

Gdynia, 09.01.2019r.

Czy polska firma pomoże Facebookowi w walce z cyberbullyingiem?

Szef Facebooka Mark Zuckerberg chce aktywnie walczyć z toksycznymi wypowiedziami rozpowszechnianymi na jego platformie. Mimo to, zachodnia prasa przedstawia Facebooka jako bezlitosną i samolubną korporację, a sam Zuckerberg odpiera te ataki i wymienia całą listę zmian jaką ostatnio wprowadził, aby ograniczyć skutki toksycznych wpisów. Z pomocą może przyjść sztuczna inteligencja z Polski, nad którą pracują inżynierowie z gdyńskiego laboratorium Samurai Labs.

W Gdyni powstała sztuczna inteligencja, która może być lekarstwem na bólczki Facebooka i innych firm technologicznych borykających się z problemem toksycznych treści. Dzięki rozwiązaniu Samurai Labs możliwa jest znaczna redukcja przemocy w sieci, z którą codziennie spotykają się osoby korzystające z mediów społecznościowych czy gier online.

Waga problemu

Cała armia moderatorów codziennie przegląda setki milionów postów zamieszczanych przez użytkowników Facebooka. Sam Zuckerberg, wśród zmian mających na celu redukcję toksycznych treści na portalu, mówi o zwiększeniu liczby pracowników zajmujących się moderacją postów do ponad 30 tys. osób. – *to ważne, że Facebook zauważa problem, bo on jest realny i badając zjawisko cyberprzemocy spotykamy się z nim na co dzień. Problem ze świata wirtualnego przenosi się także na zachowania ludzi w świecie rzeczywistym, gdzie agresja w sieci prowadzi do eskalacji przemocy przybierając różnorodną postać* – mówi Michał Wroczyński, psychiatra, a zarazem szef Samurai Labs, startupu, którego celem jest redukcja i zapobieganie przemocy w sieci z pomocą sztucznej inteligencji.

Skala toksycznych treści

Problem toksycznych wypowiedzi to złożona kwestia. Sama skala działania Facebooka jest ogromna, a do tego dochodzą inne miejsca w sieci, w których można komentować, rozmawiać czy dzielić się zdjęciami. – *tylko na Facebooku co minutę publikowane jest ponad pół miliona komentarzy. Zatrudnianie moderatorów, aby chronić użytkowników przed niepożądanymi treściami jest rozwiązaniem doraźnym i niewystarczającym. Tu potrzebne jest rozwiązanie systemowe. Do tego typu zadań najlepiej zaangażować sztuczną inteligencję.* – tłumaczy Wroczyński z Samurai Labs. Nad takim rozwiązaniem pracuje właśnie polski zespół Samurai Labs, którego oprogramowanie nie dość, że automatycznie wykrywa niepożądane treści, to w przeciwieństwie do moderatorów zablokuje treść zanim w ogóle dotrze ona do odbiorcy.

Natychmiastowa reakcja

To, z czym mierzy się Facebook i inne firmy technologiczne dostrzegające problem cyberbullyingu, to kwestia szybkiej reakcji na niepożądaną treść. Rozwiązanie opracowywane przez Polaków funkcjonuje w czasie rzeczywistym działając prewencyjnie. – *Gdy ktoś chce wysłać szkodliwy komentarz lub wiadomość na czacie, opracowywana przez nas sztuczna inteligencja analizuje ich treść i w razie*

konieczności wysłać do autora stosowne ostrzeżenie oraz blokuje daną wiadomość, jeszcze zanim dotrze ona do odbiorcy i wyrządzi krzywdę. W grach online czy na portalach społecznościowych cyberprzemoc jest niestety dość powszechna, a nam zależy na tym, aby ją ograniczyć. - mówi Wroczyński.

Przewaga technologiczna

System opracowany przez Polaków okazuje się znacznie bardziej skuteczny niż dotychczasowe rozwiązania dostępne na rynku. - *Mark Zuckerberg zakłada, że na skuteczne rozwiązanie, które będzie wykrywać przemoc w Internecie trzeba poczekać jeszcze od 5 do 10 lat. System AI Facebooka wyłapuje zaledwie 38% wypowiedzi, które uznano za toksyczne. – mówi Wroczyński. System Samurai, już na etapie prototypu, myli się czterokrotnie rzadziej niż konkurencja, osiągając wskaźnik precyzji na poziomie ponad 80%. Oznacza to, że na 5 zgłaszanych przypadków myli się tylko raz, a to dopiero początek naszej drogi – tłumaczy Gniewosz Leliwa, dyrektor ds. badań nad sztuczną inteligencją i współzałożyciel Samurai Labs.*

Rozwiązanie opracowywane przez polskich ekspertów od sztucznej inteligencji ma szansę wkrótce zaowocować wdrożeniem w pełnej skali. Obecnie twórcy pracują nad pozyskaniem partnera z segmentu e-sportu do pracy nad pilotażowym rozwiązaniem pozwalającym przetestować ich zaawansowane algorytmy. Rozwiązanie Samurai Labs wzbudziło zainteresowanie rodzimych inwestorów. W pierwszej rundzie finansowania w startup zainwestowali kapitan polskiej reprezentacji Robert Lewandowski i anioł biznesu Dawid Urban. Samurai Labs ma duże szanse na uzyskanie statusu pierwszego polskiego jednorożca, czyli firmy wycenianej na ponad miliard dolarów.