XI Konferencja Naukowa
Bezpieczeństwo w Internecie. Analityka danych
Warszawa 6-7.06.2019

Etyczne problemy analizy wielkich zbiorów danych

Jerzy Cytowski prof. UKSW

Urząd Ochrony Danych Osobowych

Naukowe Centrum Prawno-informatyczne

UKE

Wymiar danych

- Każdego dnia kreujemy 2.5 tryliona (10^{18}) bajtów danych
- Jest to ilość danych zapełniająca 10 milionów dysków blu-ray
- 90% dzisiejszych danych zostało wygenerowanych w ostatnich 2 latach

Szybkość i różnorodność danych

Co minutę:

- 72 godziny filmów jest rejestrowanych na YouTube
- 216 000 postów instagramowych
- 204 000 000 mail

50 GB/s to szacowana w 2018 roku szybkość globalnego ruchu w Internecie

80% danych to video, obrazy oraz dokumenty

90% danych jest nieuporządkowanych, tzn. nie są zgromadzone w bazach danych, ich przeszukiwanie oraz pozyskiwanie z nich wiedzy jest trudne

Wiarygodność danych

- 1 na 3 liderów biznesowych w USA nie ufa informacjom wykorzystywanym do generowania decyzji
- $3.1 * 10^{12}$ $ na rok - to straty gospodarki amerykańskiej z tytułu złej jakości danych

Konieczne jest wykorzystanie systemów sztucznej inteligencji, często autonomicznych,

w analizie danych

i wnioskowaniu na podstawie danych.

Konieczność stosowania metod sztucznej inteligencji w analizie danych

Sztuczna inteligencja osiąga sukcesy w wielu dziedzinach takich jak systemy eksperckie, przetwarzanie języka naturalnego czy rozpoznawanie obrazów. Równocześnie zaczyna się tworzyć świadomość na temat wpływu algorytmów na nasze życie. Brak kontroli nad szeroko pojętą jakością modeli uczenia maszynowego może spowodować, że wyniki generowane przez modele będą bezużyteczne lub wręcz szkodliwe. Jest to szczególnie niebezpieczne, gdy na bazie modeli podejmowane są decyzje ważne społecznie, na przykład dotyczące kar sądowych, dostępu do świadczeń socjalnych czy kredytów.

Zagrożenia wynikające z analizy danych metodami sztucznej inteligencji

Modele o dobrej skuteczności mogą wykorzystywać w predykcji czynniki, które człowiek uznałby za nieetyczne lub niewłaściwe lub przypadkowe. Jednym z głośniejszych przykładów jest sprawa systemu COMPAS (Correctional Offender Management Profiling for Alternative Sanctions).

Latem 2016 roku rozgorzała debata na temat opartego na uczeniu maszynowym narzędzia stosowanego w sądach w całych Stanach Zjednoczonych. Firma Northpointe (obecnie Equivant) stworzyła system, który na bazie wielu różnych czynników przewidywał prawdopodobieństwo tego, czy skazany popełni kolejne przestępstwo w ciągu dwóch lat od wyjścia z więzienia.

Zagrożenia wynikające z analizy danych metodami sztucznej inteligencji

System COMPAS miał uprosić pracę sędziów, uczynić ją bardziej obiektywną i pomóc odpowiednio dobierać środki karne. Osoby o mniejszej skłonności do popełniania kolejnych przestępstw mogłyby wychodzić wcześniej z więzienia, a pozostawaliby w nim przestępcy o większym ryzyku recydywy. Model COMPAS był stosowany w praktyce, wspierając decyzje sędziów i był jednym z przykładów, że uczenie maszynowe może usprawnić funkcjonowanie sądów.

Do czasu, gdy fundacja ProPublica przeprowadziła szerokie badania działania tego systemu i pokazała, że decyzje modelu nie są sprawiedliwe. Za niesprawiedliwość uznano to, że spośród więźniów, którzy nie popełnili ponownie przestępstwa, model mocno przeszacowywał ryzyko dla czarnoskórych więźniów, podczas gdy ryzyko dla białych więźniów, którzy zostali recydywistami, było niedoszacowane. Model nauczył się uprzedzeń rasowych.

