

Temat: Etyczne aspekty wykorzystania uczenia maszynowego w projektowaniu produktów cyfrowych mających na celu pozytywną zmianę społeczną

W ramach technologii dobra społecznego, szeroko wykorzystywane mogą być algorytmy uczenia maszynowego. Znajdują one zastosowanie w różnych dziedzinach, od opieki zdrowotnej po edukację i rozrywkę. Uczenie maszynowe i szerzej rozumiana sztuczna inteligencja stanowią jedno z najmocniejszych narzędzi wykorzystywanych przez nowoczesne aplikacje przemysłu 4.0. W tym celu, ich pomysłodawcy powinni kierować się etycznymi zasadami, które wpłyną na pozytywny rozwój społeczny poprzez wykorzystanie funkcjonalności stworzonych aplikacji przez użytkowników.

Uczenie maszynowe to dziedzina sztucznej inteligencji, która umożliwia analizowanie danych i uczenie się na ich podstawie określonych wzorców, w zależności od zastosowanego algorytmu. Dzięki temu można budować systemy, które potrafią rozwiązywać złożone zadania, takie jak rozpoznawanie obrazów, przetwarzanie języka naturalnego czy rekomendacje produktów. Praktycznie każda wyszukiwarka internetowa, z której korzystają codziennie miliony ludzi, korzysta z algorytmów uczenia maszynowego. Ostatnio prawdziwy renesans swojej popularności przeżywają chat-boty, generujące rozwiązania określonych problemów i korzystające ze skomplikowanych modeli językowych np. GPT-4. To tylko część zastosowań wykorzystywanych w dzisiejszym świecie. Jako, że największe organizacje świata ścigają się w celu ulepszania istniejących rozwiązań oraz tworzenia nowych potencjalnych zastosowań algorytmów uczenia maszynowego, pociąga to za sobą ogromną odpowiedzialność za nowo wprowadzane produkty.

Jednym z najistotniejszych problemów, który powoduje brak zaufania do źle dostrojonych systemów wykorzystujących algorytmy uczenia maszynowego, bywa tzw. *bias* („stronniczość”). Innymi słowy jest to błąd, jaki algorytm popełnia na danych dostarczanych do działającego systemu. Fakt, że uczenie poszczególnych klasyfikatorów, dokonuje się na danych, może dyskryminować osoby znajdujące się w jakiejś mniejszości. Ważne jest, by przy konstruowaniu modeli uczenia maszynowego uwzględniać istniejące dysproporcje w danych i starać się nie dyskryminować ludzi ze względu na rasę, płeć, poglądy. Przykładem nieprawidłowości wynikających z *biasu* może być system podejmujący decyzję o wyborze właściwego kandydata na dane stanowisko, sugerujący się czynnikami niepowiązanymi z kwalifikacjami na określony wakat.

Do budowania efektywnie działających modeli uczenia maszynowego potrzebne są dane w dużych ilościach. Najczęściej pochodzą one z istniejących już systemów informatycznych, na bazie działań podejmowanych przez użytkowników tych aplikacji. Przetwarzanie i wykorzystanie tych danych musi odbywać się pod szczególną ostrożnością, zgodnie z funkcjonującą polityką ochrony danych osobowych. Istnieją pewne regulacje prawne (np. RODO) określające prawa użytkowników w przypadku wycieku wrażliwych danych. Jedną z najbardziej znanych amerykańskich korporacji Meta została ukarana grzywną w 276 milionów dolarów w kwietniu 2021 roku po wycieku danych zawierających informacje o 533 milionach użytkowników aplikacji Facebook.

Wraz z rozwojem coraz bardziej skomplikowanych modeli uczenia maszynowego, w szczególności głębokich sieci neuronowych, problemem stało się odpowiednie zrozumienie działania generowanych przez model wyników. W celu zbudowania odpowiedniego zaufania do wykorzystania sztucznej inteligencji w nowo powstałych produktach cyfrowych, powstają różne metody próbujące wyjaśnić ich działanie w sposób generyczny (np. LIME, SHAP). Dzięki tym metodom, potencjalni klienci naszej aplikacji mogą zobaczyć, jak działa dany model i na tej podstawie zbudować zaufanie do wyników jego działania.

Kolejnym z etycznych aspektów tworzenia produktów celowych dla dobra społecznego jest odpowiedzialność. Za wytwarzanie oprogramowania z użyciem modeli uczenia maszynowego odpowiada wiele osób. Od osób zarządzających danym projektem, po programistów i zarząd spółki wprowadzającej produkt na rynek – na każdej z osobna ciąży część odpowiedzialności za działanie

końcowego produktu. Ważne jest, by po działaniu aplikacji niezgodnym z etyką, a w szczególności z prawem, osoby odpowiedzialne za niewłaściwe funkcjonowanie produktu poniosły za nie odpowiedzialność, co może się wiązać z odpowiednią karą lub zadośćuczynieniem.

Wymienione powyżej etyczne aspekty wykorzystania uczenia maszynowego w projektowaniu produktów cyfrowych są kluczowe do tworzenia pozytywnych zmian społecznych. Programiści, osoby zarządzające, właściciele firm informatycznych oraz końcowi użytkownicy aplikacji muszą wspólnie współdziałać w celu stworzenia rozwiązań, które przyniosą korzyści każdej ze stron i przyczynią się do poprawy jakości życia ludzi. Postępując zgodnie z zasadami etycznymi i będąc odpowiedzialnym za powstający produkt dla dobra społecznego, zmniejsza się ryzyko niepowodzeń biznesowych w przypadku odniesienia sukcesu przez wprowadzenie aplikacji na rynek oraz problemów prawnych, a także zwiększa się potencjalną liczbę odbiorców.

Michał Ciesielski 145325