



dr hab. Beata Zielosko, prof. UŚ  
Instytut Informatyki  
Wydział Nauk Ścisłych i Technicznych  
Uniwersytet Śląski w Katowicach

Katowice, 19 stycznia 2026 r.

## Recenzja rozprawy doktorskiej

Tytuł rozprawy: ZASTOSOWANIE GŁĘBOKIEJ EKSPLOKACJI DANYCH, W TYM DYSKRETNEGO GŁĘBOKIEGO UCZENIA, W INDUKCJI REGUŁ LOGICZNYCH ZAWIERAJĄCYCH ZŁOŻONE WARUNKI ELEMENTARNE

Autor rozprawy: **mgr inż. Cezary Maszczyk**

Promotor rozprawy: **prof. dr hab. Marek Sikora**

Dziedzina: nauki inżynierijsko-techniczne

Dyscyplina: Informatyka Techniczna i Telekomunikacja

Recenzja wykonana na podstawie uchwały Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej z dnia 28.10.2025 w przedmiocie wyznaczenia recenzentów niniejszej pracy doktorskiej (uchwała 167/2025).

### 1. Obszar problemowy pracy

Postęp technologiczny oraz złożoność i dynamika procesów zachodzących w otaczającym świecie wymuszają konieczność przetwarzania i analizy różnego rodzaju danych. W takich warunkach istotną rolę odgrywa eksploracja danych DM (data mining), stanowiąca centralny etap procesu odkrywania wiedzy z danych KDD (knowledge discovery in databases). Istotnym elementem eksploracji danych jest ocena jakości utworzonych modeli. W przypadku zadań predykcyjnych wykorzystuje się różnorodne miary, takie jak dokładność, precyzja, czułość czy miary błędu, które pozwalają ocenić skuteczność przewidywań. Jednak, równie ważna jest forma reprezentacji odkrytej wiedzy – powinna ona być nie tylko poprawna statystycznie, ale także zrozumiała i użyteczna z punktu widzenia użytkownika. W tym kontekście szczególne znaczenie mają modele interpretowalne.



Jedną z najbardziej przejrzystych i użytecznych form reprezentacji wiedzy pozyskanej w procesie eksploracji danych są reguły decyzyjne. Pozwalają one w sposób zrozumiały przedstawić zależności występujące w danych w postaci warunków i odpowiadających im decyzji. Dzięki temu możliwe jest nie tylko automatyczne wnioskowanie, ale również intuicyjna interpretacja wyników przez człowieka.

Reguły decyzyjne tworzą zróżnicowaną klasę modeli, obejmującą różne typy reguł, np. reguły klasyfikacyjne, regresyjne, rozmyte czy probabilistyczne. Różnią się one zarówno strukturą, jak i sposobem interpretacji. Kluczowym elementem każdej reguły są warunki elementarne, które mogą przyjmować postać prostych zależności logicznych, jak i bardziej złożonych form uwzględniających niepewność czy relacje między atrybutami. Różnorodne są również podejścia do indukcji reguł – od klasycznych metod eksploracji danych, takich jak algorytmy sekwencyjnego pokrywania, drzewa decyzyjne, po nowoczesne techniki wykorzystujące głębokie uczenie.

Recenzowana rozprawa doktorska dotyczy głębokich dyskretnych metod uczenia zbiorów reguł o złożonej postaci, łącząc zdolność modeli głębokich do uchwycenia złożonych zależności z interpretowalnością reprezentacji regułowej. Autor zaproponował rozwinięcie koncepcji głębokiego dyskretnego uczenia nie tylko w zakresie indukcji skomplikowanych i zagnieżdżonych formuł logicznych, ale także w kierunku wykorzystania bardziej złożonych warunków wewnątrz przesłanek reguł. Wynikiem tych prac są zaproponowane trzy metody uczenia reguł. Metoda ComplexRules pozwalająca na generowanie reguł ze złożonymi formami warunków wewnątrz przesłanek, metoda MofNRules uwzględniająca warunki składające się z  $M$  warunków logicznych oraz metoda DeepRules zapewniająca możliwość indukcji zagnieżdżonych formuł logicznych.

Autor przedstawił w pracy jasno sformułowany główny cel pracy, jakim jest *opracowanie nowych metod głębokiego dyskretnego uczenia, zastosowanych do budowy regułowych modeli danych*. Wyznaczone modele powinny skutecznie rozwiązywać problemy klasyfikacji, regresji oraz analizy przeżycia.

Zadanie to zostało uszczegółowione, obejmując cele pomocnicze dotyczące (i) złożonych warunków elementarnych występujących w regułach, (ii) indukcji złożonych wyrażeń logicznych zawierających zagnieżdżone operatory koniunkcji i alternatywy oraz (iii) eksperymentalnej weryfikacji zaproponowanych metod.

