

Dr hab. inż. Tomasz Hachaj
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej
Katedra Informatyki Stosowanej
al. A. Mickiewicza 30, Pawilon C-2
30-059 Kraków

Kraków, dnia 26 stycznia 2026 roku
POLITECHNIKA ŚLĄSKA
Biuro Rady Dyscypliny
Informatyka Techniczna i Telekomunikacja
wpłynęło dnia 30.01.26
nr zał.

Recenzja Rozprawy Doktorskiej
mgr inż. Cezarego Maszczyka
pt. „Zastosowanie głębokiej eksploracji danych, w tym dyskretnego głębokiego uczenia, w indukcji reguł logicznych zawierających złożone warunki elementarne”

Promotor: prof. hab. Marek Sikora
Politechnika Śląska

Wydział Automatyki, Elektroniki i Informatyki

Niniejsza recenzja została opracowana w oparciu o uchwałę Rady Dyscypliny Informatyka Techniczna i Telekomunikacja nr 167/2025 z dnia 28 października 2025 roku podpisanej przez prof. dr hab. inż. Andrzeja Polańskiego, Przewodniczącego niniejszej Rady; która to (uchwała) wyznacza mnie na recenzenta. Treść rozprawy wpłynęła do Sekretariatu Głównego Akademii Górniczo-Hutniczej w dniu 22 grudnia 2025 roku.

1. Ocena wyboru tematu i tez rozprawy

W pracy postawiono i starano się wykazać prawdziwość hipotezy badawczej, że możliwe jest (strona 18):

„(...) opracowanie nowych metod głębokiego dyskretnego uczenia, zastosowanych do budowy regułowych modeli danych. Wyznaczone modele powinny skutecznie rozwiązywać problemy klasyfikacji, regresji oraz analizy przeżycia.”

Autor definiuje również cele szczegółowe swojej rozprawy. Są to (strona 18):

- „1. Zaproponowanie metod pozwalających na wyznaczanie reguł, w których przesłankach pojawiają się złożone warunki elementarne.
2. Zaproponowanie metody pozwalającej na indukcję złożonych wyrażeń logicznych, zawierających zagnieżdżone operatory koniunkcji i alternatywy.
3. Eksperymentalna weryfikacja efektywności zaproponowanych metod zarówno w zakresie zdolności predykcyjnych, jak i opisowych, wyznaczanych zbiorów reguł.”

Praca została wykonana w ramach doktoratu wdrożeniowego realizowanego we współpracy z Instytutem Sztucznej Inteligencji i Cyberbezpieczeństwa należący do Sieci Badawczej Łukasiewicz, z siedzibą w Katowicach (informacja ta znajduje się na stronie 208 doktoratu).

Tematyka recenzowanej rozprawy dobrze lokuje się w obszarze zainteresowań współczesnej informatyki, w szczególności w zakresie zaawansowanych zagadnień uczenia maszynowego, wyjaśnialnej sztucznej inteligencji oraz indukcji reguł logicznych. **Wybór problemu, tezy rozprawy oraz jej celów oceniam pozytywnie.**

2. Ocena zawartości rozprawy

Przedstawiona do recenzji Praca składa się ze 236 stron maszynopisu. Na nośniku CD otrzymałem również tożsamą wersję w formacie PDF. Całość rozprawy napisana jest w języku polskim. Układ i zawartość rozdziałów merytorycznych jest zasadniczo poprawna. Dysertacja składa się z dziesięciu rozdziałów.

Pierwszy rozdział zawiera wprowadzenie do tematu pracy oraz zdefiniowana jest w nim teza.

W drugim rozdziale omówione zostały wybrane typy reguł decyzyjnych: reguły ostre (to na tego typu regułach skupia się Autor w tej pracy), reguły rozmyte, reguły asocjacyjne, reguły akcji, reguły prototypów.

W rozdziale trzecim Doktorant definiuje typy warunków elementarnych występujących w wybranych omawianych wcześniej typach reguł decyzyjnych.

Rozdział czwarty omawia stan wiedzy na temat algorytmów, które mogą być wykorzystywane do indukcji reguł.

W rozdziale piątym przedstawione zostały popularne algorytmy głębokiego dyskretnego uczenia.

W rozdziale szóstym Autor omawia autorskie algorytmy indukcji reguł, których opracowanie pozwoliło na udowodnienie tezy tej rozprawy.

