiowaniem inteligencji

W uproszczeniu, inteligencja to:

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...
 - ...część badaczy uwzględniła nawet rośliny.

Problemy z definiowaniem inteligencji

W uproszczeniu, inteligencja to:

- zdolność do pozyskiwania informacji przez percepcję lub wnioskowanie;

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...
 - ...część badaczy uwzględniła nawet rośliny.

Problemy z definiowaniem inteligencji

W uproszczeniu, inteligencja to:

- zdolność do pozyskiwania informacji przez percepcję lub wnioskowanie;
- przechowywanie informacji w formie wiedzy;

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...
 - ...część badaczy uwzględniła nawet rośliny.

Problemy z definiowaniem inteligencji

W uproszczeniu, inteligencja to:

- zdolność do pozyskiwania informacji przez percepcję lub wnioskowanie;
- przechowywanie informacji w formie wiedzy;
- umiejętność zaaplikowania zgromadzonej wiedzy w stosownym kontekście.

Czym jest inteligencja?

Brak dobrej definicji inteligencji.

- Jeszcze do nie-tak-dawna określaliśmy inteligencję jako cechę przynależną tylko ludziom.
 - Potem dodaliśmy do tego zbioru zwierzęta...
 - ...a potem również maszyny...
 - ...część badaczy uwzględniła nawet rośliny.

Problemy z definiowaniem inteligencji

W uproszczeniu, inteligencja to:

- zdolność do pozyskiwania informacji przez percepcję lub wnioskowanie;
- przechowywanie informacji w formie wiedzy;
- umiejętność zaaplikowania zgromadzonej wiedzy w stosownym kontekście.

Inteligencja jest zjawiskiem wielopłaszczyznowym, a do tego niezwykle trudnym w kwantyfikacji.

Jak zdefiniować głupotę?

Głupota to po prostu brak przejawów inteligencji.

Jak zdefiniować głupotę?

Głupota to po prostu brak przejawów inteligencji.

- “Najbardziej przemocowym elementem społeczeństwa jest ignorancja/głupota.” – Emma Goldman, filozofka i feministka z przełomu XIX i XXw.

Jak zdefiniować głupotę?

Głupota to po prostu brak przejawów inteligencji.

- “Najbardziej przemocowym elementem społeczeństwa jest ignorancja/głupota.” – Emma Goldman, filozofka i feministka z przełomu XIX i XXw.
- “[Głupota to] powolność myślenia w mowie i w czynie.” – Teofrast z Eresos, grecki uczony, III w. p.n.e.

Jak zdefiniować głupotę?

Głupota to po prostu brak przejawów inteligencji.

- “Najbardziej przemocowym elementem społeczeństwa jest ignorancja/głupota.” – Emma Goldman, filozofka i feministka z przełomu XIX i XXw.
- “[Głupota to] powolność myślenia w mowie i w czynie.” – Teofrast z Eresos, grecki uczony, III w. p.n.e.
- “Kluczowe czynniki dla głupich zachowań [...] obejmują brak myślenia przyczynowo-skutkowego [...] prowadząc do głupoty w zderzeniu z nowymi sytuacjami.” – A. Dayal ”Artificial Intelligence and Artificial Stupidity: The Inseparables”

Jak zdefiniować głupotę?

Głupota to po prostu brak przejawów inteligencji.

- “Najbardziej przemocowym elementem społeczeństwa jest ignorancja/głupota.” – Emma Goldman, filozofka i feministka z przełomu XIX i XXw.
- “[Głupota to] powolność myślenia w mowie i w czynie.” – Teofrast z Eresos, grecki uczony, III w. p.n.e.
- “Kluczowe czynniki dla głupich zachowań [...] obejmują brak myślenia przyczynowo-skutkowego [...] prowadząc do głupoty w zderzeniu z nowymi sytuacjami.” – A. Dayal ”Artificial Intelligence and Artificial Stupidity: The Inseparables”
- “[Głupotę...] charakteryzują braki w kluczowej dla inteligencji intencjonalności: zdolności do pojmowania znaczenia.”– P. Musso, ”Artificial Stupidity”

Jak zdefiniować głupotę?