Podsumowując, uważam, że podjęta w badaniach problematyka, sformułowany cel i zadania badawcze są interesujące i istotne dla rozwoju dyscypliny. Autor przeprowadził analizę istniejących podejść w zakresie realizowanego tematu badawczego, zaproponował autorskie rozwiązania oraz zweryfikował ich skuteczność, co pozwoliło na przeprowadzenie analizy porównawczej zaproponowanych algorytmów oraz ocenę opracowanych rozwiązań.

## 2. Kompozycja i zawartość pracy

Rozprawa została napisana w języku polskim i składa się z dziesięciu rozdziałów oraz bibliografii, spisu rysunków i spisu tabel. Całość jest obszerna i liczy 236 stron.

Rozdział pierwszy stanowi wstęp do tematyki rozprawy. Autor przedstawił problematykę badań, motywacje do podjęcia badań, cel pracy, główne osiągnięcia i układ pracy. Trzy kolejne rozdziały mają charakter wprowadzający i stanowią podstawę teoretyczną dla metod zaproponowanych w dalszej części pracy. Rozdział drugi przedstawia wybrane typy reguł decyzyjnych, takie jak reguły ostre, rozmyte, asocjacyjne, reguły akcji i prototypów. W rozdziale trzecim Autor omówił typy warunków elementarnych występujących w regułach decyzyjnych, w tym warunki proste, warunki z negacją, wewnętrzne alternatywy, warunki  $M$  z  $N$ , warunki skośne oraz warunki relacji atrybutów. Kolejny rozdział rozprawy omawia istniejące podejścia do indukcji reguł decyzyjnych, w tym sekwencyjne pokrywanie, konstruktywna indukcja, metody zespołowe, metody wykorzystujące sztuczne sieci neuronowe oraz zagadnienia dotyczące predykcji z wykorzystaniem modeli regułowych. Rozdział piąty rozprawy poświęcony jest tematyce głębokiego dyskretnego uczenia w kontekście reguł decyzyjnych. W rozdziale szóstym Autor omówił opracowane metody indukcji reguł: metodę indukcji reguł z warunkami złożonymi – ComplexCondition (ComplexRules), metodę  $M$ -z- $N$  – MofNRules oraz metodę indukcji reguł z zagnieżdżonymi wyrażeniami logicznymi – DeepRules. Rozdział siódmy zawiera eksperymenty wraz z opisem metodologii badań, zbiorów danych użytych w eksperymentach, wartości parametrów badanych algorytmów oraz wyniki badań obejmujące dwa aspekty: zdolności predykcyjne proponowanych metod oraz złożoność wyindukowanych zbiorów reguł. Rozdział ósmy przedstawia studium przypadków dotyczące pogłębionej analizy opracowanych metod poprzez ich zastosowanie do wybranego problemu oraz rzeczywistych danych. Analiza obejmowała popularne zbiory dotyczące problemu klasyfikacji tj. „Kwiaty irysa” oraz zbiory danych „MONK” oraz trzy studia przypadków obejmujące zbiór danych „Heart Disease”, zbiór „Ozone” badany w kontekście problemu regresji oraz zbiór „BMT-Ch” badany w ramach analizy przeżycia. Rozdział dziewiąty omawia kwestie dotyczące wdrożenia algorytmu DeepRules do platformy RuleMiner. Ostatni rozdział pracy zawiera podsumowanie i wnioski. Rozprawę kończy bibliografia obejmująca 199 pozycji. W większości są to artykuły z czasopism naukowych oraz prace konferencyjne. Obszerność bibliografii wynika zarówno z długiej historii badań nad tą tematyką, jak i z jej ponownego rozwoju w kontekście wyjaśnialnej sztucznej inteligencji (XAI, Explainable Artificial Intelligence).

Podsumowując tę część recenzji, uważam, iż praca jest obszerna, co wynika zarówno z faktu, że reguły decyzyjne są szeroko badanym zagadnieniem w literaturze, jak i ze złożoności podjętej problematyki badawczej. Rozprawa przedstawia dużą zawartość

merytoryczną, w tym naukową i eksperymentalną, związaną z przeprowadzonymi badaniami.

### 3. Opinia o rozprawie i oryginalność osiągnięć

Ważnym elementem rozprawy jest podjęta problematyka badawcza, która koncentruje się na regułach decyzyjnych jako jednej z najbardziej intuicyjnych i interpretowalnych form reprezentacji wiedzy. Jednocześnie praca dotyczy zagadnień związanych z indukcją reguł o złożonej strukturze, co stanowi aktualny i wymagający obszar badawczy, szczególnie w kontekście współczesnych metod odkrywania wzorców z danych i wyjaśnialnej sztucznej inteligencji.