Rozdział siódmy opisuje badania weryfikujące skuteczność proponowanych metod i ich porównanie pod kątem dwóch kryteriów: zdolności predykcyjnych oraz wielkości i złożoności uzyskiwanych zbiorów reguł. Porównania dokonano z algorytmami RuleKit oraz drzewami decyzyjnymi. Te dwa algorytmy będę nazywał w dalszej części tej recenzji algorytmami referencyjnymi. Dyplomant zastosował metodę walidacji krzyżowej.

Ósmy rozdział szczegółowo opisuje wyniki, jakie Autor uzyskał przeprowadzając analizę pięciu testowych zbiorów danych.

W rozdziale dziewiątym Autor omawia wdrożenie wyników swoich badań przez podmiot zatrudniający Doktoranta. Według słów Autora, ten podmiot to (strona 208): „Instytut Sztucznej Inteligencji i Cyberbezpieczeństwa należący do Sieci Badawczej Łukasiewicz, z siedzibą w Katowicach (ozn. Ł-AI), wcześniej znany jako Instytut Technik Innowacyjnych EMAG”. Opracowana w ramach doktoratu platforma RuleMiner jest obecnie ogólnodostępna, i znajduje się pod adresem: <https://app.ruleminer.ai/> (dostęp 24.01.2026). Autor deklaruje (strona 209), że „brał czynny udział w projektowaniu i implementacji całego systemu jako jeden z głównych wykonawców”. Dyplomant informuje, że do platformy RuleMiner został wdrożony opracowany przez niego algorytm DeepRules.

Ostatni, dziesiąty rozdział to podsumowanie całości pracy oraz wnioski. W rozdziale tym na niecałych sześciu stronach w skondensowany sposób opisane są najważniejsze osiągnięcia zawarte w doktoracie.

Bibliografia składa się ze 199 numerowanych pozycji. Zawiera prace wydane w czasopismach naukowych, materiałach konferencyjnych, tak zwane preprinty oraz książki.

Rozprawa zawiera wszystkie konieczne elementy, które powinny znajdować się rozprawie doktorskiej i została wykonana z należytą starannością. **Zawartość rozprawy oceniam w pełni pozytywnie.**

3. Oryginalne osiągnięcia Autora rozprawy

W recenzowanej pracy zawarto kilka wartościowych i oryginalnych koncepcji oraz rozwiązań, dokonano ich implementacji oraz uzyskano wyniki wzbogacające naszą wiedzę w zakresie omawianych zagadnień. Autor w dysertacji proponuje trzy metody uczenia opartego na regułach:

- Metodę ComplexConditions (kilkukrotnie nazwaną przez Doktoranta ComplexRules),
- Metodę MofNRules,
- Metodę DeepRules.

Metody te pozwalają na generowanie reguł zawierających warunki wewnętrznych alternatyw w przesłankach. W odróżnieniu od metod opartych na gradientowym uczeniu, wszystkie proponowane algorytmy mają charakter heurystyczny i bazują na strategii sekwencyjnego pokrywania, nie wymagając stosowania różniczkowalnych funkcji straty.

ComplexConditions wykonuje indukcję zbioru reguł w postaci dysjunkcyjnej (DNF) i bazuje na algorytmie RuleKit [74]. Rozszerza go jednak o możliwość generowania reguł zawierających różne rodzaje warunków złożonych, które pozwalają odkrywać i opisywać nowe zależności w danych. Algorytm ten jest opisany w sekcji 6.1 Dysertacji. Ma on złożoność wielomianową.

MofNRules jest rozwinięciem metody ComplexConditions i wprowadza do swojego pierwotnego warunku typu M-z-N (zdefiniowane na stronie 33 dysertacji). Opracowane zostały dwa warianty algorytmu, umożliwiające generowanie reguł zarówno z warunkami przynajmniej M-z-N, jak i dokładnie M-z-N. Wartości zarówno M jak i N mogą być kontrolowane poprzez parametry algorytmu. Uzyskiwane przy pomocy MofNRules zbiory reguł mogą zawierać wszystkie rodzaje warunków złożonych wygenerowanych przez ComplexConditions. Ich przesłanki nie muszą więc składać się wyłącznie z warunków M-z-N. Algorytm ten jest opisany w sekcji 6.2 dysertacji. Złożoność obliczeniowa drugiej iteracji indukcji została oszacowana na (trzymając się oznaczeń ze strony 25 i 33) $O(n^N)$ gdzie n to liczba przykładów uczących a N to liczba warunków logicznych warunku M-z-N.