Głupota to po prostu brak przejawów inteligencji.

- “Najbardziej przemocowym elementem społeczeństwa jest ignorancja/głupota.” – Emma Goldman, filozofka i feministka z przełomu XIX i XXw.
- “[Głupota to] powolność myślenia w mowie i w czynie.” – Teofrast z Eresos, grecki uczony, III w. p.n.e.
- “Kluczowe czynniki dla głupich zachowań [...] obejmują brak myślenia przyczynowo-skutkowego [...] prowadząc do głupoty w zderzeniu z nowymi sytuacjami.” – A. Dayal ”Artificial Intelligence and Artificial Stupidity: The Inseparables”
- “[Głupotę...] charakteryzują braki w kluczowej dla inteligencji intencjonalności: zdolności do pojmowania znaczenia.”– P. Musso, ”Artificial Stupidity”

W kontekście generatywnej sztucznej inteligencji możemy tutaj nieco bardziej skupić się na mowie – o tym za chwilę.

Wokół sztucznej inteligencji

Sztuczna Inteligencja (SI/AI) to dziedzina nauki skupiająca się na tworzeniu systemów obliczeniowych umożliwiających maszynom:

Wokół sztucznej inteligencji

Sztuczna Inteligencja (SI/AI) to dziedzina nauki skupiająca się na tworzeniu systemów obliczeniowych umożliwiających maszynom:

- percepcję środowiska;

Wokół sztucznej inteligencji

Sztuczna Inteligencja (SI/AI) to dziedzina nauki skupiająca się na tworzeniu systemów obliczeniowych umożliwiających maszynom:

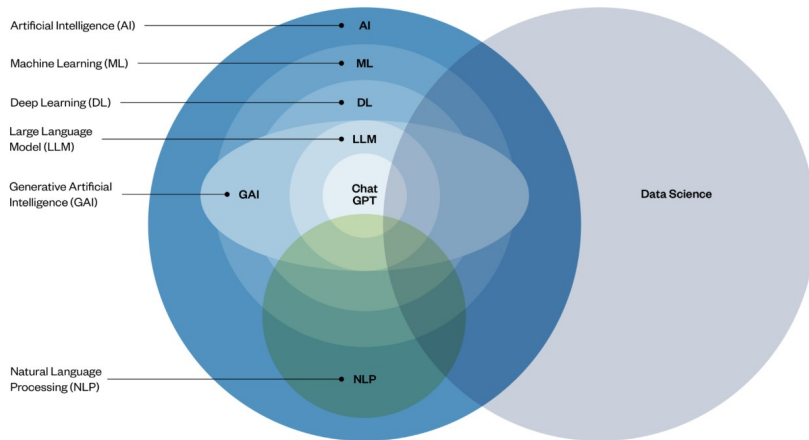
- percepcję środowiska;
- rozwiązywania problemów zdefiniowanych w danym środowisku.

Wokół sztucznej inteligencji

Sztuczna Inteligencja (SI/AI) to dziedzina nauki skupiająca się na tworzeniu systemów obliczeniowych umożliwiających maszynom:

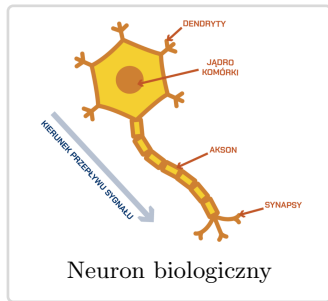
- percepcję środowiska;
- rozwiązywania problemów zdefiniowanych w danym środowisku.

Uczenie maszynowe (ML) to poddziedzina sztucznej inteligencji zorientowana na rozwój metod obliczeniowych pozwalających maszynom na naukę na podstawie danych i generalizację/udoskonalanie swojej zdolności do rozwiązywania problemów na nowe przypadki.



Z czym tak na prawdę mamy do czynienia – czyli jak te skomplikowane terminy mają się do siebie.

