

Kraków, 09.02.2026 r.

dr hab. inż. Edyta Kucharska, prof. AGH
Akademia Górniczo-Hutnicza w Krakowie
Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej

RECENZJA

rozprawy doktorskiej mgr inż. Marka Hermansy pt. Interpretowalne metody analizy niezawodności i przeżycia z zastosowaniem do predykcyjnego utrzymania ruchu

Recenzja została przygotowana w wyniku uchwały Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej z dn. 28.10.2025 r., powołującej mnie na recenzenta ww. rozprawy.

Promotorem rozprawy jest dr hab. Beata Sikora, prof. PŚ, a Promotorem pomocniczym dr inż. Łukasz Wróbel.

1. Cel i tezy pracy

Rozprawa Pana mgr inż. Marka Hermansy dotyczy aktualnego i ważnego problemu interpretowalnych metod analizy niezawodności i przeżycia dla danych cenzurowanych w kontekście zastosowania w predykcyjnym utrzymaniu ruchu oraz analizie danych medycznych. Obecność obserwacji cenzurowanych, czyli takich, dla których moment wystąpienia zdarzenia nie jest dokładnie znany, powoduje trudności w efektywnym zastosowaniu klasycznych metod analizy zagadnień z kompletnymi danymi (np. klasyfikacja, regresja,) ze względu na możliwą utratę istotnych informacji w rozpatrywanym zagadnieniu oraz może wpływać na zniekształcenie wyników analizy. Dlatego potrzebne było zastosowanie specjalistycznych metod analizy. Inny zagadnieniem było skuteczne wykorzystanie wyników metod analizy danych, zwłaszcza w odniesieniu do wyjaśnień modeli typu „czarna skrzynka”, w taki sposób, aby były zrozumiałe dla człowieka bez potrzeby dodatkowych wyjaśnień. Odpowiednie zrozumienie przez decydentów wyników analiz jest istotnym elementem procesu podejmowania kluczowych decyzji dla organizacji.

W przedstawionej do recenzji rozprawie doktorskiej podana została luka badawcza w zakresie interpretowalnych metod uczenia maszynowego dedykowanych do analizy niezawodności i przeżycia. Wskazano na brak metod łączących dobrą jakość predykcji z pełną interpretowalnością w zadaniach czas-do-zdarzenia.

Teza rozprawy została sformułowana (rozdział 1.3, str. 13) następująco:

Zastosowanie interpretowalnych metod uczenia maszynowego, w szczególności algorytmów indukcji reguł dostosowanych do specyfiki danych cenzurowanych, umożliwia opracowanie skutecznych i transparentnych narzędzi analizy niezawodności i przeżycia o szerokim zakresie zastosowań, czego potwierdzeniem jest ich potencjalna użyteczność w medycynie oraz w predykcyjnym utrzymaniu ruchu.

Jako główny cel rozprawy podano opracowanie interpretowalnych metod analizy niezawodności i przeżycia oraz przedstawienie ich potencjalnej użyteczności w zastosowaniach medycznych oraz w predykcyjnym utrzymaniu ruchu. W ramach realizacji celu głównego sformułowano cele szczegółowe obejmujące rozszerzenie podstaw integracji analizy przeżycia z indukcją reguł poprzez ujednolicenie definicji reguł oraz kryteriów oceny i agregacji predykcji w modelach i zespołach, rozwój algorytmu indukcji przeżyciowych reguł akcji wraz z opartym na nich algorytmem rekomendacji, opracowanie algorytmu przeżyciowych reguł wyjątków do identyfikacji nietypowych wzorców w danych cenzurowanych, stworzenie interpretowalnego zespołu reguł przeżyciowych łączącego zalety podejść zespołowych i modeli regułowych, przeprowadzenie walidacji metod na zróżnicowanych zbiorach danych przemysłowych i medycznych oraz pokazanie ich użyteczności na przykładach zastosowań medycznych i przemysłowych.

W mojej opinii Doktorant poprawnie określił zakres badań i tezy pracy, a także podał przykładowy zakres praktycznego wykorzystania wyników badań.

