# Wojna sztucznej inteligencji na platformie Kaggle

7 sierpnia 2017

Najlepszą obroną przeciwko złej sztucznej inteligencji może być dobra sztuczna inteligencja. Systemy SI będą przyszłością bezpieczeństwa, a używać ich będą zarówno cyberprzestępcy jak i osoby odpowiedzialne za ochronę infrastruktury IT.

Autorzy platformy Kaggle, na której twórcy modeli statystycznych i analitycznych konkurują między sobą budując modele najlepiej opisujące dostępne zestawy danych, właśnie ogłosili rozpoczęcie zawodów pomiędzy algorytmami sztucznej inteligencji. Zawody, które potrwają przez pięć miesięcy, mają na celu wyłonienie najlepszych algorytmów ofensywnych i defensywnych. „To świetny pomysł, by w jednym miejscu zgromadzić badania dotyczące zarówno technik oszukiwania systemów maszynowego uczenia się, jak systemów odpornych na tego typu ataki" – cieszy się profesor Jeff Clune z University of Wyoming, który bada granice maszynowego uczenia się.

Konkurs będzie składał się z trzech elementów. W ramach pierwszego z wyzwań zadaniem będzie oszukanie mechanizmu maszynowego uczenia się tak, by przestał właściwie pracować. Druga konkurencja zakłada próbę wymuszenia na algorytmie nieprawidłowe sklasyfikowanie danych. W ramach trzeciego wyzwania uczestnicy będą musieli stworzyć systemy jak najbardziej odporne na ataki. Wstępne wyniki zostaną zaprezentowane już pod koniec bieżącego roku.

Maszynowe uczenie się, w ramach którego maszyna samodzielnie tworzy algorytm służący do rozwiązania stojącego przed nią problemu, staje się coraz ważniejszym narzędziem pracy w wielu gałęziach przemysłu i nauki. Od dawna jednak wiadomo, że technologię maszynowego uczenia można oszukać. Na przykład udaje się to spamerom, którzy oszukują automatyczne algorytmy

wyłapujące niechciane e-maile. Czasami, ku zaskoczeniu ekspertów, okazuje się, że nawet niezwykle zaawansowane algorytmy można oszukać w bardzo prosty sposób. Zorganizowanie „bitwy algorytmów" może być dobrym sposobem na ich udoskonalenie. „Antagonistyczne maszynowe uczenie się jest trudniejsze do badania niż jego konwencjonalna odmiana. Trudno jest bowiem stwierdzić, czy to atak był tak silny, czy obrona tak słaba" – mówi Ian Goodfellow z Google Brain, które organizuje zawody.

Jako, że wspomniane algorytmy są coraz częściej wykorzystywane, rosną obawy o ich bezpieczeństwo. Tym bardziej, że maszynowe uczenie się jest coraz popularniejsze i może być wykorzystane zarówno do prowadzenia pozytywnych jak i negatywnych działań. „Bezpieczeństwo IT zmierza w kierunku maszynowego uczenia. Ci źli będą wykorzystywali maszynowe uczenie do zautomatyzowania ataków, a my do obrony" – mówi Goodfellow.

Autorstwo: Mariusz Błoński
Na podstawie: Technology Review
Źródło: KopalniaWiedzy.pl