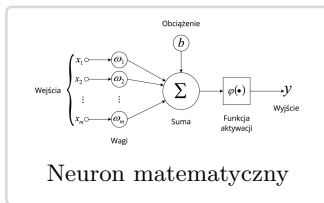
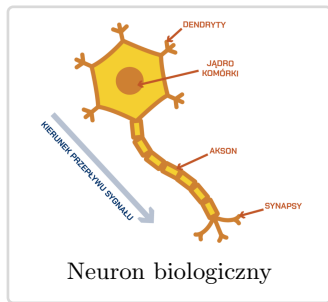
Od neuronu do sieci neuronowej



Opis procesu

Jak od pojedynczego neuronu przeszliśmy do sieci neuronowych – historia skrócona.

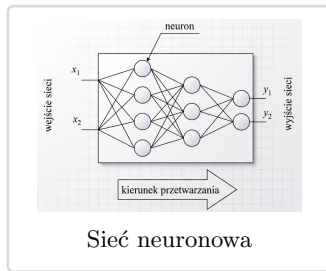
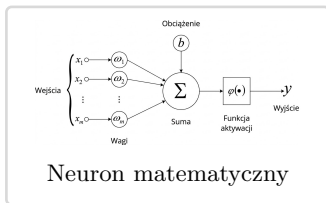
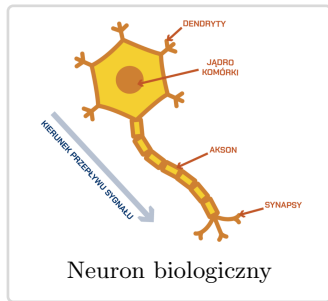
Od neuronu do sieci neuronowej



Opis procesu

Jak od pojedynczego neuronu przeszliśmy do sieci neuronowych – historia skrócona.

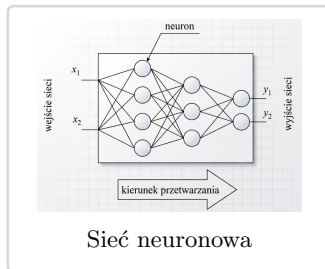
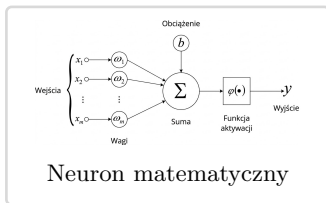
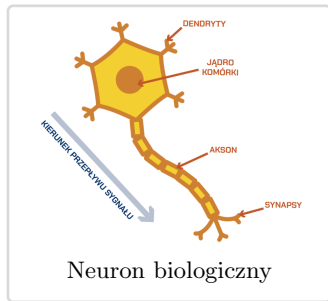
Od neuronu do sieci neuronowej



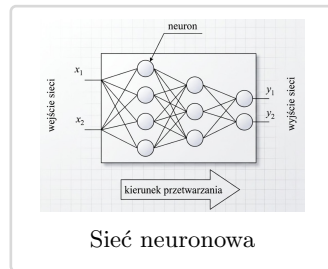
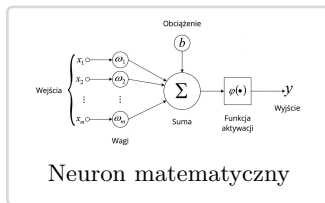
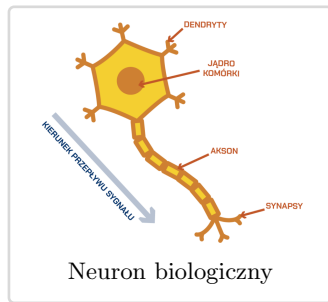
Opis procesu

Jak od pojedynczego neuronu przeszliśmy do sieci neuronowych – historia skrócona.

