

ZBUDOWAĆ WIĘZ ZE SZTUCZNĄ INTELIGENCJĄ

O (NIE)LUDZKIEJ TWARZY ALGORYTMÓW OPOWIADA
DR HAB. INŻ. PRZEMYSŁAW BIECEK, PROFESOR UCZELNI



**Wydział Matematyki
i Nauk Informatycznych**

POLITECHNIKA WARSZAWSKA

Jak zaczęła się Pana przygoda z nauką?

Na Wydziale Matematyki na Politechnice Wrocławskiej mieliśmy bardzo charyzmatycznego profesora, który prowadził seminarium z genetyki – genetyki dla statystyków. I sposób, w jaki on pokazywał, jak analiza danych może wpłynąć na bardzo ciekawe problemy, np. jak badać twardość skorupki jajka lub też ilość mleka pozyskanego od krowy, zainspirował mnie. Takie aplikacyjne problemy wydały mi się bardzo interesujące. Więc tak zaczęło się moje zainteresowanie badaniami.

Czy w związku z Pana zainteresowaniem badaniami aplikacyjnymi kariera komercyjna nie była dla Pana równie atrakcyjna?

Nie dzielę badań na komercyjne i czysto akademickie. Dlatego też od lat pracuję w obu miejscach, mam doświadczenie w pracy czysto korporacyjnej i na uczelniach badawczych. Więc mnie najczęściej interesowało przecięcie tych dwóch światów, ostatnio nazywane badaniami transdyscyplinarnymi. Badania aplikacyjne są gdzieś pomiędzy i szczególnie w obszarze sztucznej inteligencji obserwuję teraz na świecie, że to jest dosyć typowe podejście do tego problemu. Mamy firmy, które są spin-offami z uczelni – uczelnie cały czas mają prawa własności do tej własności intelektualnej. Więc to się zrodziło na uczelni, ale teraz jest normalną komercyjną działalnością. Ale są też firmy, które outsourcują jakieś badania na uczelni. Ta współpraca jest coraz bardziej ścisła.

Czym zajmuje się Pan w swojej pracy naukowej?

Pozwolę sobie powiedzieć o dwóch rzeczach. Jest jedna rzecz, z której jestem bardzo dumny, ale z jakiegoś powodu ona raczej nie porywa moich rozmówców. Jesteśmy świadkami olbrzymiej zmiany w społeczeństwie. Zmiany, gdzie umiejętność tworzenia informacji będzie kluczowa i tak jak wcześniej było ważne piśmiennictwo, umiejętność czytania czy też dodawania, to obecnie umiejętność analizy danych będzie po prostu kluczowa. I żeby zapoczątkować działania, założyłem Fundację Naukową Smarter Poland, ramach której tworzymy podręczniki, warsztaty, różne działa-

nia edukacyjne. Mamy serię komiksów o Becie i Bicie, to jest rodzeństwo. Beta jest matematyczką, a Bit jest informatykiem. Charakterologicznie są różni: Beta jest bardziej spokojna, zrównoważona. Bit jest taki bardziej do przodu, nie myśli specjalnie o tym, czy to robi dobrze, czy nie, po prostu pędzi. Ale razem rozwiązują problemy związane z analizą danych. Tworzymy materiały dla dzieci zarówno ze szkoły podstawowej, jak i średniej oraz dla ludzi już starszych, wierząc, że powinniśmy jako analitycy danych popularyzować to spojrzenie, bo to musi być częścią umiejętności globalnych, tak samo jak czytanie. To się naprawdę dzieje. Ja jestem zaskoczony, jak dużo się już pojawia w podręcznikach. Na przykład w szkole podstawowej w podręcznikach są wykresy, których kiedyś nie było, są analizy, których kiedyś nie było. To się dzieje, ale tego powinno się dzieć więcej.

Zatem czego powinniśmy się nauczyć?