2. Struktura rozprawy doktorskiej

Przedstawiona do recenzji rozprawa zawiera 183 strony, w tym m.in. 3 strony spisu treści, 153 strony tekstu zasadniczego, 18 stron bibliografii (153 pozycji literaturowych, w tym 2 pozycje, w których Doktorant jest współautorem). W ramach pracy umieszczono również spisy skrótów, rysunków, tabel i algorytmów.

Rozprawa posiada uporządkowaną strukturę. Zasadnicza część rozprawy składa się z siedmiu rozdziałów, w tym wprowadzenia i podsumowania. W kolejnych rozdziałach omówiono podstawy teoretyczne, założenia metodologiczne, rozwiązania algorytmiczne oraz potencjalne zastosowania interpretowalnych metod analizy niezawodności i przeżycia w medycynie oraz w predykcyjnym utrzymaniu ruchu.

We wprowadzeniu (Rozdział 1) Doktorant wskazał, dlaczego tematyka badań z zakresu rozprawy jest ważna i aktualna. Przedstawia motywację podjętych badań, formułuje tezę, główny cel rozprawy i sześć celi szczegółowych oraz zakres pracy. Wskazuje zagadnienia, która stanowią elementy oryginalne rozprawy i jego wkład naukowy.

Rozdział drugi przedstawia tło i zakres wykorzystania analizy przeżycia i niezawodności w środowisku przemysłowym. W szczególności uwzględniono predykcyjne utrzymanie ruchu bazujące na analizie danych typu czas-do-zdarzenia. Omówiono cztery stosowane strategie utrzymania ruchu (strategia do wystąpienia uszkodzenia, strategia prewencyjnego utrzymania ruchu, strategia utrzymania na podstawie stanu technicznego, strategia predykcyjnego utrzymania ruchu) i wykorzystywane technologie. Scharakteryzowano dane w predykcyjnym utrzymaniu ruchu, które odzwierciedlają specyfikę procesów przemysłowych i warunków pracy urządzeń, ze

szczególnym uwzględnieniem cenzurowania obserwacji. Podane zostały przykłady zastosowań metod przeżyciowych w przemyśle i medycynie.

Rozdział trzeci stanowi część teoretyczną pracy i wprowadza podstawowe pojęcia analizy niezawodności oraz analizy przeżycia. Przedstawiono w nim funkcje i rozkłady statystyczne wykorzystywane w modelowaniu czasu do zdarzenia, metody estymacji oraz klasyczne modele przeżycia, takie jak estymator Kaplana–Meiera, model proporcjonalnych hazardów Coxa czy parametryczne modele czasu życia. Zaprezentowano również przykłady zastosowań tych metod w inżynierii niezawodności, co stanowi punkt wyjścia dla dalszych rozważań i badań prezentowanych w pracy.

Rozdział czwarty poświęcono zagadnieniom interpretowalności i objaśnialności modeli uczenia maszynowego, szczególnie w kontekście zastosowań przemysłowych i medycznych. Podano definicje interpretowalności i objaśnialności modeli. Omówiono różnice pomiędzy modelami interpretowalnymi a rozwiązaniami typu „czarna skrzynka”. Scharakteryzowano metody objaśnialnego uczenia maszynowego: globalne (analiza ważności cech, wykresy zależności częściowej: służące zrozumieniu zachowania modelu przeżycia na całym zbiorze danych) i lokalne (LIME - Local Interpretable Model-agnostic Explanations, SHAP - SHapley Additive exPlanations, wyjaśnienia kontrfaktyczne - służące wyjaśnianiu pojedynczych predykcji). Ponadto scharakteryzowano cztery główne kategorie interpretowalnych metod uczenia maszynowego: drzewa decyzyjne, reguły logiczne, modele regresyjne oraz metody oparte na instancjach.