Od neuronu do sieci neuronowej



Od neuronu do sieci neuronowej



Opis procesu

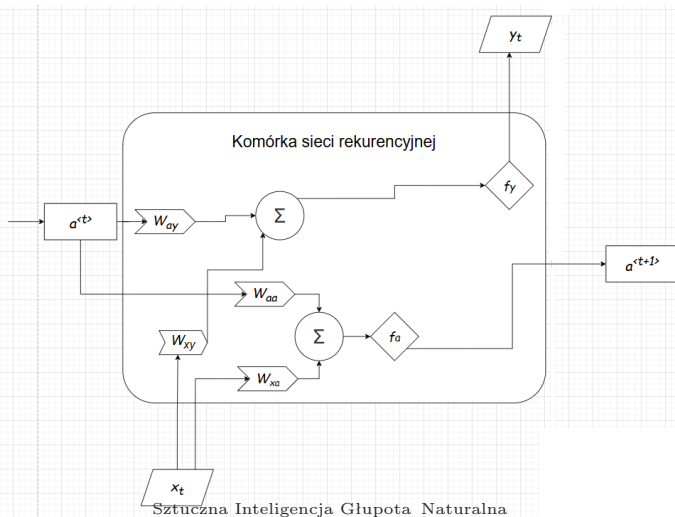
Jak od pojedynczego neuronu przeszliśmy do sieci neuronowych – historia skrócona.

Istotne ograniczenie i jak mu zaradzić

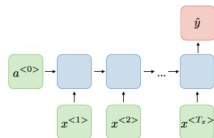
Jak obejść ograniczenie związane z ustalonym rozmiarem wejścia?

Istotne ograniczenie i jak mu zaradzić

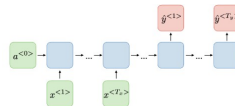
Jak obejść ograniczenie związane z ustalonym rozmiarem wejścia?



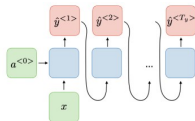
Po co sieci rekurencyjne?



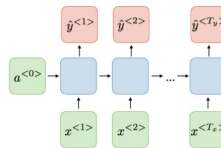
Klasyfikacja dokumentów



Tłumaczenie tekstu



Generowanie muzyki



Analiza gramatyczna tekstu

Generatywna AI

- Pierwotnie – podejście adwersarialne (Policjant vs Falszerz);

Generatywna AI

- Pierwotnie – podejście adwersarialne (Policjant vs Falszerz);
- Później – transformery, tj. w uproszczeniu – nauka przez autouzupełnianie.

Generatywna AI

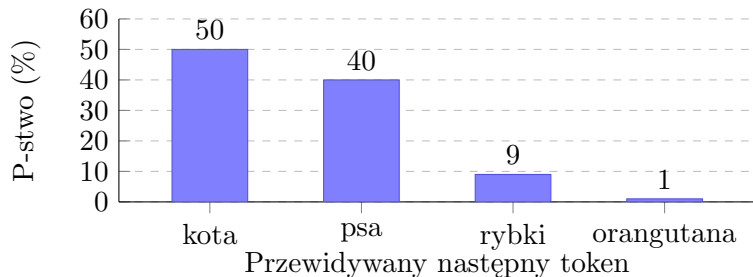
- Pierwotnie – podejście adwersarialne (Policjant vs Falszerz);
- Później – transformery, tj. w uproszczeniu – nauka przez autouzupełnianie.

Ala ma

- Pierwotnie – podejście adversarialne (Policjant vs Falszerz);
- Później – transformery, tj. w uproszczeniu – nauka przez autouzupełnianie.

Ala ma

Wyjście z warstwy Softmax (Logika "Statystycznej Papugi")

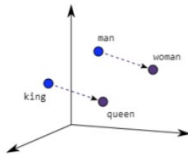


Jak działa nasza rozmowa z LLMem?

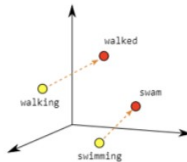


Ciekawostka: Model nie "myśli" całymi zdaniami, lecz statystycznie dobiera najbardziej prawdopodobny następny element układanki.

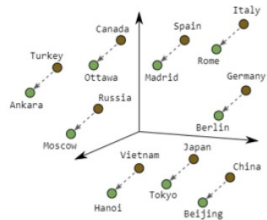
Word2Vec



Male-Female



Verb Tense



Country-Capital

Jak wygląda proces zanurzania słów w przestrzeń znaczeniową na przykładzie Word2Vec.

Halucynacja?

Niech S będzie alfabetem, S^* – zbiorem skończonych ciągów symboli. Załóżmy, że dana jest przestrzeń wyrażeń empirycznie weryfikowalnych $G \subset S^*$ i powiązanych z nimi zapytań empirycznie weryfikowalnych $Q_e \subset S^*$.