Rekomendacje

Przyjęcie przejrzystości zasad projektowania w odniesieniu do sposobu gromadzenia i wykorzystywania danych wejściowych w algorytmach sztucznej inteligencji.

Często wady algorytmów są powodowane faktem, że dane wejściowe nie reprezentują dobrze próby. W ten sposób może być wprowadzona stronniczość do określonych grup ludzi, ich poglądów lub zachowań. Przejrzystość w sposobie gromadzenia danych w algorytmicznych systemach podejmowania decyzji jest niezbędna dla ich wiarygodności.

Rekomendacje

Należy skoncentrować badania nad algorytmami sztucznej inteligencji wyjaśnialnymi. W ten sposób można zwiększyć przejrzystość systemów.

Można zaproponować różne możliwe rozwiązania:

1. budowa przejrzystych modeli (możliwe, że na bazie wiedzy wydobytej z czarnych skrzynek); możliwe jest uczenie równocześnie algorytmu w modelu typu czarna skrzynka i wyjaśnianie lub modyfikacja algorytmu w celu zwiększenia jego przejrzystości,
2. analiza wyuczonych już czarnych skrzynek w celu zrozumienia tego, jak ich decyzje zależą od danych wejściowych oraz diagnostyka modelu,
3. analiza zagadnień związanych z prywatnością danych, sprawiedliwością modeli i niedyskryminacją.

Rekomendacje

Eksperymentowanie.

Podobnie jak w przypadku badań klinicznych nowych leków, systemy sztucznej inteligencji powinny być wielokrotnie testowane i oceniane w dobrze monitorowanych testach przed ich wprowadzeniem na rynek. W takich eksperymentach należy wyraźnie zbadać, czy interakcja między jednostkami i systemami sztucznej inteligencji (np. robotami) spełnia normy bezpieczeństwa i prywatności istot ludzkich. Powinny one również dostarczać jasnego komunikatu na temat tego, jak należy zmodyfikować projekt systemu sztucznej inteligencji w celu spełnienia tych zasad.

Rekomendacje

Opracowanie ram regulacyjnych i etycznych dla odpowiedzialności rozproszonej.

Ramy te powinny zawierać jasne standardy i zalecenia w stosunku do narzuconych zasad odpowiedzialności, które ułatwiają ochronę zarówno użytkownikom, jak i producentom poprzez skuteczne i sprawiedliwe mechanizmy podziału ryzyka.

Rekomendacje

Zwiększenie świadomości społecznej.

Ponieważ algorytmy sztucznej inteligencji wnikają coraz bardziej w nasze życie, powinniśmy być dobrze poinformowani o ich przydatności i potencjalnych zagrożeniach.
W tym celu powinny być zaprojektowane programy edukacyjne i szkoleniowe.
W ten sposób osoby nie tylko będą świadome niebezpieczeństw, ale także zmaksymalizują korzyści z takich systemów.

Rekomendacje

Powinien zostać opracowany spójny kodeks etyczny projektowania i eksploatacji systemów autonomicznych (na przykład na poziomie UE), oparty na wspólnych humanistycznych wartościach. Zasady kodeksu powinny być przestrzegane przez projektantów systemów sztucznej inteligencji, firmy, władze publiczne, organizacje pozarządowe i użytkowników.

Rekomendacje

Zakaz śmiertelnej broni autonomicznej.

Konieczne jest opracowanie zakazu projektowania śmiercionośnej broni autonomicznej i wsparcie w tym względzie odpowiednich inicjatyw Organizacji Narodów Zjednoczonych.

Bibliografia

- Michał Araszkiewicz, Sztuczna inteligencja i prawo do wyjaśnienia, Kwartalnik Trzeci Sektor, Nr 44 (4/2018).
- The ethics of Big Data: Balancing economic benefits and ethical questions of Big Data in the EU policy context, Study of European Economic and Social Committee