

## ROZUMIENIE TEKSTU PISANEGO



### GPT-2: ALGORYTM DO TWORZENIA TEKSTÓW

Proszę przeczytać tekst i odpowiedzieć na poniższe pytania.

Organizacja OpenAI, którą finansowo wspierał miliarder Elon Musk stworzyła algorytm służący do generowania gotowych tekstów. GPT-2 (tak nazwano algorytm) opiera się na sztucznej inteligencji i jest w stanie po otrzymaniu próbki tekstu napisać dłuższy i spójny tekst w tym samym stylu: od plotek, przez bieżące wydarzenia polityczne i odkrycia naukowe, po prozę inspirowaną np. twórczością George'a Orwella. Oprogramowanie bazuje na danych pochodzących z kilku milionów stron internetowych. Duża baza ucząca sprawia, że generuje naprawdę dobre teksty i wygrywa z konkurencją.

Początkowo GPT-2 został zablokowany, ponieważ badacze uznali, że dostarcza zbyt dobre teksty i jest trochę niebezpieczny. Obawiano się, że będzie wykorzystywany w niewłaściwy sposób, na przykład do generowania fake newsów, czyli wprowadzających w błąd lub szkodliwych treści.

Obawy te były uzasadnione, gdyż istniały już prostsze, zdolne do napisania spójnego zdania programy (boty), wykorzystywane w celu rozpowszechniania fake newsów (większość kont na Twitterze prowadzą takie właśnie boty). Jednak dotychczasowe generatory tekstu nie radziły sobie ze składnią dłuższych zdań, a w połowie akapitu potrafiły zmienić temat.

1. Do czego służy oprogramowanie GPT-2?
2. Jakie teksty potrafi wygenerować GPT-2?
3. Czym różni się GPT-2 od innych, podobnych programów?
4. Jakie błędy popełniały dotychczasowe generatory tekstu?
5. Dlaczego oprogramowanie GPT-2 może być niebezpieczne?
6. Do czego mogłoby się przydać GPT-2 w przypadku prowadzenia kampanii dezinformacyjnej?
7. Czy GPT-2 posługuje się językiem polskim?

Nowy program jest tych wad praktycznie pozbawiony i aby przeprowadzić kampanię dezinformacyjną opartą na wielu tekstach, jest w stanie w okamgnieniu wyprodukować setki przekonujących, a wprowadzających w błąd wiadomości.

Po kilku miesiącach OpenAI opublikowało ograniczoną wersję modelu GPT-2, aby sprawdzić czy ich oprogramowanie będzie używane przez osoby rozpowszechniające fake newsy.

Ponieważ nic takiego się nie wydarzyło, więc startup zdecydował się na udostępnienie pełnego modelu zawierającego 1,5 miliarda parametrów.

Dostępny jest on na stronie [talktotransformer.com](http://talktotransformer.com). Badacze ocenili wiarygodność generowanych tekstów na 6,91 w skali od 0 do 10. Ocena ta może wzrosnąć wraz z rozmiarem modelu.

Niestety algorytm nie generuje tekstów w języku polskim. Na zwykłe "dzień dobry" reaguje w dziwny sposób:

:((sad face))

*Zgadzysz otkondzieńszy stry i wyglądenie podkło o 22 januari 2015 powinieności bardzo wydać się dobrającym i karwa!*  
All Rights Reserved.