Halucynacja?

Niech S będzie alfabetem, S^* – zbiorem skończonych ciągów symboli. Załóżmy, że dana jest przestrzeń wyrażeń empirycznie weryfikowalnych $G \subset S^*$ i powiązanych z nimi zapytań empirycznie weryfikowalnych $Q_e \subset S^*$.

- Stanem faktycznym względem funkcji prawdziwości $f : Q_e \rightarrow \mathcal{P}(G)$ nazywamy zbiór $G_f = \{(q, r) : q \in Q_e, r \in f(q)\}$, gdzie $f(q)$ jest zbiorem wszystkich możliwych sformułowań poprawnej odpowiedzi na pytanie/polecenie q .

Halucynacja?

Niech S będzie alfabetem, S^* – zbiorem skończonych ciągów symboli. Załóżmy, że dana jest przestrzeń wyrażeń empirycznie weryfikowalnych $G \subset S^*$ i powiązanych z nimi zapytań empirycznie weryfikowalnych $Q_e \subset S^*$.

- Stanem faktycznym względem funkcji prawdziwości $f : Q_e \rightarrow \mathcal{P}(G)$ nazywamy zbiór $G_f = \{(q, r) : q \in Q_e, r \in f(q)\}$, gdzie $f(q)$ jest zbiorem wszystkich możliwych sformułowań poprawnej odpowiedzi na pytanie/polecenie q .
- Niech dany będzie model językowy $LM : S^* \times \mathbb{N} \rightarrow S^*$.

Halucynacja?

Niech S będzie alfabetem, S^* – zbiorem skończonych ciągów symboli. Załóżmy, że dana jest przestrzeń wyrażeń empirycznie weryfikowalnych $G \subset S^*$ i powiązanych z nimi zapytań empirycznie weryfikowalnych $Q_e \subset S^*$.

- Stanem faktycznym względem funkcji prawdziwości $f : Q_e \rightarrow \mathcal{P}(G)$ nazywamy zbiór $G_f = \{(q, r) : q \in Q_e, r \in f(q)\}$, gdzie $f(q)$ jest zbiorem wszystkich możliwych sformułowań poprawnej odpowiedzi na pytanie/polecenie q .
- Niech dany będzie model językowy $LM : S^* \times \mathbb{N} \rightarrow S^*$. Jeśli istnieją takie $w \in Q_e$ i $n \in \mathbb{N}$, dla których $LM(w, n) \notin f(w)$, to mówimy, że model LM halucynuje, a $LM(w)$ nazywamy halucynacją.

Halucynacje i ich typy

Model językowy LM halucynuje
wewnętrznie, gdy

$$\exists_{w \in Q_e} [\text{LM}(w) \notin f(w) \quad \wedge \quad \text{LM}(w) \in G].$$

Pytanie:

Halucynacje i ich typy

Model językowy LM halucynuje
wewnętrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \wedge LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny
fotosyntezy?”

Odpowiedź modelu:

Halucynacje i ich typy

Model językowy LM halucynuje
wewnętrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \quad \wedge \quad LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny
fotosyntezy?”

Odpowiedź modelu:

”Produktem ubocznym fotosyntezy jest
dwutlenek węgla.”

Stan faktyczny:

Halucynacje i ich typy

Model językowy LM halucynuje
wewnętrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \quad \wedge \quad LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny
fotosyntezy?”

Odpowiedź modelu:

”Produktem ubocznym fotosyntezy jest
dwutlenek węgla.”

Stan faktyczny:

”Produktem ubocznym fotosyntezy jest
tlen.”

Halucynacje i ich typy

Model językowy LM halucynuje wewnątrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \wedge LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny fotosyntezy?”

Odpowiedź modelu:

”Produktem ubocznym fotosyntezy jest dwutlenek węgla.”

Stan faktyczny:

”Produktem ubocznym fotosyntezy jest tlen.”

Model językowy LM halucynuje zewnątrznie, gdy

$$\exists_{w \in G} [LM(w) \notin f(w) \wedge LM(w) \notin G].$$

Halucynacje i ich typy

Model językowy LM halucynuje wewnątrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \wedge LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny fotosyntezy?”

