

Algorytm automatycznie nie przetwarzające dane i podejmujące decyzje, oparte na technologii sztucznej inteligencji (AI), stają się powoli rzeczywistością także w Polsce. Jednym z ostatnich zastosowań jest określenie zdolności kredytowej (scoring) w oparciu o dane z oficjalnych rejestrów lub historii rachunku bankowego (co stało się możliwe w wyniku wdrożenia dekretu PSD2). Warto zatem zwrócić uwagę, iż (oprócz innych wymagań prawnych) twórcy takich algorytmów powinni zwracać uwagę na ryzyko dyskryminacji w wyniku działania algorytmu.

Algorytm mnie nie lubi!

Jakkolwiek doświadczenia z działaniem systemów sztucznej inteligencji są na razie raczej ograniczone, to ujawni się już jeden z problemów prawnych związanych z systemami automatycznie przetwarzającymi dane i podejmującymi decyzje – otóż, okazuje się, że rezultaty działania tych systemów mogą wskazywać na dyskryminację „zaszytą” w algorytmie. To zaś pociąga za sobą możliwość wystąpienia ważnych skutków prawnych, w szczególności roszczeń osób dyskryminowanych.

Praktyka kilku ostatnich lat pokazuje szereg przypadków, gdy mechanizm działania AI dyskryminował niektóre osoby (np. z powodów ich rasy, pochodzenia, płci) – na przykład: a) w lutym 2018 r. naukowiec z MIT ogłosił, że trzy z ostatnich systemów AI rozpoznających płeć na podstawie zdjęć miały 99-procentową skuteczność w rozpoznawaniu płci – ale tylko w odniesieniu do białyh mężczyzn; w przypadku, na przykład, ciemnoskórych kobiet skuteczność spadła do 35 procent, gdyż systemy podobały im się bardziej niż mężczyznom; b) algorytm PredPol jest używany w kilku stanach USA do przewidywania, kiedy i gdzie może dojść do przestępstwa – w oparciu o dane statystyczne (chodzilo o uniknięcie wystąpienia części patroli policyjnych do regionów, gdzie faktycznie nie popełnia się najczęściej przestępstw, ale co do których panują uprzedzenia, że rzekomo w tych regionach musi być duża przestępczość, gdyż, na przykład, mieszka tam wiele osób czarnoskórych); jednakże symulacja w 2016 r. przeprowadzona przez NGO w oparciu o prze-

stępstwa narkotykowe w Oakland wykazała, że algorytm nakazywał relatywnie najczęściej wysłanie patroli policyjnych do regionów zamieszkałych przez mniejszości rasowe, podczas gdy przedstawia także częścię popełnianą w innych regionach; prawdopodobnie algorytm nie był uczony w oparciu o statystyki faktycznie popełnionych przestępstw, ale w oparciu o raporty policyjne z patroli;

c) badania z 2015 r. wykazały, że system reklam Google szedł do wniosku, iż w przypadku Palestyńczyka z budlożerem w tle jest bardziej prawdopodobne, że szykuje się on do ataku.

Te proste przykłady pokazują, jak dużym problemem może być kwestia mechanizmów dyskryminujących w przypadku systemów AI – taki sam problem może zaistnieć na przykład w przypadku systemów wspierających diagnozę pacjentów, oceniających preferencje żywieniowe lub polityczne.

ich przez systemy AI. Należy się zatem spodziewać, że niobawem wiele firm z tego typu usług będzie korzystać – warto zatem, by były świadome zagrożen, jakie tkwią w takim rozwiązaniu.

AI dyskryminuje – i co z tego?

Konsekwencje prawne uznania, iż algorytm został zbudowany lub funkcjonuje w oparciu o dyskryminujące

puszczają się recydywy. Skarżący twierdził, że sąd, między innymi, dopuścił się dyskryminacji w oparciu o płeć i rasę skazanego. Zaskarzenie tym razem okazało się nieskuteczne, ale problem prawny, jaki sygnalizuje ta właśnie sprawa, pozostaje nierozstrzygnięty. Należy się zatem liczyć, iż w przyszłości tego typu zarzuty będą się powtarzały i przynajmniej w niektórych przypadkach sądy mogą uznać, iż wadliwy algorytm skutkuje wadliwością decyzji lub roz-

Czy sztuczna inteligencja może być do mnie uprzedzona?

