

Kraków, 16 czerwca 2021r.

dr hab. inż. Michał Bereta, Prof. PK
Prodzikan ds. Nauki Wydziału Informatyki i Telekomunikacji
Politechnika Krakowska
mbereta@pk.edu.pl

Recenzja Rozprawy Doktorskiej

Tytuł

Definiowanie realne w maszynowym rozpoznawaniu wzorców

Praca doktorska napisana na seminarium logiki filozoficznej, Lublin 2021

Dyscyplina

Filozofia

Autor

ks. Jakub Sławomir Płonka

Miejsce

Katolicki Uniwersytet Lubelski Jana Pawła II, Wydział Filozofii, Instytut Filozofii

Promotorzy

dr hab. Marek Lechniak, prof. KUL
oraz dr hab. Paweł Karczmarek, prof. Politechniki Lubelskiej

Cel badawczy rozprawy i jego znaczenie

Recenzowana praca doktorska jest pracą w dyscyplinie filozofii, jednak zawiera w sobie bardzo wiele wątków związanych z działem informatyki, zwanym uczeniem maszynowym. Można powiedzieć, że praca jest w swoim charakterze interdyscyplinarna, co bardzo mocno podnosi jej wartość.

Jak sam Autor określa: "*Głównym celem niniejszej pracy jest uzasadnienie twierdzenia: maszynowe rozpoznawanie wzorców w aspekcie filozoficznym jest przede wszystkim definiowaniem realnym. Główny wątek wyznacza też temat pracy, jest nim definiowanie realne w maszynowym rozpoznawaniu wzorców.*"

Aby uzasadnić postawioną przez siebie tezę, Autor przedstawia przede wszystkim podstawowe informacje o uczeniu maszynowym (ang. *machine learning*), które jest najprężniej rozwijającą się dziedziną sztucznej inteligencji. Bardzo dokładnie omówione są filozoficzne idee potrzebne do uzasadnienia tezy pracy. Przedstawiony jest historyczny i krytyczny przegląd idei definicji w filozofii, w szczególności definicji realnych. Szeroko omówione zostały sposoby rozumienia substancji i istoty w różnych nurtach myśli filozoficznej. Filozoficzne idee przedstawione w pracy obejmują wiele epok i nurtów, począwszy od starożytnych myśli Arystotelesa, poprzez idee średniowieczne i nowożytne, aż



do osiągnięć filozofii współczesnej. Autor w sposób bardzo szczegółowy bada w swojej pracy, które idee, i w jakim zakresie, mogą dostarczyć intelektualnych narzędzi mogących wspomóc nas w trudnym i niejednoznacznym procesie interpretacji efektów działania algorytmów uczenia maszynowego. Można powiedzieć, że dziedzina uczenia maszynowego pędzi do przodu, ale w niewielkim stopniu zastanawia się nad sobą. Wynika to niewątpliwie z wielkiego nacisku na użyteczność praktyczną opracowywanych algorytmów uczenia maszynowego. W zamyśle swoich twórców, głównie programistów-inżynierów, mają one przede wszystkim działać. Oczywiście, chęć zrozumienia źródeł sukcesów uczenia maszynowego towarzyszy jego twórcom, lecz często ma ona charakter wyłącznie utylitarny, wynikający z chęci dalszego ulepszania technik obliczeniowych. W ogólności brak jest chęci głębszej, filozoficznej właśnie zadumy nad najnowszymi osiągnięciami techniki. Recenzowana praca stara się w pewnym stopniu uzupełnić ten brak.

Autor podjął się w mojej opinii zadania bardzo ambitnego. Połączenie obu wyżej wymienionych dziedzin wiedzy, uczenia maszynowego i filozofii, w postawionej przez siebie tezie, pozwoliło Autorowi na zmierzenie się z tematem ciekawym, naukowo wartościowym ale i trudnym. Od razu warto tu zaznaczyć, że Autor zdecydowanie poradził sobie z postawionym sobie zadaniem. Tematyka badawcza, która została przedstawiona w pracy, jest w mojej opinii unikalna, co wynika właśnie z próby intelektualnych poszukiwań wspólnych elementów odwiecznych myśli filozoficznej oraz osiągnięć uczenia maszynowego, które rozwija się prężnie, ale jest dziedziną młodą, zaledwie kilkudziesięcioletnią. Nie istnieje wiele opracowań, które podejmowałyby tę tematykę. Widać to również w recenzowanej pracy, w której Autor odnosi się wprawdzie do nielicznych (z konieczności) prac, które podejmują tę tematykę (np. prace Watanabego), jednak w rzeczywistości wyruszając na tę intelektualną wyprawę zmuszony był samodzielnie wypracować spójny pogląd pozwalający dopasować odpowiednie myśli i idee filozofii do potrzeb interpretacyjnych osiągnięć uczenia maszynowego. Przyznać należy, że w dużej mierze udało się to Autorowi osiągnąć. Biorąc pod uwagę, że najbardziej interesujące opracowania tego typu (prace Watanabego) mają już swoje lata, tym bardziej warto docenić istotność powstałej pracy.