DeepRules pozwala na indukcję reguł z zagnieżdżonymi wyrażeniami logicznymi koniunkcji i alternatyw w przesłankach, nawiązując do koncepcji głębokiego dyskretnego uczenia. DeepRules pozwala na generowanie reguł zawierających wszystkie rodzaje warunków złożonych stosowanych we wcześniej wspomnianych algorytmach. Główna różnica względem wcześniejszych metod polega na iteracyjnym dodawaniu do reguły składników zamiast pojedynczych warunków. Zastosowanie zagnieżdżonych koniunkcji i alternatyw w przesłankach pozwala operować na mniejszej przestrzeni poszukiwań, redukując przy tym czasy obliczeń. Algorytm ten jest opisany w sekcji 6.3 dysertacji. Ma on złożoność wielomianową.

Zaproponowane algorytmy zostały przedstawione przy pomocy pseudokodów w sposób zbliżony do implementacji, Autor na przykład uwzględnia „flagi”, które mają za zadanie wskazywać, czy algorytm będzie brał pod uwagę lub nie pewne funkcjonalności. W tabeli 6.6 na stronie 114 znajduje się zestawienie opracowanych metod, w których Autor porównuje sposoby generowania reguł, typ stosowanej konstruktywnej indukcji, zdolność algorytmu do generacji warunków zanegowanych itd. Umieszczenie tej tabeli było bardzo dobrym posunięciem, ponieważ pozwala czytelnikowi na porównanie najważniejszych cech zaproponowanych przez Doktoranta algorytmów.

W trakcie przeprowadzonych eksperymentów zostało użytych 74 (30 klasyfikacyjnych, 30 regresyjnych i 14 dotyczących analizy przeżycia) publicznie dostępnych, tabelarycznych zbiorów danych. Charakterystyki tych zbiorów zostały podane w tabelach 7.1, 7.2, 7.3 i 7.4 pracy. Do oceny działania zaproponowanych algorytmów został użyty szereg współczynników takich jak BAcc, C2, RRMSE, IBS, test statystyczne Friedmana, Nemenyi i Wilcoxa oraz inne metody, dobrane odpowiednio do dyskutowanego problemu. Zastosowane przez Doktoranta podejście do oceny porównawczej algorytmów jest w pełni poprawne i uzasadnione. Autor dokonał również oceny, jak wprowadzenie

złożonych form warunków i przesłanek wpłynęło na złożoność uzyskiwanych zbiorów reguł, liczbę reguł, średnią liczbę warunków w przesłance oraz sumaryczną liczbę warunków we wszystkich regułach. Jest to najobszerniejsza część pracy i zajmuje strony od 117 do 177. Na podstawie przeprowadzonych badań Autor zaobserwował, że w niektórych eksperymentach proponowane algorytmy osiągnęły statystycznie lepsze wyniki od algorytmów referencyjnych biorąc pod uwagę problemy klasyfikacji, regresji i analizy przeżycia. Ciekawa, kompleksowa dyskusja na ten temat znajduje się na stronach 176-177 Dysertacji.

Dobrym pomysłem było napisanie rozdziału ósmego, w którym Autor szczegółowo opisuje wyniki eksperymentalne zaproponowanych w dysertacji metod dla pięciu różnych zbiorów danych. Wyniki te pozwalają czytelnikowi na szczegółowy wgląd w uzyskane wyniki, które w rozdziale siódmym rozważane były zbiorczo, najczęściej jako zagregowane obliczenia z walidacji krzyżowej. Podobnie jak w poprzednim rozdziale dyskusja wyników została przeprowadzona w sposób zrozumiały i rzetelny. Podsumowując swoje badania na podstawie przeprowadzonych eksperymentów Doktorant zauważa, że żadna z zaproponowanych przez niego metod nie jest uniwersalnie najlepsza dla wszystkich typów problemów, oraz, że wyniki pokazują, że skuteczność tych metod jest w dużej mierze zależna od charakterystyki danych. Jest to słuszne spostrzeżenie, prawdziwe dla praktycznie wszystkich algorytmów z grupy metod uczenia maszynowego. Przeprowadzone przez Autora badania pokazały, że opisane metody mają potencjał do generowania zbiorów reguł o mniejszej złożoności niż algorytmy referencyjne, a ich wyniki predykcyjne pozostają przy tym na ogół na zbliżonym poziomie. **Autor w ten sposób udowodnił tezę badawczą rozprawy.**

W dysertacji zaproponowano również dalsze kierunki badawcze takie jak:

- opracowanie metod optymalizacji zapisu reguł,
- dodanie odpowiednich parametrów do opracowanych metod, parametry pozwolą na ograniczenie, które atrybuty mogą być ze sobą porównywane,
- uwzględnienie wiedzy domenowej eksperta w procesie indukcji reguł.