Odpowiedź modelu:

”Produktem ubocznym fotosyntezy jest dwutlenek węgla.”

Stan faktyczny:

”Produktem ubocznym fotosyntezy jest tlen.”

Model językowy LM halucynuje zewnątrznie, gdy

$$\exists_{w \in G} [LM(w) \notin f(w) \wedge LM(w) \notin G].$$

”Ile wyniesie płaca minimalna w Polsce w 2040 roku? Podaj tylko konkretną kwotę z dokładnością do 200 zł.

Odpowiedź modelu:

Halucynacje i ich typy

Model językowy LM halucynuje wewnątrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \wedge LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny fotosyntezy?”

Odpowiedź modelu:

”Produktem ubocznym fotosyntezy jest dwutlenek węgla.”

Stan faktyczny:

”Produktem ubocznym fotosyntezy jest tlen.”

Model językowy LM halucynuje zewnątrznie, gdy

$$\exists_{w \in G} [LM(w) \notin f(w) \wedge LM(w) \notin G].$$

”Ile wyniesie płaca minimalna w Polsce w 2040 roku? Podaj tylko konkretną kwotę z dokładnością do 200 zł.

Odpowiedź modelu:

”Przewidywana płaca minimalna w Polsce w 2040 roku wyniesie około 11 600 zł.”

Stan faktyczny:

Halucynacje i ich typy

Model językowy LM halucynuje wewnątrznie, gdy

$$\exists_{w \in Q_e} [LM(w) \notin f(w) \wedge LM(w) \in G].$$

Pytanie:

”Jaki gaz powstaje jako produkt uboczny fotosyntezy?”

Odpowiedź modelu:

”Produktem ubocznym fotosyntezy jest dwutlenek węgla.”

Stan faktyczny:

”Produktem ubocznym fotosyntezy jest tlen.”

Model językowy LM halucynuje zewnątrznie, gdy

$$\exists_{w \in G} [LM(w) \notin f(w) \wedge LM(w) \notin G].$$

”Ile wyniesie płaca minimalna w Polsce w 2040 roku? Podaj tylko konkretną kwotę z dokładnością do 200 zł.

Odpowiedź modelu:

”Przewidywana płaca minimalna w Polsce w 2040 roku wyniesie około 11 600 zł.”

Stan faktyczny:

”Nie jest możliwe podanie sensownej prognozy płacy minimalnej w Polsce na 2040 rok z taką dokładnością.”

- **Co kupisz:** Na piętrze ze sprzętem RTV/AGD mają mnóstwo taniej elektroniki, w tym tajskie i chińskie trymery za ułamek ceny, ale też podstawowe modele znanych marek. Kiedyś sam kupowałem tam najprostszą maszynkę za ok. 300 THB (35 zł) w kryzysowej sytuacji i dała radę!

To jest pewnie ciężko zaklasyfikować, ale łatwo doświadczyć.

To skąd halucynacje?

- Dane:

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy
(niedoreprezentowanie);

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy
(niedoreprezentowanie);
 - Dezaktualizacja wiedzy;

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy (niedoreprezentowanie);
 - Dezaktualizacja wiedzy;
- Trening:

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy (niedoreprezentowanie);
 - Dezaktualizacja wiedzy;
- Trening:
 - MLE+RLHF (estymacja największej wiarygodności);

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy (niedoreprezentowanie);
 - Dezaktualizacja wiedzy;
- Trening:
 - MLE+RLHF (estymacja największej wiarygodności);
 - Memoryzacja korpusowa;

To skąd halucynacje?

- Dane:

- Niskiej jakości dane z sieci;
- Dziury w wiedzy (niedoreprezentowanie);
- Dezaktualizacja wiedzy;

- Trening:

- MLE+RLHF (estymacja największej wiarygodności);
- Memoryzacja korpusowa;

- Architektura:

To skąd halucynacje?

