

COMPETING RISK ANALYSIS - GRAPHICAL REPRESENTATION OF VARIABLE INFLUENCE

Małgorzata Krętowska

Faculty of Computer Science, Białystok University of Technology, Białystok, Poland

Abstract: In the paper the possibilities of assessing the variable influence on the failure occurrence is shown. Ensemble of dipolar survival trees is used as a prediction tool. The technique is able to cope with censored data (data with incomplete observations) as well as with competing risks data. The results are presented on the base of two real datasets for which the influence of discrete and continuous variables is examined. To this purpose, the cumulative incidence functions and the quartiles of CIF functions are applied.

Keywords: competing risks, survival analysis, ensemble of survival trees, dipolar criterion

DANE Z KONKURENCYJNYM RYZYKIEM - GRAFICZNA REPREZENTACJA WPŁYWU CZYNNIKÓW RYZYKA

Streszczenie W pracy przedstawione zostały możliwości graficznej weryfikacji hipotez dotyczących wpływu poszczególnych cech na czas wystąpienia porażki. Jako narzędzie prognostyczne zostały wykorzystane predyktory złożone, w których dipolowe drzewa przeżycia służą jako pojedyncze predyktory. Algorytm tworzenia predyktorów złożonych wykorzystuje informację pochodzącą z obserwacji cenzorowanych, jak również jest przystosowany do danych z konkurencyjnym ryzykiem.

Eksperymenty zostały wykonane przy użyciu dwóch zbiorów danych: zbiór opisujący pacjentki z rakiem piersi i drugi - opisujący pacjentów z chłoniakiem grudkowym. Pierwszy z analizowanych zbiorów posłużył jako przykład do badania wpływu zmiennych dyskretnych. W tym celu wyznaczone zostały dystrybuanty (ang. cumulative incidence function) dla wyróżnionych dwóch zdarzeń konkurencyjnych i dwóch cech: rodzaju leczenia oraz typu histologicznego raka. W przypadku zbioru z chłoniakiem grudkowym badane były cechy ciągłe: wiek oraz wartość hemoglobiny. Analiza tych danych opierała się na wyznaczeniu wartości kwartyła pierwszego oraz mediany z funkcji dystrybuanty, wyznaczonej dla czasu nawrotu choroby.

Słowa kluczowe: dane z konkurencyjnym ryzykiem, analiza przeżyć, predyktory złożone, kryterium dipolowe

Artykuł zrealizowano w ramach pracy badawczej S/WI/2/08.