Są to interesujące oraz istotne plany. Sformułowanie ich świadczy o głębokim zrozumieniu badanej dziedziny wiedzy i dojrzałości naukowej Doktoranta.

Wymienione osiągnięcia są oryginalne i znaczące. Na nich zatem przede wszystkim opieram ogólnie pozytywną ocenę rozprawy. Udowadniają one postawioną tezę rozprawy oraz stanowią wkład w rozwój algorytmów uczenia maszynowego, wyjaśnialnej sztucznej inteligencji oraz indukcji reguł logicznych

Autor wykazał się również aktywnością publikacyjną. W bibliografii doktoratu znajduje się dwie publikacje, których jest współautorem. Są to pozycje [112]: praca konferencyjna w ramach uznanej w środowisku naukowym międzynarodowej siedemnastej edycji konferencji Conference on Computer Science and Intelligence Systems (FedCSIS), oraz artykuł [113] wydany w czasopiśmie Machine Learning and Knowledge Extraction (WOS impact factor 6 dla roku 2025). W obu tych pracach Doktorant jest pierwszym autorem. Oznacza to, że doktorant bierze aktywny udział w życiu naukowym.

4. Uwagi dyskusyjne i krytyczne

Jak już wcześniej wspomniałem, przedłożona do recenzji rozprawa pod względem merytorycznym i redakcyjnym napisana jest w większości poprawnie. Doktorant w sposób dokładny i wyczerpująco omawia rozwiązany problem badawczy. Miejscami praca przyjmuje formę zbliżoną do podręcznika akademickiego – moim zdaniem nie ma konieczności wprowadzania pojęć, które nie będą używane w późniejszym wywodzie i nie są konieczne do zrozumienia tematu, na przykład różne rodzaje reguł w rozdziale 2, algorytmy indukcji reguł w rozdziale 4, algorytmy dyskretnego uczenia w rozdziale 5, do których nigdy później Doktorant już się nie odnosi. Praca dzięki temu mogłaby być znacznie krótsza, w dalszym ciągu mieć bogatą bazę literaturową świadczącą o zgłębieniu tematu przez Autora

i równocześnie nie tracić przy tym walorów merytorycznych. Również przy tak długiej pracy korzystne byłoby dodanie indeksu użytych symboli.

Szkoda, że nie wszystkie definicje oraz równania są numerowane. Znacznie ułatwiłoby to zarówno czytelnikowi jak i Autorowi dysertacji ściśle odwoływanie się do konkretnych efektów pracy, bez konieczności stosowania języka opisowego.

Praca zawiera pewne nieścisłości, potknięcia językowe oraz drobne błędy, których obecność nie zaskakuje w tak długim tekście. Nie stanowią one przeszkody w zrozumieniu wywodu Autora jak również nie wpływają znacząco na jakość merytoryczną pracy. Jako recenzent mam obowiązek je wskazać. Dla porządku wymienię je w kolejności pojawiania się w tekście:

1. Strona 14: „Symboliczne metody uczenia maszynowego, inherentnie interpretowalne, świetnie wpisują się w tę koncepcję” – moim zdaniem lepiej brzmi: „Symboliczne metody, których działanie jest bezpośrednio interpretowalne, świetnie wpisują się w tę koncepcję”.
2. Strona 15: „Problem klasyfikacji polega na zgadnięciu” – moim zdaniem lepiej brzmi: „przewidzeniu”.

Warto zaznaczyć, że definicja regresji wprowadzona przez autora na stronie 15 nie uwzględnia regresji funkcji z inną przeciwdziedzina niż $\mathbb{R}$. Jest to pewne ograniczenie, które staje się jasne w dalszej części pracy, w której Autor precyzuje przedmiot swoich dociekań.

