

Sztuczna inteligencja weszła w świat niemowlaka. Dostłownie. Niezwykły eksperyment

© Piotr Cieśliński

Gazeta Wyborcza, 02.02.2024

https://wyborcza.pl/7,75400,30657885,sztuczna-inteligencja-odbiera-swiat-oczami-i-uszami-dziecka.html#s=S.embed_article-K.C-B.1-L.1.zw



Naukowcy z New York University sfilmowali proces uczenia się mowy z perspektywy małego dziecka za pomocą lekkiej kamery, zainstalowanej na jego głowie. A potem to nagranie użyli do szkolenia sztucznej inteligencji (NYU's Center for Data Science)

Eksperyment miał sprawdzić, czy sztuczna inteligencja jest w stanie nauczyć się mówić w ten sam sposób co małe dziecko. Czyli słysząc i widząc to samo co rozpoczynający przygodę z życiem maluch.

Systemy sztucznej inteligencji (AI - Artificial Intelligence), takie jak GPT-4, świetnie radzą sobie z ludzkim językiem. Muszą jednak być trenowane na podstawie astronomicznych ilości danych wejściowych. Potrzebują znacznie więcej przykładów, niż uczące się mówić dzieci.

Najlepsze systemy AI są szkolone na bilionach słów. Dzieci, jak się szacuje, w czasie nauki mowy słyszą ich nawet milion razy mniej.

Z tego względu badacze skłaniali się do wniosku, że uczenie i rozwój człowieka przebiegają zupełnie inaczej, niż dzieje się to w obecnych systemach AI. Mimo że struktura sieci neuronowych będących podstawą systemów sztucznej inteligencji przypomina budowę mózgu.

Naukowcy z Uniwersytetu Nowego Jorku (NYU) postanowili to sprawdzić. Spróbowali wyszkolić model sztucznej inteligencji nie na ogromnych zbiorach tekstu z internetu, ale wyłącznie na danych wizualnych i dźwiękowych, które w czasie nauki mowy otrzymuje dziecko. Byli ciekawi, jak wiele ich AI zdoła się nauczyć.

Przeprowadzili eksperyment.

Wyposażyli niemowlę w lekką kamerę na głowie i rejestrowali to, co ono widzi i słyszy od szóstego miesiąca życia do drugich urodzin.

Czyli w czasie, gdy większość dzieci zaczyna mówić.

Następnie stworzyli multimodalną sieć neuronową i szkolili ją jedynie na fragmentach nagrań uzyskanych z tej kamery. Okazało się, że ich model AI z powodzeniem zdołał się nauczyć słów i pojęć obecnych w codziennym doświadczeniu niemowlaka.

W najnowszym numerze "Science" autorzy eksperymentu wykazują, że ich model oparty na sieci neuronowej może się nauczyć znacznej liczby słów i pojęć, wykorzystując ograniczone wycinki tego, czego doświadczyło niemowlę. W danych treningowych AI znalazły się bowiem nagrania obejmujące tylko ok. 1 proc. godzin czuwania dziecka (w sumie 61 godzin wideo).

Ale nawet to wystarczyło do efektywnej nauki języka.

- Po raz pierwszy pokazujemy, że sieć neuronowa wytrenowana na realistycznych pod względem rozwojowym danych wejściowych pochodzących od pojedynczego dziecka może się nauczyć łączyć słowa z ich wizualnymi odpowiednikami - mówi Wai Keen Vong, pierwszy autor publikacji.

Materiał filmowy zawierał blisko ćwierć miliona słów (tj. liczbę wypowiedzianych słów, wiele z nich wielokrotnie), które były powiązane z klatkami wideo przedstawiającymi to, co dziecko widziało, gdy te słowa były wypowiadane. Dane obejmowały szeroki zakres różnych aktywności dziecka, m.in. posiłków, czytania książek i zabaw.

Badacze z NYU szkolili swoją multimodalną sieć neuronową za pomocą dwóch oddzielnych modułów: jeden odbierał pojedyncze klatki wideo, a drugi transkrybowane na tekst słowa, jakie były wypowiadane do dziecka.

Te dwa moduły połączono i przeszkolono przy użyciu algorytmu, który wyszukiwał powiązania między obrazem i słowami.

Na przykład, gdy rodzic mówił coś do dziecka, algorytm przyjmował, że użyte słowa prawdopodobnie odnosiły się do czegoś, co było w zasięgu wzroku malucha.

- Dzięki temu model otrzymywał wskazówkę, które słowa należy skojarzyć z konkretnymi obiektami - tłumaczy Vong. - AI stopniowo uczyła się, które słowa pasują do poszczególnych elementów wizualnych lub koncepcji.

Po przeszkoleniu modelu naukowcy przetestowali go tak, jak sprawdza się wyniki nauki słów u niemowląt. AI miała np. wybrać, który z czterech przedstawionych jej obrazów pasuje do danego słowa. Wyniki pokazały, że sztuczna inteligencja była w stanie nauczyć się znacznej liczby słów i pojęć obecnych w codziennym życiu dziecka.

Co więcej, niektóre z wyuczonych słów AI potrafiła uogólnić - odnieść je do obrazów, z którymi nie miała do czynienia podczas szkolenia.

- Być może uchwyciliśmy, jak przebiega proces uczenia się pierwszych słów przez dziecko - uważa psycholog Brenden Lake, współautor publikacji.

AI poniosła jednak kilka spektakularnych porażek, np. nie potrafiła uchwycić pojęć "pokój" albo "ręce". Ręce często pojawiały się w danych wizualnych, ale wymawiane były rzadko, najczęściej podczas zabawy na plaży, a więc AI mylnie utożsamiała słowo "ręce" z piaskiem.

Naukowcy są też świadomi ograniczeń swojego modelu AI. Nauka u dzieci jest rozciągnięta w czasie, a ta sztuczna inteligencja dostała ciąg nieruchomych klatek wideo, co prawdopodobnie ograniczyło jej zdolność uczenia się czasowników.

Poza tym dzieci są ucieleśnionymi, aktywnymi uczniami, podczas gdy AI nie odbiera żadnych wrażeń z ciała, uczy się pasywnie.

Wreszcie AI miała dostęp tylko do transkrybowanych wypowiedzi, była pozbawiona takich przydatnych wskazówek, jak intonacja czy akcent.

Mimo to zdaniem amerykańskich badaczy eksperyment może zmienić nasze rozumienie wczesnego przyswajania języka i pojęć u dzieci.

Wciąż trwają bowiem debaty na temat tego, czy do nauki mowy człowiek wykorzystuje jakieś wrodzone struktury w mózgu. Czy nasze mózgi są specjalnie przystosowane i przygotowane do przyswojenia języka, czy też po prostu jest to nauka skojarzeniowa, taka jak w przypadku modelu AI amerykańskich naukowców.

- Wygląda na to, że wystarczają skojarzenia - mówi Brenden Lake.

Can AI Learn Language the Way Babies Do? – film na YT - <https://youtu.be/tFZcXayvCsw>

© Piotr Cieśliński - Dziennikarz i redaktor naukowy, z wykształcenia fizyk, pracuje w "Gazecie Wyborczej" od roku 1991, popularyzuje odkrycia z nauk ścisłych, kosmosu, matematyki i techniki