Krok po kroku. Na początek nauczmy się czytać wykresy. Ale naprawdę myślę, że to, co jest szalenie ważne, to odczytywanie relacji: co to dokładnie znaczy, co można badać, czego nie można badać. Myślę, że naukowcy często bardzo krytycznie podchodzą do zależności, które widzą w swoich danych i z dużą pokorą te zależności interpretują. Ale poza sferą nauki tej pokory nie widać i ona prowadzi do wielu błędnych wniosków. Tak, żebyśmy trochę lepiej potrafili wyciągać wnioski z informacji przeczytanej

w gazecie czy w innych miejscach. To są prozaiczne kwestie typu dociekanie, czy są źródła podane. Mamy informację i ktoś może być świetny w odczytywaniu wykresów, ale nie jest krytyczny w analizie źródeł, nie zauważy, że na wykresie są prezentowane złe dane i wyniki są zupełnie bezsensowne. Nie ma znaczenia, czy potrafi czytać, czy nie potrafi czytać. Ja muszę dosyć dobrze zrozumieć cały proces analizy danych, rozumieć ważność pojedynczych elementów tego procesu, żeby móc oceniać, które wnioski są uprawnione, a które nie. Sama nauka o danych powstała relatywnie niedawno – kilkanaście lat temu, może dwadzieścia czy trzydzieści, więc my jeszcze nie wiemy, jak dobrze uprawiać tę naukę. Ale na pewno ona będzie ważna.

W jaki sposób to wpłynie na codzienność, na jakość życia ludzi?

Świat dotąd mogliśmy rozumieć przez historię i wiele książek komunikowało rzeczywistość przez biografie, historie, fabuły, ale coraz więcej mamy dostępnych dużych zbiorów danych i możemy część tego świata rozumieć przez dane. Komunikacja danych wymaga innych narzędzi niż historie fabularne. Cały czas nie wiadomo jakich, ale jednak to są zupełnie innego rodzaju historie. Myślę, że zaspokajanie naturalnej ciekawości tego, co się dzieje na świecie będzie możliwe dzięki tym dodatkowym narzędziom. Wyobrażam sobie, że ludzie, którzy często słuchają muzyki, mają lepiej wyrobiony słuch i mają

przyjemność z odsłuchiwania koncertu filharmonii. Osoby, które mają wyrobione zmysł plastyczny odczuwają przyjemność z chodzenia do galerii. Tak samo osoby, które wykształcą sobie umiejętność czytania danych będą odczuwały taką przyjemność z lepszego zrozumienia danych, które dookoła obserwują.

Edukowanie w zakresie analizy danych jest działalnością obok Pana pracy naukowej. Czym zajmuje się Pan naukowo?

Mnie interesuje sztuczna inteligencja, szczególnie aspekty współdziałania. Często myślimy o tej sztucznej inteligencji na zasadzie: robot wykonuje nasze rozkazy, przynosi nam kawę albo coś podobnego. Ale widzimy też, że ta przyszłość związana z robotami się nie sprawdziła. Dużo większe oddziaływanie na nasze życie mają automatyczne algorytmy, które decydują o tym, jakie reklamy widzimy na portalach, która wiedza, o której godzinie nas obudzić w danym dniu tygodnia i które nam mówią wiele innych rzeczy.

Na przykład co powinniśmy kupić.

Tak. Algorytmy są kluczowe. Te algorytmy są automatyczne, a to, w jaki sposób powinny współdziałać z człowiekiem jest zagadką. To mnie bardzo interesuje w tej chwili. Chodzi o to, żeby ludzie cały czas byli w stanie nie być tylko biernymi odbiorcami treści, być może nawet niesprawdzonych. Chciałbym, żeby człowiek był w stanie wejść w interakcję, w dyskurs,

w rozmowę z tą skrzynką i zrozumieć za i przeciw. Nie chciałbym, żeby maszyny mówiły człowiekowi, co ma robić, ale żeby były one wirtualnymi asystentami, którzy przedstawiają różne perspektywy, a człowiek rozumiejąc, jakie są za i przeciw, podejmie najlepszą decyzję.

Jak człowiek powinien komunikować się z maszyną?

To jest trudne, bo ludzie mają swój język, którym się posługują. Maszyny mają swój język wnioskowania, a to są bardzo różne języki. Więc to jest wyzwanie: jak zbudować interfejs pomiędzy człowiekiem a maszyną. Zauważa je cały świat. Obecnie w Unii Europejskiej też toczy się olbrzymia dyskusja na temat odpowiedzialności sztucznej inteligencji. Ta inteligencja powinna nie tylko być efektywna, ale też musi być zgodna z jakimiś wartościami. Jak te wartości opisać? Czy w ogóle da się opisać? Czy da się zweryfikować? To są niebanalne rzeczy, ale też bardzo interesujące.