3. Strona 15: Co we wzorze na dole tej strony oznaczone jest symbolem „X” i „f”?
4. Strona 16: „którego prawdziwość można ustalić na podstawie zmiennych należących do A” – chodzi zapewne o uporządkowany zbiór wartości atrybutów przyjmujących wartości w dziedzinie wyznaczonej przez A.
5. Strona 16, definicja c – tę definicję słowną można by zastąpić przez definicję matematyczną gdzie c jest funkcją o odpowiedniej dziedzinie i przeciwdziedzinie. To samo dotyczy zbiorów P i N. Nie jest do końca jasne, czy C jest zbiorem wszystkich możliwych warunków logicznych, czy też jest zbiorem warunków logicznych występujących pośród zdefiniowanych reguł.
6. Strona 17: „Zastosowanie bardziej złożonych warunków logicznych w symbolicznych metodach uczenia maszynowego może prowadzić do uzyskiwania prostszych, mniej rozbudowanych, a przy tym łatwiejszych w zrozumieniu dla człowieka modeli” – prostszych, ale w porównaniu do jakich innych modeli?
7. Strona 17: „(...) pozwalają uzyskiwać relatywnie wysoką jakość predykcji” – trafniejszym słowem jest „skuteczność” (effectiveness).
8. Strona 31: „W literaturze znaleźć można również metody używające negacji wyżej wspomnianych warunków prostych. W przypadku atrybutów numerycznych warunek taki przyjmuje postać $a \neq v_i$.” Autor nie podaje, w jakich pracach można znaleźć tego typu zapis negacji warunków prostych, ale jeśli warunki te mają postać nierówności, tak jak jest to podane na stronie 30, to ten zapis jest nieintuicyjny i mylący. Dużo lepsza jest definicja znajdująca się kolejnym zdaniu, która używa alternatyw dopełnień. Zanegowany warunek rozumiany jako brak przynależności do przedziału $a \notin \langle v_1, v_2 \rangle$ jest w pełni zrozumiały. Z dalszej części wywodu wynika, że v_i w $a \neq v_i$ nie jest granicą przedziału a pojedynczą liczbą, ale jest to niespójne ze wcześniejszą definicją warunków prostych ze strony 30. Możliwe jest też, że Autor miał na myśli atrybuty nominalne.
9. Strona 31: „Wśród najbardziej popularnych metod indukcji reguł, zawierających tego typu warunki, wymienić można między innymi algorytm C2N (...) Są one również [warunki – przypisek recenzenta] wspierane w wersji drugiej algorytmu RuleKit” – zapewne chodzi o to, że RuleKit pozwala na używanie wymienionych wcześniej typów warunków.
10. Strona 36, 37, wzór (3.4) – operator $\odot$ był zdefiniowany wcześniej na stronie 30 i oznaczał operator relacji w warunkach prostych dla atrybutów numerycznych i nie uwzględniał równości