- Dane:

- Niskiej jakości dane z sieci;
- Dziury w wiedzy (niedoreprezentowanie);
- Dezaktualizacja wiedzy;

- Trening:

- MLE+RLHF (estymacja największej wiarygodności);
- Memoryzacja korpusowa;

- Architektura:

- Konwersacja wykracza poza rozmiar okna kontekstowego;

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy (niedoreprezentowanie);
 - Dezaktualizacja wiedzy;
- Trening:
 - MLE+RLHF (estymacja największej wiarygodności);
 - Memoryzacja korpusowa;
- Architektura:
 - Konwersacja wykracza poza rozmiar okna kontekstowego;
- Zachowanie w czasie inferencji:

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy (niedoreprezentowanie);
 - Dezaktualizacja wiedzy;
- Trening:
 - MLE+RLHF (estymacja największej wiarygodności);
 - Memoryzacja korpusowa;
- Architektura:
 - Konwersacja wykracza poza rozmiary okna kontekstowego;
- Zachowanie w czasie inferencji:
 - Pochlebstwo/Niejasne instrukcje/Wysoka temperatura modelu;

To skąd halucynacje?

- Dane:
 - Niskiej jakości dane z sieci;
 - Dziury w wiedzy (niedoreprezentowanie);
 - Dezaktualizacja wiedzy;
- Trening:
 - MLE+RLHF (estymacja największej wiarygodności);
 - Memoryzacja korpusowa;
- Architektura:
 - Konwersacja wykracza poza rozmiary okna kontekstowego;
- Zachowanie w czasie inferencji:
 - Pochlebstwo/Niejasne instrukcje/Wysoka temperatura modelu;

Czy jesteśmy skazani na głupotę?

- Duże modele językowe nie "rozumieją" języka/tekstów w ten sposób, jaki my;

Czy jesteśmy skazani na głupotę?

- Duże modele językowe nie "rozumieją" języka/tekstów w ten sposób, jaki my;
- Większości z nas nie stać na dotrenowanie modelu w domowym zaciszu;

Czy jesteśmy skazani na głupotę?

- Duże modele językowe nie "rozumieją" języka/tekstów w ten sposób, jaki my;
- Większości z nas nie stać na dotrenowanie modelu w domowym zaciszu;
- Musimy polegać na innych sztuczkach:

Czy jesteśmy skazani na głupotę?

- Duże modele językowe nie "rozumieją" języka/tekstów w ten sposób, jaki my;
- Większości z nas nie stać na dotrenowanie modelu w domowym zaciszu;
- Musimy polegać na innych sztuczkach:
 - RAG;

Czy jesteśmy skazani na głupotę?

- Duże modele językowe nie "rozumieją" języka/tekstów w ten sposób, jaki my;
- Większości z nas nie stać na dotrenowanie modelu w domowym zaciszu;
- Musimy polegać na innych sztuczkach:
 - RAG;
 - Chain of Thought;

Czy jesteśmy skazani na głupotę?

- Duże modele językowe nie "rozumieją" języka/tekstów w ten sposób, jaki my;
- Większości z nas nie stać na dotrenowanie modelu w domowym zaciszu;
- Musimy polegać na innych sztuczkach:
 - RAG;
 - Chain of Thought;
 - Self-refine;

Czy zatem AI jest głupie?

- Różnorodne modele sztucznej inteligencji bardzo ułatwiają nam życie (diagnostyka, prognozy itd.);
- LLMy są dobre w symulowaniu procesów myślowych, ale nie myślą w klasycznym sensie;
- Istnieją modele wnioskujące, które symulują procesy myślowe i logiczne wnioskowanie – ale przeciętny człowiek nie jest w stanie z nimi ”porozmawiać” (AI-Hilbert, Lean);
- AI staje się progresywnie mądrzejsze od strony architekturnej i głupsze od strony danych uczących;
- AI obnaża naszą głupotę (a raczej intelektualne lenistwo).

Serdecznie dziękuję za uwagę!

Serdecznie dziękuję za uwagę!

To jest dobry moment na pytania

Serdecznie dziękuję za uwagę!

To jest dobry moment na pytania
(i omówienie zadania).

Serdecznie dziękuję za uwagę!

To jest dobry moment na pytania
(i omówienie zadania).

Filip Turoboś
filip.turobos@p.lodz.pl