Człowiek może mieć poczucie przedmiotowego traktowania przez maszynę.

Przedmiotowe traktowanie to po pierwsze. Weźmy jako przykład rozmowy z automatami na infolinii. Ja się czuję gorzej, że nie rozmawiam z człowiekiem tylko z algorytmem, który w ogóle mnie nie rozumie. Ja do niego mówię, a on nie rozumie,

„Jesteśmy świadkami olbrzymiej zmiany w społeczeństwie. [...] Tak jak wcześniej było ważne piśmiennictwo, umiejętność czytania czy też dodawania, to obecnie umiejętność analizy danych będzie po prostu kluczowa”.

pyta się mnie trzy razy, czy dobrze rozumiał „dziewięć”. Oczywiście są wyraźne powody ekonomiczne, dla których te automaty są wprowadzane, ale z punktu widzenia naszego poczucia wartości, to wydaje mi się, że nie odbieramy tego najlepiej. To źle, że widzimy postęp technologiczny jako coś przeciwko nam, coś, co obniża naszą jakość życia zamiast ją podnosić. Człowiek musi być w środku – nie możemy go krzywdzić, tylko powinniśmy go wspierać.

Proszę podać przykład z Pana pracy, który może pokazać, że sztuczna inteligencja wspiera człowieka, że to on jest w środku.

Mój Zespół podjął się inicjatywy wpisującej się w walkę z pandemią COVID-19. W ramach tej inicjatywy wzięliśmy na warsztat kilka problemów. Wreszcie nauka jest potrzebna tu i teraz, nie na zasadzie – za 50 lat, tylko dzisiaj. Jednym z elementów inicjatywy jest rozpoznawanie zmian na zdjęciach płuc – wspieranie radiologów pod kątem weryfikacji, jakie to są zmiany, czy to jest włóknienie, czy to pozostałość po gruźlicy, czy to jest nowotwór,

czy to jest zmiana tymczasowa. Inne działanie, które jest jeszcze bardziej namacalne, to model predykcyjny, który jest w stanie oceniać na podstawie różnych charakterystyk, jakie są szanse, że dana osoba przejdzie w sposób trudny COVID. Zrobiliśmy model, który funkcjonuje na zasadzie strony internetowej. Pokazywaliśmy to nawet w telewizji i pisały o nim gazety. Więc to dotarło do różnych grup osób. Analizując informacje, takie jak wiek, płeć, choroby towarzyszące, model oceniał, jakie jest indywidualne spersonalizowane ryzyko ciężkiego przejścia choroby. Są bardzo ciekawe aplikacje naszego kalkulatora. Wiem, jaka grupa jest najbardziej narażona na ciężkie przejście choroby i gdy mam pewne możliwości interwencji, na przykład szczepienia czy pulsoksymetri, to mogę podejmować uzasadnione działania. Powiedzmy, że mam 100 000 pulsoksymetrów i komu mam je dać? Model mi podpowie, dla kogo będą najbardziej użyteczne.

Czy to rozwiązanie jest wykorzystywane teraz w trakcie pandemii?

Ono jest wykorzystywane, ale trzeba pamiętać, że te końcowe decyzje nie są oparte o jedno rozwiązanie.

„Trzeba patrzeć na społeczeństwo pod kątem tego, jak technologia na nie oddziałuje. Nie sztuką jest tworzyć narzędzia, ale też jako twórcy narzędzi powinniśmy być odpowiedzialni też za to, co te narzędzia robią, jak zmieniają świat”.