- i nierówności natomiast dla atrybutów nominalnych możliwy był tylko operator równości. Która z tych dwóch definicji jest obowiązująca w reszcie pracy?
11. Strona 42: „W konstruktywnej indukcji, sterowanej hipotezami (...)” – przecinek jest niepotrzebny.
 12. Strona 54, oraz Strona 55 Tabel 5.1: „Eksperymenty przeprowadzone przez autorów tej pracy na innych popularnych benchmarkowych zbiorach danych nie potwierdziły jednak tych wniosków. Wyniki tychże eksperymentów przedstawione zostały w tabeli 5.1.” Z kontekstu wynika, że chodzi o eksperymenty przeprowadzone przez autorów pracy [139] (Qiao, L., Wang, W., Lin, B. Learning accurate and interpretable decision rule sets from neural networks. AAAI Conference on Artificial Intelligence, 2021.,) jednak w pracy [139] nie widzę tych wyników. Czy eksperymenty przedstawione w Tabeli 5.1 przeprowadził Doktorant? Metryka BAcc jest zdefiniowana dopiero na stronie 119 (wzór (7.1), jest tam podana inna pisownia: Bacc, ale zakładam, że jest to ta sama wielkość) a informacje o zbiorach danych, na których przeprowadzony był test na stronie 123 w Tabeli 7.1. Tabela 5.1 powinna więc znaleźć się znacznie później w pracy.
 13. Strona 57: „Poniżej został rozpisany przykład indukcji tego typu reguły dla klasy 1.” – jaką klasę 1 Autor ma tu na myśli?
 14. Strona 62: „W ten sposób możliwe jest osiągnięcie zadowalających wyników, nawet gdy część reguł w zbiorze sama w sobie nie posiada wysokiej jakości.” – na tym etapie Autor nie zdefiniował czym jest „jakość reguł”. Dokonuje tego na stronie 66 w Algorytmie 2. Dla jasności wyводу powinno się wprowadzać definicję pojęcia przed jego pierwszym użyciem. „Sama w sobie” jest redundancją językową.
 15. Strona 75 równanie (6.2) oraz strona 76 równanie (6.3) – Autor nie ustrzegł się kolizji oznaczeń. Na stronie 16 P i N to przykłady pozytywne i negatywne w zbiorze D, na stronie 18 występuje warunek M-z-N (N nie ma tu nic wspólnego z N ze strony 16), na stronie 25 N jest parametrem algorytmu [2], na stronie 70 N jest parametrem kontrolującym maksymalną liczbę składników alternatywy itd. Jak zdefiniowane jest więc p, P, n i N w równaniach (6.2) i (6.3), które są istotnym składnikiem algorytmu ComplexConditions?
 16. Strona 78 rysunek 6.1 – do tej pory atrybuty były oznaczane jako a_1 , a_2 itd. (patrz strona 15). Teraz Autor zastępuje je kolejnymi literami alfabetu „a” i „b”. Warto byłoby zachować spójność wyводу.
 17. Strona 80: „gdzie m i n oznaczają kolejno liczbę kolumn i wierszy w treningowym zbiorze danych” – na stronie 15 mówiono o przykładach uczących i atrybutach. Które z nich to kolumny a które wiersze?
 18. Strona 86, Twierdzenie 1: jeśli bias jest równy $1-M$, to zgodnie z (6.10): $y = f(\sum_{i=1}^N a_i + 1 - M)$ a nie $y = f(\sum_{i=1}^N a_i - M)$ jak jest w pracy.
 19. Strona 90, Twierdzenie 2: na stronie 25 mamy „Zbiorem częstym nazywamy zbiór elementów, pojawiający się minimalnie w pewnej określonej liczbie transakcji”. Jak zdefiniowane jest pojęcie „często spełniony warunek M-z-N”? Można się tego domyslać z informacji na stronie 89 ale ścisła definicja tam nie pada.
 20. Strona 97 „Przed przedstawieniem ostatniej z metod proponowanych przez autora niniejszej pracy, aby lepiej wyjaśnić jej sens, wróćmy na moment do rodzajów warunków złożonych stosowanych przez dwie wcześniejsze” – dysertacja tworzy ciąg logiczny, w którym nie ma potrzeby ponownie definiować ani przypominać wcześniej zdefiniowanych pojęć. W razie potrzeby należy odwołać się do odpowiednich definicji, twierdzeń bądź przykładów podając ich unikalny numer.
 21. Strona 101: „Możliwe jest zdefiniowanie osobnych miar wykorzystywanych do wyboru najlepszej postaci reguły DNF i CNF (patrz podsekcja 3.4).” oraz „Algorytm umożliwia także stosowanie różnych miar jakości podczas generowania reguł CNF i DNF (patrz podsekcja 3.4).” Co Autor ma na myśli pisząc „osobne miary”? Czy chodzi o równanie (6.2) i (6.3)?
 22. Strona 103, równania (6.18) oraz (6.19) – brak wytłumaczenia, co oznaczają symbole $\hat{\pi}(t|x_i)$, $\hat{G}(y_i)$, $\hat{G}(t)$ zastosowane w tych równaniach.
 23. Strona 107: na liście parametrów znajduje się dwa razy parametr o nazwie „dnf_quality_measure” ale ma dwie różne definicje. Która z tych dwóch wersji jest właściwa?