Istotnym osiągnięciem Autora jest rozszerzenie idei głębokiego dyskretnego uczenia w kierunku indukcji złożonych i zagnieżdżonych formuł logicznych, a także wykorzystania rozbudowanych warunków w przesłankach reguł. W rezultacie opracowano trzy nowe metody indukcji reguł oparte na podejściu sekwencyjnego pokrywania i każda z metod generuje zrozumiały dla człowieka opis danych w postaci zbioru reguł decyzyjnych.

Metoda ComplexRules polega na rozszerzeniu klasycznego procesu indukcji reguł decyzyjnych poprzez wprowadzanie złożonych warunków elementarnych w przesłankach reguł. Pozwala to na tworzenie bardziej zwartych opisów danych niż w przypadku klasycznych algorytmów regułowych, które lepiej odwzorowują zależności występujące w danych. Rodzaje generowanych warunków złożonych to: warunki przedziałów, negacje warunków, negacje warunków przedziałów, warunki relacji atrybutów, warunki zbioru dyskretnego, warunki wewnętrznej alternatywy dla atrybutów numerycznych. Generowanie takich warunków odbywa się w sposób kontrolowany, na podstawie jakości uzyskiwanych reguł.

Osiągnięciem Autora w ramach metody MofNRules jest wprowadzenie do procesu indukcji reguł decyzyjnych warunków typu M-z-N (M-of-N), które umożliwiają wyrażanie zależności polegających na spełnieniu co najmniej M spośród N warunków, z wykorzystaniem koncepcji zbiorów częstych. W odróżnieniu od klasycznych koniunkcji, gdzie wymagane jest spełnienie wszystkich warunków, podejście to pozwala na bardziej elastyczne modelowanie złożonych zależności występujących w danych bez konieczności generowania wielu alternatywnych reguł. Warunki typu M-z-N stanowią uogólnienie klasycznych warunków logicznych oraz nawiązują do modeli perceptronowych, co pozwala łączyć interpretowalność reguł z większą zdolnością predykcyjną. Ponadto warunki w metodzie MofNRules mogą być zarówno warunkami prostymi, jak i warunkami złożonymi występującymi w metodzie ComplexRules.

Osiągnięciem Autora w ramach metody DeepRules jest możliwość indukcji mieszanych zbiorów reguł, w których jednocześnie występują reguły z przesłankami w postaci koniunkcyjnej CNF (conjunctive normal form) i dysjunkcyjnej DNF (disjunctive normal form). W klasycznych podejściach reguły mają zazwyczaj jednolitą strukturę, natomiast proponowana metoda umożliwia współistnienie różnych form logicznych w ramach jednego modelu. Pozwala to na lepsze dopasowanie formy reguł do charakteru danych oraz bardziej efektywne odwzorowanie złożonych zależności występujących w danych.

Ważnym i interesującym aspektem pracy są badania dotyczące weryfikacji opracowanych metod indukcji reguł oraz studium przypadków obejmujące analizę rzeczywistych zbiorów danych w problemach klasyfikacji, regresji i analizy przeżycia, a także porównanie z modelami referencyjnymi w postaci drzew decyzyjnych oraz algorytmem RuleKit. Rozbudowany zakres eksperymentów, w tym analiza istotności statystycznej uzyskanych wyników, pozwolił na kompleksową ocenę efektywności oraz zastosowalności proponowanych metod w zróżnicowanych problemach analitycznych.

Należy także zwrócić uwagę, iż Doktorant realizował badania w ramach piątej edycji programu „Doktorat wdrożeniowy”. Opracowany algorytm DeepRules został zaimplementowany i wdrożony do platformy RuleMiner w Instytucie Sztucznej Inteligencji i Cyberbezpieczeństwa należącym do Sieci Badawczej Łukasiewicz, z siedzibą w Katowicach.

Podsumowując, uważam, że zrealizowane w ramach przedstawionej rozprawy badania stanowią ważny wkład w rozwój dyscypliny, zwłaszcza w kontekście wyjaśnianych metod sztucznej inteligencji. Rozprawa prezentuje dobry poziom naukowy i zasługuje na ocenę pozytywną.

