

# AI coraz bliższa ludzkiemu sposobowi przyswajania wiedzy

18 marca 2024

Naukowcy z Uniwersytetu Nowojorskiego opracowali model uczenia maszynowego, który naśladuje sposób, w jaki dzieci uczą się języka. Wykorzystując nagrania wideo i audio rejestrowane z perspektywy małego dziecka, model z powodzeniem nauczył się dopasowywać słowa do odpowiadających im obrazów. Wysoka, dwukrotnie większa niż w przypadku dużo większych modeli skuteczność pozwala przypuszczać, że jesteśmy bliscy zrozumienia tego, w jaki sposób dzieci zaczynają rozumieć język i go używać. Dotychczas poruszaliśmy się w obszarze niepotwierdzonych, a jedynie uprawdopodobnionych teorii.

„Jest wiele różnych teorii na temat tego, jak dzieci przyswajają język. Z jednej strony mamy teorię naturalnego przyswajania języka – wielu badaczy twierdzi, że dzieci rodzą się ze specjalnymi elementami czy wiedzą wbudowaną w mózg, która w wyjątkowy sposób pozwala nam uczyć się języka. Z drugiej strony mamy teorię bazującą na wychowaniu, zgodnie z którą dzieci uczą się języka głównie na podstawie doświadczeń zmysłowych. Nie chodzi tu o zdolności wrodzone, ale doświadczenia zdobywane codziennie, które pozwalają nam przyswajać na przykład język angielski” – wyjaśnia w rozmowie z agencją Newseria Wai Keen Vong z Centrum Analityki Danych Uniwersytetu Nowojorskiego.

W swojej pracy naukowcy oparli się w przeważającej części na drugiej z wymienionych teorii. Opracowując model, chcieli jednocześnie głębiej zrozumieć proces wczesnego przyswajania języka.

Zbiór danych będących bazą do pracy zebrano za pomocą kamery montowanej na głowie i noszonej przez jedno dziecko. Rejestrowała ona nagrania wideo i audio, towarzysząc dziecku w

okresie pomiędzy 6. a 25. miesiącem życia. Zbiór danych obejmował 600 tys. klatek wideo w połączeniu i 37,5 tys. transkrybowanych wypowiedzi. Kluczowe było to, by badanie nie odbywało się w warunkach laboratoryjnych, lecz jak najbliższych naturalnemu środowisku.

„To największy model tego rodzaju, jaki dotychczas powstał. Istnieją już zbiory danych zawierające nagrania ludzi mówiących do swoich dzieci, ale żaden z nich nie obejmował tak długiego okresu i nie zawierał materiałów wideo przedstawiających to, co dziecko widzi w danym momencie. Nasze dane dają wyjątkową perspektywę na kwestię przyswajania języka. Ponadto połączyliśmy te dane z multimodalną siecią neuronową. Model jest dość ogólny w zakresie sposobu uczenia się. Przesyła się dane i proste aktualizacje reguł, na podstawie których przebiega uczenie się. Fakt, że uczenie się zachodzi, dostarcza nowego spojrzenia na wczesne przyswajanie języka, jest ono inne od wcześniejszego podejścia w tej kwestii. Dużo większy nacisk jest położony na uczenie się i na to, co można dzięki niemu osiągnąć” – podkreśla Wai Keen Vong.

Wydajność modelu została oceniona na podstawie kategoryzacji polegającej na przypisywaniu słów do odpowiadających im wizualizacji. Dokładność klasyfikacji wyniosła 61,6 proc. Dla porównania przeprowadzono testy na kontrastowej sieci neuronowej obrazu i tekstu CLIP, wyszkolonej na dużo większym zasobie danych. W tym przypadku dokładność klasyfikacji wyniosła zaledwie 34,7 proc.

„Po raz pierwszy potwierdzono, że nasz model może się uczyć tych słów na podstawie rzeczywistych doświadczeń dzieci. Ta możliwość uczenia się potwierdza słuszność koncepcji, że może być to jeden ze sposobów przyswajania tego rodzaju pojęć przez dzieci. Nie możemy stwierdzić tego kategorycznie, ale jest to bardzo sugestywny i interesujący kierunek dalszych badań” – zapowiada badacz.

Według „Precedence Research” światowy rynek uczenia

maszynowego przekroczył w 2022 roku 38 mld dol. przychodów. Do 2032 roku jego wielkość urośnie do ponad 771 mld dol.

Źródło: [Newseria.pl](https://www.newseria.pl)