Podsumowując, postawiony w pracy problem badawczy jest nie tylko unikalny w swojej interdyscyplinarności, ale również ważny, zwłaszcza w świetle szybko rozwijających się technik uczenia maszynowego, które, czy tego chcemy czy nie, zmieniają świat. Filozofia, która ma za swój cel zrozumienie świata, nie może ignorować tego postępu, jeśli nie chce się narażać na zarzut, że nie opisuje już świata aktualnego. W tym kontekście znaczenie recenzowanej pracy jest nie do przecenienia.

Struktura pracy

Praca składa się ze wstępu, sześciu rozdziałów oraz zakończenia. Zawiera również wykaz rysunków oraz bibliografię. We wstępie przedstawiono cele badawcze i główną tezę pracy, przedstawiono strukturę pracy. W rozdziale pierwszym przedstawiono główne zagadnienia związane z uczeniem maszynowym, szerzej ze sztuczną inteligencją i problemy filozoficzne z nią związane. Przedstawiono klasyczne dylematy i rozważania, jak np. test Turinga. Wykazano również, że prac filozoficznych związanych z tą dyscypliną jest stosunkowo niewiele, co dodatkowo motywuje powstanie recenzowanej pracy. W rozdziale drugim przedstawiono zagadnienie definicji realnych. Przedstawiono bardzo szeroki zakres myśli filozoficznych, nie zawsze zgodnych ze sobą. W rozdziale

trzecim przedstawiono różne rodzaje rozumienia kategorii istoty. Uzasadniono, dlaczego definiowanie istotowe zajmuje szczególnie ważną rolę w podjętej tematyce badawczej. Uwzględniono tu imponująco wiele różnych myśli, od arystotelesowskiej, poprzez logikę modalną aż po elementy fenomenologii. W rozdziale czwartym przedstawiono główne etapy związane z opracowywaniem algorytmów uczenia maszynowego, takie jak ekstrakcja i selekcja cech, redukcja wymiarowości, wybór modelu klasyfikacyjnego, jego ocena. W rozdziale piątym przedstawiono dwa konkretne problemy klasyfikacji biometrycznej, rozwiązywane przez algorytmy uczenia maszynowego.

Rozdział szósty to zdecydowanie najciekawszy i najbardziej wartościowy rozdział całej pracy. Wątki i pojęcia rozwijane w poprzednich rozdziałach w końcu spotykają się i Autor jest w stanie przedstawić swoją interpretację definiowania realnego w uczeniu maszynowym. Jest ono przekonujące i wystarczająco dobrze uzasadnia postawioną w pracy tezę.

Całość pracy podsumowana jest w zakończeniu, które stanowi zwarte podsumowanie tez i osiągnięć badawczych zawartych w pracy.

Bibliografia składa się z 178 dobrze dobranych pozycji. Praca jest obszerna, składa się z 234 stron.

Wiedza i wkład Autora, oryginalność pracy

Autor może pochwalić się unikalną, dużą wiedzą, która łączy w sobie dogłębną wiedzę filozoficzną z dokładnym i bardzo aktualnym rozeznaniem w metodach i osiągnięciach uczenia maszynowego. Przedstawiając myśli i idee filozoficzne, Autor nie skupia się jedynie na tych, które najbardziej pasują do jego tezy. Wprost przeciwnie, dostajemy w pracy wyczerpujący przekrój historyczny różnych ujęć problemów definicji, definiowania realnego, istotowego. Idee te często są ze sobą sprzeczne, niekiedy pokrywają się częściowo, czasami w zasadzie prezentują to samo podejście, jednak w wyniku zastosowania różnego aparatu pojęciowego, zdają się nie zgadzać ze sobą. Autor bardzo dobrze wyczuwa te niuanse, wykrywa istotne, wspólne podłoża omawianych idei, potrafi krytycznie oanieść się do wielu z nich.