Rozdział piąty koncentruje się na regułach logicznych jako interpretowalnej metodzie analizy danych cenzurowanych i predykcji zdarzeń. Zawiera przegląd metod indukcji reguł i zespołów reguł, a także przedstawia podstawy integracji algorytmów regułowych z analizą przeżycia. Ponadto, znaczna część rozdziału dotyczy autorskiego wkładu naukowego. Doktorant zaprezentował cztery autorskie algorytmy opracowane w ramach rozprawy, które zostały opisane formalnie oraz zilustrowane pseudokodem: pokryciowy algorytm indukcji przeżyciowych reguł akcji, algorytm rekomendacji przeżyciowych reguł akcji, algorytm indukcji przeżyciowych reguł wyjątków i metoda interpretowalnego zespołu reguł przeżyciowych.

Rozdział szósty prezentuje wyniki eksperymentów oraz studia przypadków stanowiących empiryczną weryfikację proponowanych metod. Podano kryteria oceny efektywności zaproponowanych algorytmów. Scharakteryzowano zbiory danych medycznych i przemysłowych, na których przeprowadzono badania. Przedstawiono uzyskane rezultaty dla czterech zaproponowanych algorytmów oraz porównano z metodami referencyjnymi (SR survival rules, ST survival tree, RSF random survival forest, regularyzowana regresja Coxa z karą elastic net, estymator Kaplana-Meiera).

Ostatni siódmy rozdział zawiera podsumowanie pracy, w którym zestawiono najważniejsze wyniki i osiągnięcia, omówiono potencjalne obszary zastosowań opracowanych metod w medycynie i predykcyjnym utrzymaniu ruchu oraz wskazano kierunki dalszych badań. Przedstawiono również ograniczenia zaproponowanych rozwiązań oraz możliwe kierunki ich rozwoju.

3. Uwagi krytyczne

Rozprawa napisana jest poprawnym językiem, styl prowadzonego wywodu odpowiada opracowaniu naukowemu. Kolejność rozdziałów ułatwia śledzenie toku rozumowania autora. W bibliografii pracy wykorzystano zarówno pozycje starsze, stanowiące fundament metod analizy danych cenzurowanych (m.in. modele Coxa, estymator Kaplana-Meiera, survival forests), jak i aktualne pozycje literatury związane z podejmowaną problematyką. Tematyka cytowanych prac jest ściśle związana z obszarami badawczymi podjętymi w rozprawie doktorskiej. W bibliografii uzupełnienia o najnowsze publikacje wymaga obszar zagadnień XAI (Explainable Artificial Intelligence) i praktycznych narzędzi uczenia maszynowego dla przeżycia.

Na podstawie lektury przedstawionej rozprawy doktorskiej nasuwają się następujące uwagi o charakterze ogólnym i edytorskim.

3.1. Uwagi ogólne

- Pewne wątpliwości może budzić użycie pojęcia „narzędzia” w tezie rozprawy, które sugeruje rozwiązania aplikacyjne, podczas gdy zasadniczy wkład pracy ma charakter metodologiczny i dotyczy przede wszystkim rozwoju algorytmów oraz metod analizy.
- W kilku miejscach powracają podobne uzasadnienia dotyczące znaczenia interpretowalności modeli, co niekiedy wydłuża tekst bez istotnego zwiększenia jego wartości informacyjnej.
- W zaproponowanych algorytmach wykorzystano strategię wspinaczki (ang. hill climbing) do iteracyjnego przeszukania przestrzeni meta-przykładów w mD i zidentyfikowania zestawu zmian atrybutów prowadzący do pożądanej modyfikacji krzywej przeżycia. Zastosowanie strategii wspinaczki upraszcza proces przeszukiwania przestrzeni rozwiązań, jednak metoda ta jest podatna na zatrzymywanie się w lokalnych optimum. Czy zastosowanie tej strategii nie ogranicza możliwości odnalezienia lepszych, globalnych rozwiązań, szczególnie w przypadku silnie niejednorodnej przestrzeni meta-przykładów?
- Przeprowadzone badania walidacyjne bazują na 27 zbiorach danych, przy czym 11 dotyczy zbiorów przemysłowych i 18 zbiorów medycznych. Choć wykorzystano dostępne dane, nadal widoczny jest brak szerokiego zestawu benchmarków dla tej klasy problemów, co ogranicza możliwość pełnej generalizacji wyników. Część wykorzystanych do badań zbiorów danych bazuje danych symulacyjnych. Czy zestaw danych jest wystarczająco reprezentatywny dla ogólnych wniosków dotyczących efektywności zaproponowanych algorytmów? Czy wyniki nie są częściowo zależne od specyfiki wybranych domen, a zastosowanie danych syntetycznych w części eksperymentów nie ogranicza możliwość pełnego przeniesienia wyników na środowiska produkcyjne?
- Czy w wybranych do badaniach danych rozpatrywano różne warunki środowiskowe, czy założono niezmiennie warunki?
- W badaniach zastosowano etap wstępnego przetwarzania danych (str. 117). Warto doprecyzować, jaki wpływ miał on na uzyskane rezultaty, w szczególności czy przeprowadzono analizę wrażliwości wyników na przyjęte procedury przygotowania danych. Istotne byłoby również wskazanie, w jakim stopniu zastosowane operacje mogły wpłynąć na jakość generowanych reguł oraz stabilność uzyskiwanych modeli.