PRZEMYSŁAW WIERZBICKI

Już na etapie projektowania algorytmu AI powinno się przeanalizować możliwe warianty i efekty jego działania – także w zakresie najpoważniejszych kryteriów dyskryminujących.



MAT PRAS

częścię wyświetał ogłoszenia o wysoko płatnych stanowiskach pracy, na które można aplikować, mężczyznom niż kobietom, a wynik wyszukiwania grafik do zapytania „CEO” pokazywały prawie trzykrotnie mniej zdjęć kobiet, niż wynikałoby to z faktycznej proporcji kobiet wśród CEO w USA, co zdaniem badaczy, może utwierdzać użytkowników internetu w uprzedzeniu, że kobiety nie nadają się do sprawowania kierowniczych ról;

d) w październiku 2017 r. izraelska policja aresztowała Palestyńczyka, który umieszczał na Facebook swoje zdjęcie na tle budlożera z podpisem, który z kolei automatycznie system tłumaczenia używany przez Facebooka błędnie przetłumaczył na „zaatakujmy ich” (tymczasem oryginalny podpis w języku arabskim oznaczał „dłżej dobyte”, ale był bardzo podobny właśnie do słowa „zaatakujmy ich”) – wyglądało zatem na to, iż algorytm Facebooka, skontrolowany z dwoma możliwymi wariantami tłumaczenia, do-

W naszym kraju za chwilę z pewnością rozpowszechnią się systemy analizujące dane i wspierające decyzje w odniesieniu do zdolności kredytowej osób ubiegających się o pożyczkę lub kredyt w banku (scoring). Wejście w życie przepisów implementujących dyrektywę PSD2 otworzyło bankowe bazy danych dla

” W przyszłości usłyszymy o imponujących odszkodowaniach dla poszkodowanych

podmiotów” trzecich. Wystarczy zgoda osoby ubiegającej się o pożyczkę, żeby potencjalny pożyczkobiorca sprawdził, jak kształtowały się obroty ewentualnego pożyczkobiorcy na koniach w innych bankach (czy, na przykład, nie było tam zajęć komorniczych) albo w biurach informacyjnych kredytowej. Przetwarzanie tak dużych ilości danych z pewnością będzie bardziej wydajne w przypadku przetwarzania

rozwiązania, mogą być dwójakie:

a) osoby doknięte „dyskryminującym” działaniem systemu, opartyymi na efektach działania „dyskryminującego” systemu AI, mogą podważać fakcie decyzje lub rozstrzygnięcia; Jakkolwiek w Polsce to dopiero plesn odległej przyszłości;

b) w październiku 2017 r. izraelska policja aresztowała Palestyńczyka, który umieszczał na Facebook swoje zdjęcie na tle budlożera z podpisem, który z kolei automatycznie system tłumaczenia używany przez Facebooka błędnie przetłumaczył na „zaatakujmy ich” (tymczasem oryginalny podpis w języku arabskim oznaczał „dłżej dobyte”, ale był bardzo podobny właśnie do słowa „zaatakujmy ich”) – wyglądało zatem na to, iż algorytm Facebooka, skontrolowany z dwoma możliwymi wariantami tłumaczenia, do-

Na zakończenie warto pamiętać, iż kwestia dyskryminacji jest tylko jednym z istotnych zagadnień prawnych, z jakimi muszą się liczyć podmioty tworzące lub obsługujące się algorytmami opartymi na AI – innym bardzo istotnym zagadnieniem w tym wypadku jest zapewnienie zgodności z RODO, które jako pierwszy europejski akt prawny bezpośrednio i wyraźne adresuje kwestię automatycznego przetwarzania danych osobowych i wyniki przetwarzania automatycznej decyzji.

Przemysław Wierzbicki
Advokat, co-founder Nowy Wymiar Prawa, uczestnik
HUMANINT Winter School on AI
2019 „Human behaviour and
Machine Intelligence”
organizowanej przez Joint
Research Center Komisji
Europejskiej

Rzecz o Polityce

W KAŻDĄ ŚRODĘ



Kto, czym i jak gra w polskiej polityce? Najbardziej skuteczni politycy, najlepiej prowadzone kampanie. Wnikliwe analizy, zaskakujące opinie

CZYTAJ WIĘCEJ NA rp.pl/rzeczopolityce