Również jeśli chodzi o zaprezentowaną wiedzę Autora z zakresu uczenia maszynowego, należy podkreślić, że jest ona imponująca, zwłaszcza jeśli weźmiemy pod uwagę, że prezentowana jest w pracy w gruncie rzeczy z filozofii. Autor przedstawia nie tylko klasyczne podejścia do uczenia maszynowego, ale również jest świadomy aktualnych osiągnięć uczenia głębokiego (ang. *deep learning*), które jest dominującym podejściem w ostatnich latach. Autor dostrzega, że uczenie głębokie jest w istotny sposób nowym wyzwaniem również dla analizy filozoficznej. Potrafi uzasadnić to szczegółami technicznymi, które wskazują na jego dobre zrozumienie podstaw funkcjonowania uczenia głębokiego.

Praca tryska wielką liczbą odnośników, dodatkowych informacji, zawartych zwłaszcza w przypisach. Duża liczba przypisów nie męczy jednak. Przeciwnie, wyraźnie daje odczuć, że wiedza Autora jest szeroka, a Autor musi wręcz sam się ograniczać w jej przedstawianiu.

Wkład Autora w rozwój wiedzy jest w mojej ocenie istotny. Warto podkreślić, że uczenie maszynowe rozwija się niesamowicie szybko ostatnimi laty. Zakres zmian w skali światowej, jaki generuje,

powoduje, że nie można go ignorować. Rozwój uczenia maszynowego nie może być ignorowany w naukach technicznych, biznesie, bezpieczeństwie finansowym i militarnym państw. Wpływa na życie codzienne milionów ludzi. Wykorzystywane jest w automatyzacji procesów podejmowania decyzji przez instytucje, co ma wpływ na życiowe decyzje i możliwości poszczególnych ludzi. Wykorzystywane jest do automatyzacji możliwości bojowych sprzętu wojskowego, co wiąże się z niebezpieczeństwem łatwiejszego zwolnienia się z odpowiedzialności moralnej za wojskowe działania operacyjne. Wszystko to, i jeszcze wiele więcej, jest wynikiem coraz większego wykorzystania praktycznego metod uczenia maszynowego. Wiąże się to z koniecznością wielu analiz i badań mających na celu określenie czegoś, co można nazwać moralnością sztucznej inteligencji. Recenzowana praca w oczywisty sposób nie porusza wszystkich tych zagadnień, nie mierzy się ze wszystkim problemami. Stanowi jednak istotną podwalinę pod potencjalne przyszłe prace mogące powstawać na przecięciu filozofii, etyki oraz metod sztucznej inteligencji. Wydaje się uzasadnione twierdzenie, że rozważając wiele powyższych problemów, podstawowym wymogiem i koniecznym narzędziem będzie zrozumienie, jak działają metody uczenia maszynowego, jakimi narzędziami intelektualnym, wypracowanymi przez wieki rozwoju myśli filozoficznej, możemy je badać. Można stwierdzić, że recenzowana praca jest jednym z pierwszych podłoży pod przyszłe badania w tym zakresie. Można powiedzieć więcej: po przeczytaniu recenzowanej pracy i biorąc pod uwagę unikalną wiedzę Autora łączącą tak odmienne od siebie dyscypliny, chciałoby się powiedzieć "więcej!". Być może w przyszłości powstaną nowe prace Autora, które podejmą trudne zagadnienia filozoficzne i moralne związane z rozwojem sztucznej inteligencji. Biorąc pod uwagę to, co zaprezentował Autor w recenzowanej pracy, jest on jak najbardziej predysponowany do dalszych badań w tym zakresie.

Poprawność językowa i strukturalna pracy

Praca ma dobrze zaplanowaną strukturę, odpowiednią do przedstawianej treści. Układ rozdziałów jest przemyślany i w gruncie rzeczy odpowiedni. Zastanawiać się jednak można, czy niektóre rozdziały, zwłaszcza drugi i trzeci, nie są zbyt długie i zbyt szczegółowe. Wiele informacji w nich zawartych jest powtórzonych w rozdziale szóstym. Można by się zastanowić, czy nie bardziej zasadne byłoby przedstawić najpierw zawartość związaną z uczeniem maszynowym (rozdziały czwarty i piąty) a następnie podczas analizy różnych idei definiowania istotowego, jak również alternatywnych, przedstawiać je od razu w kontekście i w odniesieniu do uczenia maszynowego. Nie należy jednak uwagi tej traktować jako zarzutu dyskwalifikującego pracę.

Praca napisana jest bardzo starannym językiem i poddana została odpowiedniemu formatowaniu. Pojawiają się nieliczne błędy redakcyjne (takie jak powtórzenia części wyrazów w zdaniu). Uporczywie często powtarzającym się błędem jest stosowanie określenia "ilość" do rzeczowników przeliczalnych. Powyższe zarzuty nie umniejszają jednak wartości i jakości pracy, którą oceniam na bardzo wysoką.