- Widoczne są znaczne różnice w uzyskanych wynikach pomiędzy danymi przemysłowymi a medycznymi, co sugeruje, że skuteczność proponowanych metod może zależeć od charakterystyki analizowanej domeny. W pracy warto byłoby przeprowadzić pogłębioną analizę przyczyn tych rozbieżności, uwzględniając m.in. różnice w strukturze danych, poziomie cenzurowania, liczbie obserwacji, jakości pomiarów czy dynamice procesów prowadzących do zdarzenia. Pozwoliłoby to lepiej określić zakres stosowalności proponowanych metod oraz warunki, w których osiągają one największą skuteczność.
- Złożoność obliczeniowa proponowanych metod nie została szczegółowo przeanalizowana (str. 157: „Analiza złożoności obliczeniowej nie była przedmiotem niniejszej rozprawy”), a w kontekście potencjalnych zastosowań przemysłowych jest to istotny aspekt. Przedstawione ilustracje działania poszczególnych algorytmów w rozdziale 5 (podrozdziały 5.2.3, 5.3.2, 5.4.3, 5.5.3) dotyczą prostych przykładów. Wybór takich ilustracji jest dobry ze względu na łatwość zrozumienia proponowanego podejścia. Ponadto w rozdziale 6 przedstawione są studia przypadków, jednak tylko po jednym dla proponowanych algorytmów, co nie pokazuje ewentualnych problemów i ograniczeń dla różnych przykładów praktycznych. Czy zatem zaproponowane algorytmy skalują się do bardzo dużych zbiorów danych? Czy mogą być powszechnie stosowane w systemach przemysłowych?

3.2. Uwagi polemiczne i dyskusyjne

- W zastosowaniach praktycznych w medycynie i przemyśle często wykorzystuje się modele o charakterze „czarnej skrzynki”, jeżeli oferują wyraźnie lepszą jakość predykcji bez konieczności dokładnego wyjaśniania sposobu działania metody. Doktorant przyjmuje, że interpretowalność modeli jest kluczowa w analizie przeżycia. Czy dążenie do interpretowalności powinno być zawsze nadrzędne wobec skuteczności predykcyjnej?
- W rozprawie podano ogólne wzory i parametry dla modeli regułowych (np. wzór 5.1, 5.2, 5.5 itd.). W praktyce jednak duże zbiory reguł stają się trudne do analizy przez użytkownika. Czy wraz ze wzrostem liczby reguł model pozostaje interpretowalny i gdzie leżą granice praktycznego stosowania interpretowalności?
- W zaproponowanym podejściu reguły opisują współwystępowanie określonych konfiguracji atrybutów i sugerowanych akcji. Natomiast w zastosowaniach medycznych i technicznych często oczekuje się wskazań przyczynowych, a nie tylko proponowanie rekomendacji bez wskazania przyczyn. Jak można uwzględnić takie oczekiwania?
- Jaka jest możliwość praktycznego zastosowania uzyskanych wyników? Jak wyglądałaby integracja proponowanych metod z istniejącymi systemami przemysłowymi lub klinicznymi? Czy użytkownicy końcowi są w stanie efektywnie korzystać z generowanych reguł?
- Doktorant w podsumowaniu wskazuje krótko na ograniczenia przeprowadzonych badań m.in. koncentracja na wybranych klasach reguł: regułach akcji, regułach wyjątków oraz zespole reguł (str. 175, Rozdział 7 „Podsumowanie”). W jakich sytuacjach proponowane metody nie powinny być stosowane? Z drugiej strony jakie są inne obszary zastosowań proponowanych algorytmów poza zastosowaniami w predykcyjnym utrzymaniu ruchu oraz analizie danych medycznych?