Uczestniczymy z innymi grupami w cyklicznych spotkaniach z Ministerstwem Zdrowia i z Polskim Zakładem Higieny. To są spotkania, gdzie dyskutujemy co tydzień ważne wyzwania pandemii. Także przewidywanie konsekwencji wdrażania technologii jest niebanalne. Jedną sprawą jest wymyśleć, jakie jest ryzyko, a druga sprawa – dużo trudniejsza – to jest przewidzenie konsekwencji wdrożenia tej innowacji. Na pewno nie mogę sobie przypisać jakiejś wielkiej zasługi, ale mamy z innymi grupami gorącą linię z PZH i te wyniki są wykorzystywane. Modele analizują rozwój pandemii i starają się przewidzieć liczbę zgonów. Jeżeli ja dzisiaj wiem, kto jest chory, to doskonale wiem, kto umrze za tydzień. To jest bardzo łatwo powiedzieć, bo my już wiemy, kto zachoruje. Nie wiem, czy będzie to Kowalski czy Wiśniewski, ale wiem, że ktoś o danym profilu.

Przykład użyteczności nauki i do tego w słusznej sprawie.

Nie mogę niestety, jak lekarze, powiedzieć, że uratowałem jedno życie, bo tego człowiek nie wie, ale mam naprawdę przekonanie, że pośrednio, wspierając tych lekarzy albo decydentów, mamy deltę pozytywną. Czas nie jest zmarnowany.

To był jeden z przykładów zastosowania sztucznej inteligencji. Gdzie jeszcze algorytmy mogą być dla nas użyteczne?

Akurat COVID to jest coś rzeczywiście nowego, ale mamy też bardzo wiele modeli wytworzonych dla onkologii, gdzie staramy się określić ryzyko nieskuteczności terapii w zależności od biomarkerów genetycznych, więc w obszarze służby zdrowia aplikacji mamy dosyć dużo. Drugi obszar, gdzie modele predykcyjne są między nami, są wśród nas i które są bardzo widoczne, to jest obszar ryzyka w finansach, czyli decyzja, kto dostanie kredyt, a kto nie. Brzmi bardzo prozaicznie – zwykły kredyt, ale czasami to jest czyjeś być albo nie być. Dostępność kredytu to bardzo ważny mechanizm oddziaływania na gospodarkę. Dużo się mówi o tym, czy jeżeli decyzja kredytowa dyskryminuje ze względu na wiek albo płeć, albo rasę, to czy to jest w porządku. W Stanach Zjednoczonych jest to istotny temat. W Polsce dyskryminacja rasowa raczej nie istnieje, bo u nas właściwie nie ma wielu ras, ale mamy cały czas kwestię wieku. O osobach starszych dużo więcej wiemy, więc to ryzyko jest niewielkie albo ktoś ma pracę i spłacał wszystkie kredyty przez 20 lat i wiemy, że spłacił albo nie spłacał i nie spłaci. Grupą, która jest często dyskryminowana, jest grupa osób młodych, gdzie my nie mamy historii i nie wiemy właściwie, czy im dać kredyt, czy im nie dać kredytu. Więc często ta grupa jest dyskryminowana i nie dostaje kredytów, chociaż ich potrzebuje, bo też najczęściej niewiele ma.

Czy decyzja algorytmu nie jest weryfikowana przez bankowca?

Oczy bankowca patrzą albo nie patrzą. Coraz częściej decyzje są zautomatyzowane dla tych małych kredytów. Jeżeli jest to kredyt na 10 milionów, to na pewno linia decyzyjna banku jest bardzo długa, ale w przypadku małych kredytów, które często dotyczą osób, które naprawdę są w potrzebie, wiele decyzji jest zautomatyzowanych. Tam nie ma żadnych oczu bankowca – tam algorytmy podejmują decyzje automatycznie.

Algorytmy się mylą?

Trzeba patrzeć na społeczeństwo pod kątem tego, jak technologia na nie oddziałuje. Nie sztuką jest tworzyć narzędzia, ale też jako twórcy narzędzi powinniśmy być odpowiedzialni też za to, co te narzędzia robią, jak zmieniają świat. Dzisiaj nie mamy dobrych rozwiązań, które są w stanie badać te sekwencje. Sięgając trochę wstecz, gdy były pierwsze maszyny wprowadzane w rewolucji przemysłowej, gdy maszyna oderwała rękę pracownikowi, to najczęściej była to wina pracownika. Dlatego, że maszyna nie pracowała przez tydzień, trzeba było ją naprawić. Stwierdzenia, że to jest wina pracodawcy, że nie zagwarantował warunków pracy bezpiecznych dla pracownika, wymagało zmiany w społeczeństwie. Po pierwsze, producent maszyny musi zadbać o warunki pracy, a po drugie – pracodawca. Dzisiaj, gdy zapytamy o to, kto jest winny za złe decyzje automatyczne, nie usłyszy-

my odpowiedzi – myślę, że nie ma takiej odpowiedzi.