Ocena pracy

Pracę oceniam jaką bardzo dobrą. Stanowi ważne i unikalne połączenie głębokiej wiedzy filozoficznej oraz technicznej, związanej z uczeniem maszynowym, w jego bardzo aktualnym ujęciu.

Komentarze krytyczne i punkty dyskusyjne

Nasuują się następujące uwagi o charakterze krytycznym.

1. W rozdziale pierwszym brakuje podanej w sposób jawny definicji sztucznej inteligencji. Autor zdaje się przyjmować jakieś intuicyjne rozumienie tego terminu. Powoduje to pewne zdziwienie, jako że co do terminu nie ma zgodności pośród naukowców. Dodatkowo czasami rozróżnia się "różne poziomy" sztucznej inteligencji, np. "twardą" sztuczną inteligencję. W pracy filozoficznej, brak precyzji w zakresie definicji, o czy jest mowa, tym bardziej zwraca uwagę. Stwierdzenie, że, jak pisze Autor, "*Sztuczną inteligencję można określić jako jedną z dziedzin informatyki, natomiast rozpoznawanie wzorców jako jedną z dziedzin sztucznej inteligencji.*", to jednak trochę za mało.

2. W rozdziale pierwszym przywołano test Turinga i poglądy Searle'a. To dość klasyczne dyskusje odnoszące się do możliwości skonstruowania "twardej" sztucznej inteligencji. Brakuje w tym miejscu choćby przywołania nowszych i istotnych głosów w tym zakresie. Warty wspomnienia jest tu choćby Roger Penrose i jego książka "*Nowy umysł cesarza. O komputerach, umyśle i prawach fizyki*".

3. Omawiając zagadnienia związane z uczeniem maszynowym, nie udało się Autorowi uciec od pewnych uproszczeń i nieścisłości. Przedstawione przykładowe problemy klasyfikacyjne są liniowo lub nieliniowo, ale jednak separowalne. Sugerują one, że w uczeniu maszynowym mamy zazwyczaj do czynienia z dostępnymi cechami (atrybutami), które pozwalają oddzielić od siebie klasy obiektów, a problemem jest głównie znalezienie odpowiednio skomplikowanej (nieliniowej) funkcji. Co jednak w przypadku, gdy klas nie da się idealnie oddzielić? Można to sobie wyobrazić jako sytuację, w której chmury punktów reprezentujących obiekty różnych klas mieszają się w pewnym obszarze przestrzeni atrybutów. Powoduje to, że podejmowanie decyzji klasyfikacyjnej dla obiektów z tego obszaru jest obarczone inherentną i nieusuwalną niepewnością, jeśli nie jesteśmy w stanie dostarczyć innych, lepszych cech. Transformacje przestrzeni cech, ich selekcja, nie rozwiązują tu problemu. Teoretyczne rozważania w uczeniu maszynowym mówią o tym, że optymalne decyzje klasyfikacyjne (co nie oznacza bezbłędne!) podejmowane są tu przez tzw. klasyfikator bayesowski. Jednak problemy związane z jego praktyczną realizacją powodują, że rzadko kiedy można go skonstruować. Podsumowując, algorytmy uczenia maszynowego zmuszone są podejmować decyzje zazwyczaj dla danych w takich problemach. Nie można zatem wprowadzić tu analogii pomiędzy definiowaniem klasy obiektu przez algorytm i przez podejście filozoficzne, które szuka cech unikalnych, jednoznacznych dla danej klasy. W mojej opinii, zagadnienie to jest ciekawe i niewystarczająco podjęte w recenzowanej pracy. Przedstawione przykłady zdają się nawet niebezpiecznie sugerować, że omówiona analogia występuje zawsze, co w mojej opinii nie jest prawdą.

4. Wydaje się, że Autor za mocno akcentuje aspekt projektowania algorytmu uczenia maszynowego przez człowieka, natomiast za mało jest podkreślane to, że to w procesie uczenia, które od pewnego momentu jest automatyczne, wytwarzane są cechy, abstrakcje i końcowe modele decyzyjne. Niezależność tego procesu od człowieka wydaje się być istotna w przedstawianych rozważaniach.