3.3. Uwagi edytorskie

Rozprawa jest przygotowana na wysokim poziomie edytorskim. W tekście rozprawy wystąpiły jedynie nieliczne niedociągnięcia, jak przykładowo:

- Błędy stylistyczne, przykładowo (np. str. 44 „Do najważniejsze metody w tej kategorii przedstawiono poniżej.”, str. 52 „Reguły w ramach zbioru są niezależne, dzięki czemu można ją analizować konieczności rozumienia całego modelu.”, str. 111.
- Umieszczenie rysunków i tabel zawierających skróty i oznaczenie wielkości znacznie przed ich wyjaśnieniem, znacznie utrudnia ich szybką interpretację.
- Rysunek 5.2 – ucięte słowo „docelowej”
- str. 68. wprowadzenie zbioru danych *Maintenance*, a wskazanie źródła dopiero na str. 113
- Kilka razy brakuje spacji.
- Rysunek 6.1 Krzywe KM dla przemysłowych zbiorów danych z Tabeli 6.1 pokazuje wyniki w różnej skali (inna dla dwóch zbiorów). Choć w tekście zawarto wyjaśnienia to warto je również umieścić w podpisie rysunku.
- W kilku tabelach zostały pominięte wyniki, np. brak zbioru *Aircraft* w tabeli 6.15. Brak taki został uzasadniony. Jednak informacja o pominięciu zbioru danych oprócz tekstu powinna znaleźć się również w opisie tabeli.
- Spis literatury – opis bibliografii powinien być ujednolicony, np. w niektórych miejscach podawani są wszyscy autorzy (nazwisko i inicjał imion), a w innych tylko pierwszy autor. Przykładem jest pozycja [117] Badura, J. i in. „Separate-and-Conquer Survival Action Rule Learning”. W: *Knowledge-Based Systems* 280 (list. 2023), s. 110981. issn: 09507051. Doi: 10.1016/j.knosys.2023.110981. (Term. wiz. 22. 09. 2024). W pozycji tej, podanej w taki sposób nie widnieje nazwisko Doktoranta, co jest istotne dla recenzowanej rozprawy.
- W bibliografii występują dwie pozycje [14] Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. 3:e upplagan. Christoph Molnar, 2025. isbn: 978- 3-911578-03-5. oraz [90] Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Second edition. Christoph Molnar, 2022. 1 s. isbn: 979-8-4114-6333-0. Wyglądają one jak ten sam tytuł “Survival Analysis: Techniques...” w dwóch wydaniach/rekordach — warto ujednolicić lub uzasadnić wybór obu.

4. Ogólna ocena rozprawy doktorskiej

W przedstawionej rozprawie doktorskiej wskazana została zidentyfikowana luka badawcza w obszarze interpretowalnych metod uczenia maszynowego dla danych typu czas-do-zdarzenia z cenzurowaniem. Autor podejmuje aktualny i istotny problem badawczy, dotyczący transparentności modeli decyzyjnych, zwłaszcza w zastosowaniach takich jak medycyna czy utrzymanie ruchu. Interpretowalność modeli zwiększa ich potencjalną akceptację w środowiskach eksperckich. Praca wpisuje się w rozwijający się nurt interpretowalnego uczenia maszynowego (Explainable AI).