Nie ma odpowiedzialności?

To jest wyzwanie, którego rozwiązanie musi znaleźć interdyscyplinarna grupa ludzi. To nie jest temat tylko dla prawników, bo oni często też nie wiedzą, co jest możliwe technicznie. To nie jest temat tylko dla inżynierów, bo oni nie widzą szerszego kontekstu. Za mało mamy dobrych narzędzi analizujących konsekwencje nowych technologii, szczególnie w obszarze zdrowia i obszarze linii kredytowych. Tworzymy narzędzia, które pozwalają analizować konsekwencje. Jedne we współpracy z lekarzami, a drugie we współpracy z ludźmi z finansów. Żeby wejść w te dziedziny, trzeba mieć lata doświadczeń. My nie mamy kompetencji, żeby wypowiadać się o mikroekonomii czy makroekonomii, ale możemy pomóc w analizie narzędzi dla ekonomistów.

Często my jako społeczeństwo nie mamy świadomości zakresu, w jakim oddziałuje na nas sztuczna inteligencja.

Myślę, że algorytmy naprawdę oddziałują na nas w sposób dla nas niewyobrażalny. Nikt sobie nie wyobraża. Tak jak kryzys finansowy na rynku mieszkaniowym w USA był zaskoczeniem olbrzymim, tak samo dzisiaj ryzyka związane z algorytmami są niedoceniane. I w najbliższych latach musimy jakoś rozpracować, co z nimi zrobić, w jaki sposób umówić się, co wolno, czego nie wolno, co jest dobre

dla nas, co jest złe. Na pewno w technologii jest olbrzymi potencjał, ale też w niektórych obszarach może być ryzyko i musimy też te ryzyka wykrywać, zarządzać nimi i eliminować je.

Czyli wyzwaniem na dziś jest etyka współdziałania człowieka i algorytmu?

Bezpieczeństwo i etyka. Etykę czasami za bardzo kojarzy się z jakimiś przekonaniem, a tutaj chodzi o nasze czyste bezpieczeństwo. O bardzo prozaiczne rzeczy, które są nam dużo bliższe niż dywagacje światopoglądowe. To bardzo pragmatyczne rzeczy. Na przykład czy coś przekłada się na nasze lepsze zdrowie za pięć lat czy na nasze bezpieczeństwo finansowe. Książka, którą bardzo cenię, czytałem ją kilka lat temu, ale wracam do niej bardzo często, to książka Cathy O’Neil „Broń matematycznej zagłady”. Ona opisuje bardzo fajnie przykłady różnych technologii, które nazywa bronią matematycznej zagłady. Pokazuje, jak algorytmy matematyczne mogą krzywdzić różne osoby. Oczywiście to nigdy nie jest tak, że matematyka krzywdzi. To złe użycie algorytmów może doprowadzić do krzywdy. I to się dzieje. Opisuje sytuację, kiedy świetni nauczyciele są zwalniani z pracy, bo system omyłkowo powiedział, że oni nie są świetni. Nikt nie wie, dlaczego, ale nauczyciel zostaje zwolniony.

Dlaczego temat bezpieczeństwa algorytmów dopiero teraz znajduje swoje miejsce w rozważaniach naukowych?

Po pierwsze kiedyś nie mieliśmy tak dużych zbiorów danych, a bez dużych zbiorów danych nie można budować bardzo złożonych modeli. Kiedyś nie mieliśmy też tak dużej mocy obliczeniowych. Postęp, który się wydarzył w obszarze sprzętu do budowania modeli, jest olbrzymi, niewyobrażalny. Teraz to już jest wyścig, który się obserwuje w postępie co kilka miesięcy, a nie co kilka lat. Pokazuje się, że miliardy różnych forów z całego Facebooka czy Twittera są przetwarzane, żeby zbudować jakiś model. Pięć lat temu zasób danych był niebotycznie mniejszy i moc obliczeniowa była niebotycznie mniejsza, więc nie było modeli. Dzisiaj mamy modele, a jeszcze nie wiemy, jak je walidować i weryfikować, więc dopiero dzisiaj docieramy do miejsca, gdzie uświadamiamy sobie, jak groźną technologię tworzymy. Jesteśmy gotowi, żeby zaproponować rozwiązania, żeby ta technologia była bezpieczniejsza.