5. Uczenie głębokie (*deep learning*) jest oczywiście wspomniane w pracy. W mojej opinii jest to jednak niewykorzystany potencjał do dyskusji. Sposób konstruowania reprezentacji obiektów, ten właśnie aspekt dochodzenia do definicji realnej przez model głęboki, jest tak odmienny od

funkcjonowania np. drzewa decyzyjnego czy klasyfikatora najbliższego sąsiada, że zaryzykowałbym stwierdzenie, że to modele głębokie powinny być podstawą rozważań, a przynajmniej dominować w pracy. Pozostaje mieć nadzieję, że jeśli Autor będzie kontynuował pracę badawczą w tym zakresie, podda modele głębokie bardziej szczegółowej analizie.

6. Ciekawe pytanie, które może się pojawić podczas czytania pracy jest następujące: czy filozofia jest potrzebna uczeniu maszynowemu? Jak sam Autor wspomina, filozofia jest nauką o poznaniu rzeczywistości bez aspektu koniecznej użyteczności praktycznej. Filozofia nie może zatem zignorować rozwoju sztucznej inteligencji i uczenia maszynowego, które staje się ważnym elementem rzeczywistego świata. A czy rozwój uczenia maszynowego będzie czekał na rozwój filozofii? Jeśli nie, to jakie mogą być konsekwencje? Na ile, według Autora, filozofia jest, czy mogłaby być, przydatna w a) rozwoju sztucznej inteligencji, b) jej społecznym okiełznaniu i oswojeniu?

7. Jednym z imponujących osiągnięć uczenia głębokiego są tzw. modele generatywne, realizowane zazwyczaj przez tzw. sieci neuronowe generatywno-antagonistyczne (GAN, ang. *Generative Adversarial Network*). Nie rozwiązują one problemów klasyfikacji. Są natomiast w stanie wygenerować nowe treści, podobne do tych, którymi były uczone. Nowe treści są często nie do odróżnienia jakościowo od oryginalnych, rzeczywistych przykładów. Jak interpretować możliwości tych modeli w kontekście definiowania realnego, istotowego? Czy jest tu różnica w stosunku do procesu definiowania realnego realizowanego przez model klasyfikacyjny podczas jego uczenia?

8. Niektóre idee filozoficzne twierdzą, że podobieństwa obiektów, abstrakcje, istnieją jedynie w umysłach postrzegających ich ludzi. Pytanie: na ile uczenie maszynowe wspiera tezę, że abstrakcje, podobieństwa istnieją realnie, a nie są jedynie wytworem naszych umysłów? Skoro niezależny model obliczeniowy może je wytworzyć, czy nie jest to argument za ich obiektywnym istnieniem poza umysłem ludzkim? Czy może, przewrotnie, można by stwierdzić, że fakt, że uczenie maszynowe jest w stanie wytworzyć te abstrakcje, przeprowadzić niezbędną analizę oraz syntezę, świadczy o tym, że model uczenia maszynowego powstaje jako coś zbliżonego do naszych umysłów, przynajmniej na pewnym fundamentalnym poziomie? Jakie miało by to znaczenia dla rozwoju filozofii?

Powyższe uwagi w żaden sposób nie dyskwalifikują całości pracy, która, jak podkreślę jeszcze raz, jest unikalna, cenna i bardzo interesująca. W jednej pracy nie sposób zmieścić wszystkiego. Autor zdaje sobie sprawę z mnogości zagadnień i wyzwać badawczych. Jak pisze w ostatnim akapicie pracy:

"Osobno można rozpatrywać takie zagadnienia, jak opisane w pierwszym rozdziale, podobieństwo sztucznej inteligencji do umysłu ludzkiego. Zagadnienia etyczne, które nie zostały opisane w pracy, mogą również odnosić się do definiowania realnego. Wiele kontrowersji budzi pytanie o to, co może wchodzić w zakres kompetencji i odpowiedzialności sztucznej inteligencji. W pewien sposób zagadnienie to będzie powiązane z pytaniem o to, na ile sztuczna inteligencja adekwatnie rozpoznaje rzeczywistość."

Jest to bardzo dobre i zarazem inspirujące podsumowanie. Pokazuje, że Autor zdaje sobie sprawę z istotności już przeprowadzonych badań i ich wpływu na dalsze wyzwania. Jednocześnie kierunkowskazy badawcze na przyszłość są wyraźnie ustawione.

Wnioski końcowe

Podsumowując, stwierdzam, że praca "*Definiowanie realne w maszynowym rozpoznawaniu wzorców*" autorstwa ks. Jakuba Sławomira Płonki, spełnia wymagania stawiane rozprawom doktorskim. Oceniam ją jako bardzo dobrą i wnioskuję o dopuszczenie jej do publicznej obrony.

Michał Bereta