

Interpretowalne metody analizy niezawodności i przeżycia z zastosowaniem do predykcyjnego utrzymania ruchu

Marek Hermansa

Promotor: dr hab. Beata Sikora, prof. PŚ
Promotor pomocniczy: dr inż. Łukasz Wróbel

Streszczenie rozprawy doktorskiej

Tematem pracy są interpretowalne metody analizy niezawodności i przeżycia dla danych cenzurowanych typu czas-do-zdarzenia. Powszechnie stosuje się zarówno klasyczne, interpretowalne metody statystyczne (np. estymator Kaplana-Meiera, model proporcjonalnych hazardów Coxa), jak i bardziej złożone techniki uczenia maszynowego, w tym drzewa i ich zespoły oraz modele typu *black box*. W literaturze brakuje jednak podejść, które jednocześnie wprost uwzględniają cenzurowanie, zachowują interpretowalność procesu decyzyjnego i oferują dobrą skuteczność predykcyjną. Problem ten ma znaczenie w zastosowaniach o podwyższonych wymaganiach (np. medycyna, predykcyjne utrzymanie ruchu), gdzie transparentność i możliwość weryfikacji podstaw decyzji są istotne dla wdrożenia modeli.

Celem pracy było opracowanie interpretowalnych metod analizy niezawodności i przeżycia opartych na regułach logicznych, dostosowanych do specyfiki danych cenzurowanych, oraz ocena ich potencjalnej użyteczności w medycynie i w predykcyjnym utrzymaniu ruchu. W pracy zaproponowano cztery oryginalne algorytmy: pokryciowy algorytm indukcji przeżyciowych reguł akcji, algorytm rekomendacji oparty na przeżyciowych regułach akcji, algorytm przeżyciowych reguł wyjątków oraz interpretowalny zespół reguł przeżyciowych.

Praca składa się z siedmiu rozdziałów. Rozdział 1 przedstawia wprowadzenie, lukę badawczą, cele i tezę. Rozdział 2 omawia kontekst zastosowań w przemyśle oraz specyfikę danych w predykcyjnym utrzymaniu ruchu. Rozdział 3 stanowi część teoretyczną i przedstawia pojęcia i metody analizy niezawodności i przeżycia. Rozdział 4 omawia interpretowalność i objaśnialność w uczeniu maszynowym. Rozdział 5 przedstawia reguły logiczne dla danych cenzurowanych oraz cztery autorskie algorytmy. Rozdział 6 prezentuje wyniki eksperymentów i studium przypadku dla każdej z metod. Rozdział 7

przedstawia wnioski, wkład i kierunki dalszych badań.

Pokryciowy algorytm indukcji przeżyciowych reguł akcji adaptuje strategię *separate-and-conquer* do danych cenzurowanych. Maksymalizuje rozbieżność krzywych przeżycia grup spełniających i niespełniających regułę, kontroluje pokrycie i nakładanie pokryw oraz ogranicza złożoność zbioru reguł. Algorytm rekomendacji przekłada przeżyciowe reguły akcji na zalecenia dla pojedynczych obiektów, wykorzystując meta-tablicę i mechanizm rozstrzygania konfliktów między regułami.

Algorytm przeżyciowych reguł wyjątków definiuje trójskładnikową strukturę: regułę bazową (CR), regułę referencyjną (RR) oraz regułę wyjątku (ER). Reguła bazowa opisuje typowy wzorec w danych. Reguła RR jest dobierana tak, aby jej krzywa Kaplana-Meiera nie różniła się istotnie od CR, co pozwala odfiltrować pozorne wyjątki. Istotne różnice ujawniają się dopiero dla reguły wyjątku (ER), tj. koniunkcji $CR \wedge RR$, która odznacza się inną krzywą przeżycia niż pozostałe reguły (weryfikacja w parach CR-ER, RR-ER testem log-rank). Tak sformułowana procedura redukuje liczbę pozornych wyjątków i dostarcza uzasadnienie, dlaczego reguła jest wyjątkiem.

Interpretowalny zespół reguł przeżyciowych agreguje wiele modeli regułowych, redukując wariancję predykcji przy zachowaniu przejrzystości procesu decyzyjnego na poziomie reguł. Umożliwia śledzenie aktywacji reguł oraz ważności zmiennych objaśniających. Ograniczeniem zespołu jest brak skalarne-go wskaźnika ryzyka, co utrudnia ocenę zdolności dyskryminacyjnej, oraz wzrost złożoności wraz z liczbą estymatorów bazowych.

Walidację empiryczną przeprowadzono na 27 zbiorach (medycznych i przemysłowych). Dla zespołów oceniano zintegrowany wskaźnik Briera (IBS), dla reguł akcji i wyjątków test log-rank oraz metryki struktury i pokrycia, a dla algorytmu rekomendacji spójność S_α , błąd średniokwadratowy (MAE) i pokrycie. Zbiory reguł akcji dawały wysoki odsetek istotnych reguł (powyżej 95% reguł dla poziomu istotności 0.01 w przypadku większości zbiorów); rekomendacje osiągały wskaźnik spójności $S_{0.05}$ rzędu 0.70–1.00, MAE 0.04–0.35 i pokrycie 90–100%; reguły wyjątków wykryto w 16 z 27 zbiorów; zespół reguł uzyskał IBS porównywalny z RSF i Cox oraz w 96% przypadków niższy niż KM, szczególnie dla zbiorów o wysokim cenzurowaniu.

Wyniki łącznie potwierdzają tezę, że zaproponowane w pracy interpreto-walne algorytmy regułowe dostosowane do danych cenzurowanych stanowią skuteczne i transparentne metody analizy niezawodności i przeżycia, o potencjalnej użyteczności w praktyce medycznej i przemysłowej.