Zatem na początku człowiek zachłusnął się możliwościami data science, a teraz przyszedł moment na refleksję, co z tym dalej, czy to na pewno idzie w dobrą stronę?

Technologia ma taką bardzo fajną Krzywą Hype'u Gartnera. Ta krzywa rośnie na początku, gdy dana technologia ma olbrzymi entuzjazm, później jest to szczytowe napompowane

„Dzisiaj lekarz potrzebuje trzech godzin, żeby przeanalizować dokumentację medyczną i podjąć decyzje, a wspomagając go technologią ułatwimy mu pracę tak, że analiza będzie zajmowała mu pół godziny”.

oczekiwanie, że nam to wszystkie problemy rozwiąże, a dalej każda technologia przechodzi w rozczarowanie, które się musi wydarzyć. Wiadomo, później musi się odbić i musi wejść na poziom produktywności, kiedy my wiemy, jak tego używać bezpiecznie. Sztuczna inteligencja na Krzywej Hype'u Gartnera przechodzi z fazy napompowanych oczekiwań do fazy obawy. Ale to nie jest tak, że będzie źle. Musimy wyjść z fazy obawy i wejść w poziom produkcji, gdzie będziemy potrafili używać tej technologii w sposób bezpieczny.

Jak już nauczymy się bezpiecznie wykorzystywać algorytmy, to gdzie one znajdą zastosowanie?

Algorytm ma duży potencjał. Weźmy przykład ochrony zdrowia. Mamy dużą liczbę nadmiarowych zgonów, którą często przypisuje się mniejszej dostępności specjalistów, ponieważ jest ich mało. Nasi specjaliści pracują w Norwegii czy w Niemczech. Aby tak rzadki zasób, jakim jest praca specjalisty wykorzystać, musimy go wspomóc technologią. Dzięki sztucznej inteligencji będziemy w stanie zwiększyć produktywność człowie-

ka, szczególnie eksperta, który musi podjąć ważne decyzje. Dla przykładu: dzisiaj lekarz potrzebuje trzech godzin, żeby przeanalizować dokumentację medyczną i podjąć decyzje, a wspomagając go technologią ułatwimy mu pracę tak, że analiza będzie zajmowała mu pół godziny. Być może większa produktywność lekarzy przełoży się na dalszą poprawę sytuacji zdrowotnej całego społeczeństwa. Trudno mi jest przewidywać tak szeroko, ale liczę na to, że algorytmy pomogą zwiększyć produktywność człowieka i to, co dzisiaj on robi w ciągu dwóch godzin, będzie mógł zrobić w ciągu pięciu minut.

Czy pojawia się problem braku zaufania do algorytmów i do decyzji podejmowanych przez algorytmy? Czy lekarze, z którymi Pan współpracuje mają wątpliwości co do tego, czy można zaufać decyzjom algorytmu?

Oczywiście, ale wydaje mi się, że tak jest z każdą technologią. Gdy zaczęły jeździć pociągi, to przed pociągiem musiał iść człowiek, który machał flagą: „uwaga, jedzie pociąg”. Ludzie bali się, że krowy nie będą dawały mleka. Zawsze przed nową technologią jest strach i to jest normalna rzecz. Tak samo jest z nieufnością. Trzeba pokazać korzyść i po jakimś czasie, może za pokolenie, a może za miesiąc, część z tych rozwiązań się przyjmie. Każda technologia budzi niepokój. Patrząc na inne technologie, też widzimy, ile trwa niepokój i jak go rozładowywać. Dzisiaj mamy pociągi Shinkansen, które pędzą 500 km/h i nikt nie martwi się, że krowa będzie

dawała gorsze mleko, ale to trwało. Zawsze trwa.

Rozmowę przeprowadziła:
Magda Matysiak
(DBA CZliTT PW)