Oryginalność badań przeprowadzonych przez Doktoranta polega na opracowaniu czterech autorskich algorytmów, które zostały opisane formalnie oraz zilustrowane pseudokodem:

- pokryciowy algorytm indukcji przeżyciowych reguł akcji, który koncentruje się na generowaniu przeżyciowych reguł akcji na podstawie danych treningowych z wykorzystaniem strategii sekwencyjnego pokrywania,
- algorytm rekomendacji przeżyciowych reguł akcji, traktujący wygenerowane reguły jako dane wejściowe do procesu tworzenia spersonalizowanych zaleceń dla nowych obiektów,
- algorytm indukcji przeżyciowych reguł wyjątków, bazujący na strategii *separate-and-conquer*, w której proces wzrostu reguły bazowej jest rozszerzony o wyszukiwanie wyjątków,
- metoda interpretowalnego zespołu reguł przeżyciowych, łącząca interpretowalność modeli opartych na regułach z dokładnością technik uczenia zespołowego, która rozszerza klasyczny paradygmat lasów losowych, zastępując drzewa decyzyjne modelami reguł przeżyciowych, dostosowanymi do analizy danych cenzurowanych.

Skuteczność zaproponowanych rozwiązań została zweryfikowana na zróżnicowanych zbiorach danych, odzwierciedlających problemy analizy przeżycia i niezawodności, a uzyskane rezultaty porównano z odpowiednimi metodami referencyjnymi, takim jak przykładowo dla proponowanej metody interpretowalnego zespołu reguł przeżyciowych: SR — indukcja reguł przeżyciowych z estymacją funkcji przeżycia, ST — pojedyncze drzewo przeżycia z podziałami opartymi na statystyce log-rank i estymacją Kaplana-Meiera w liściach, RSF — losowy las drzew przeżycia z bootstrapowaniem próbek i losowym, Cox — regularyzowana regresja Coxa z karą *elastic net*) oraz estymator Kaplana-Meiera.

Wyniki badań przedstawione przez Autora rozprawy wykazały użyteczność zastosowania interpretowalnych metod uczenia maszynowego, w szczególności algorytmów indukcji reguł dostosowanych do specyfiki danych cenzurowanych do opracowanie algorytmów analizy niezawodności i przeżycia i ich potencjalnej użyteczności w medycynie oraz w predykcyjnym utrzymaniu ruchu. Tym samym potwierdziły tezę rozprawy.

Wniosek końcowy

Podsumowując, Doktorant podejmuje problem łączenia jakości predykcji z możliwością wyjaśnienia decyzji w obszarach wymagających transparentności. Zaproponowane algorytmy rozwijają istniejące podejścia regułowe i adaptują je do specyfiki danych cenzurowanych, co stanowi nową jakość w literaturze przedmiotu. Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinę. Doktorant wykazał się umiejętnością samodzielnego prowadzenia pracy naukowej i przedstawił oryginalne rozwiązanie problemu naukowego dotyczącego interpretowalnych metod analizy niezawodności i przeżycia dla danych cenzurowanych, a uzyskane wyniki można użyć w zastosowaniach medycznych oraz w predykcyjnym utrzymaniu ruchu. Jednak według mnie konieczne jest przebadanie w dalszych badaniach innych zbiorów danych z innych obszarów, aby potwierdzić ogólność proponowanych rozwiązań.

Mimo przedstawionych uwag krytycznych stwierdzam, że rozprawa doktorska mgr inż. Marka Hermansy pt. „Interpretowalne metody analizy niezawodności i przeżycia z zastosowaniem do predykcyjnego utrzymania ruchu” **spełnia** warunki stawiane rozprawom doktorskim, określone w *Ustawie z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce* (Dz. U. 2023 poz. 742).

Zgodnie z dokonaną pozytywną oceną rozprawy **wnioskuję o przyjęcie** rozprawy doktorskiej mgr inż. Marka Hermansy i dopuszczenie jej do publicznej obrony w dyscyplinie *informatyka techniczna i telekomunikacja*.

Edyta
Kucharska

Elektronicznie
podpisany przez Edyta
Kucharska
Data: 2026.02.16
16:27:47 +01'00'