24. Strona 116: „Problem wysokiej złożoności obliczeniowej został częściowo zaadresowany poprzez parametry algorytmów (...)” – moim zdaniem lepiej brzmi: „Autor zaproponował metodę pozwalającą ograniczyć problem wysokiej złożoności obliczeniowej poprzez użycie parametrów algorytmów (...)” .
25. Strona 116: „(...) realnych zbiorów danych (...)” – moim zdaniem lepiej brzmi: „(...) zbiorów danych o stosunkowo dużej liczbie przykładów uczących (...)”
26. Strona 118 „dla wszystkich 10 foldów walidacji krzyżowej” – moim zdaniem lepiej brzmi: „dla dziesięciokrotnej walidacji krzyżowej”.
27. Strona 118 „W przypadku problemu analizy przeżycia ich liczba zbiorów była mniejsza i wyniosła 14.” – moim zdaniem lepiej brzmi: „W przypadku problemu analizy przeżycia użyto 14 zbiorów testowych.”
28. Strona 119: Zgodnie z [28] balanced accuracy ma wzór: $\frac{1}{2} \left(\frac{TP}{P} + \frac{TP}{N} \right)$
29. Strona 126: „(patrz równanie (6.2).” – brak domknięcia nawiasu „)”.
30. Strona 131, rysunek 7.3: czy napis „Korel.” Widoczny w podpisach wierszy i kolumn oznacza (6.3)?
31. Strona 176: „W przypadku regresji z kolei każda z reguł (...)” – moim zdaniem lepiej brzmi: „W przypadku regresji, każda z reguł (...)”
32. Strona 196 oraz strona 197 rysunek 8.3: „Na jego podstawie widać, że rozkład ten przypomina rozkład Gaussa” – czy ten rozkład jest rzeczywiście normalny? Stawiając tego typu tezę w pracy doktorskiej, należałoby ją zweryfikować odpowiednim testem statystycznym, na przykład Shapiro-Wilka albo Kołmogorowa-Smirnowa.
33. Strona 210: Autor nie odnosi się do rysunku 9.1 w tekście swojej pracy. Rysunek przedstawia zrzut ekranu formularza konfiguracyjnego algorytmu DeepRules w aplikacji RuleMiner.
34. Strona 214: link na dole strony nie ma daty dostępu. W dniu 24.01.2026 link był aktywny.

Uwagi redakcyjne:

35. Brak konsekwentnego podpisywania tabel – czasem podpisy znajdują się nad tabelami (np. w wypadku tabeli 6.5, 6.6, 7.1, 7.2 itd.) a czasem pod nimi (np. w tabeli 2.1, 5.1, 6.1, 6.2, 6.3, 6.4, 7.10, 7.11 itd.).
36. Opracowanie wyników w tabelach – niezależnie od tego, czy na danym miejscu po przecinku są zera czy nie, należy wypisać wszystkie cyfry dla założonej dokładności dziesiętnej. Na przykład jeśli założymy, że wypisujemy trzy cyfry (tak jak jest w Tabeli 8.17), zamiast 7.54 powinno zostać napisane 7.540 itd. W niektórych tabelach Autor stosuje się do tej reguły (np. w tabeli 5.1, 7.8, 7.9, 7.15 itd.) a w niektórych nie (na przykład w tabeli 7.10, 7.11, 7.12, 7.13, 7.14 itd.).

Jeszcze raz zaznaczam, że wymienione powyżej uwagi nie wpływają w sposób bardzo istotny na poznawcze oraz użytkowe wartości proponowanych rozwiązań i metodologii ich oceny.

5. Podsumowanie

Przytoczone powyżej uwagi polemiczne nie umniejszają wartości merytorycznej pracy, która stanowi oryginalny wkład Autora w zagadnienia związane z uczeniem maszynowym, wyjaśnianą sztuczną inteligencją oraz indukcją reguł logicznych.

Podsumowując recenzję stwierdzam, że moja generalna opinia o rozprawie doktorskiej „Zastosowanie głębokiej eksploracji danych, w tym dyskretnego głębokiego uczenia, w indukcji reguł logicznych zawierających złożone warunki elementarne” mgr inż. Cezarego Maszczyka jest pozytywna. Uważam, że przedstawiona do recenzji Rozprawa zawiera samodzielne i oryginalne, proponowane przez Autora rozwiązanie trudnego i aktualnego problemu badawczego, o potencjalnie licznych praktycznych zastosowaniach. Przedstawiona teza rozprawy została dowiedziona a postawione cele i

zadania pracy zrealizowane. Odpowiada to ustawowym wymaganiom stawianym rozprawom doktorskim. **Na tej podstawie wnioskuję o dopuszczenie rozprawy do publicznej obrony w celu uzyskania przez jej Autora stopnia doktora w dziedzinie nauk inżynieryjno – technicznych w dyscyplinie naukowej informatyka techniczna i telekomunikacja.**

Tomasz Hachaj