#### 4. Uwagi i problemy do dyskusji

1. W procesie konstruowania drzew decyzyjnych wskazano jedynie kryterium podziału (Gini) w pakiecie scikit-learn, natomiast brakuje informacji o pozostałych parametrach indukcji drzewa (np. maksymalna głębokość, minimalna liczba próbek w węźle). Jakie ustawienia przyjęto i czy badano ich wpływ na uzyskane wyniki?
2. W pracy przyjęto określone wartości parametrów dla poszczególnych metod indukcji reguł, mimo że ich dobór może istotnie wpływać na wyniki. Czy rozważano zastosowanie strategii przeszukiwania hiperparametrów zamiast ustalania ich „na sztywno”? Na przykład, dlaczego podczas wyszukiwania zbiorów częstych ograniczono się jedynie do trzech wartości parametru wsparcia (0.1, 0.15 oraz 0.2)?
3. W treści zabrakło odniesienia do istotnych własności miary wsparcia dla zbiorów częstych, w szczególności jej antymonotoniczności, która umożliwia efektywne ograniczanie przestrzeni poszukiwań. Uwzględnienie tego aspektu pozwoliłoby lepiej uzasadnić przyjęte podejście oraz jego efektywność obliczeniową.

4. W pracy zastosowano algorytm FP-Growth do generowania zbiorów częstych, jednak nie uzasadniono wyboru tej metody. Biorąc pod uwagę, że istnieją również inne podejścia, warto wyjaśnić, jakie przesłanki przemawiały za wyborem FP-Growth oraz czy porównywano go z alternatywnymi algorytmami pod względem efektywności lub jakości uzyskanych wyników?
5. Uzasadnienie, że porównanie wyłącznie z RuleKit jest wystarczające ze względu na wcześniejsze porównania tej metody z innymi, wydaje się niewystarczające. Pośrednie odniesienie do wyników z innych prac nie zastępuje bezpośredniego porównania w tych samych warunkach eksperymentalnych, co w pewnym stopniu utrudnia pełniejszą ocenę skuteczności proponowanego podejścia.
6. Zabrakło w pracy jasnego powiązania między zasadą MDL (ang. minimum description length) a kwestią zwięzłego opisu wiedzy, w szczególności ich wspólnego dążenia do minimalizacji opisu i uchwycenia regularności występujących w danych.
7. W pracy stosowane jest niespójne nazewnictwo – ta sama metoda określana jest zarówno jako ComplexConditions, jak i ComplexRules. Warto ujednolicić terminologię i konsekwentnie stosować jedną nazwę w całym tekście, aby uniknąć niejednoznaczności, np.: na str. 18 „Metoda ComplexRules skupia się na generowaniu reguł z bardziej złożonymi formami warunków wewnątrz przesłanek”, tytuł rozdziału 6 to: Metoda indukcji reguł z warunkami złożonymi – ComplexConditions, na str. 212 „Każda z trzech metod – ComplexConditions, MofNRules oraz DeepRules...”

#### Uwagi redakcyjne:

Praca jest starannie napisana. Występują pewne uchybienia i drobne błędy redakcyjne wymienione poniżej.

- brak objaśnień we wzorach 6.17, 6.18, 6.19, str. 103.
- w jednym miejscu występują podpisy nad tabelą, w innym pod tabelą np. Tabela 6.6 oraz Tabela 7.11.
- tytuł podrozdziału Posumowanie, str. 113.
- „Z użyciem zbioru D” – drugi element rys. 6.3, str. 88
- takich formuł logiczny – str. 42
- „eksperta . Może” – str. 43
- „Backpropagation). [196].” – str. 53
- „z swej pracy” – str. 39
- „wadze (linia 20).” – str. 73 - brak wskazania że chodzi o Algorytm 4
- „posiada on jednak kilka dodatkowych, pozwalających” – str. 74
- „kroku 6. mogą” – str. 92

## 5. Wnioski

Pan mgr inż. Cezary Maszczyk przedstawił rozprawę doktorską stanowiącą oryginalne rozwiązanie problemu naukowego z zakresu opracowania metod indukcji reguł decyzyjnych zawierających złożone warunki elementarne, w kontekście eksploracji danych i dyskretnego głębokiego uczenia. Opracowane metody zostały szeroko zweryfikowane w sposób eksperymentalny, a wyniki dotyczące predykcji oraz reprezentacji wiedzy zostały porównane z modelami referencyjnymi. Autor wykazał się wiedzą w zakresie tematyki rozprawy oraz umiejętnością pracy naukowej. Recenzowana rozprawa doktorska spełnia ustawowe wymagania stawiane kandydatom do stopnia naukowego doktora, określone w art. 187 ust. 1 i 2 Ustawy z dnia 20 lipca 2018r. Prawo o szkolnictwie wyższym i nauce (Dz.U. 2024 r. poz. 1571, z późn. zm.). Wniosuję do Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej o dopuszczenie jej do publicznej obrony.