Abstract

Omneya Amr Attallah

Handling High Level of Censoring For Endovascular Aortic Repair Prediction

Censoring is the unique characteristic in survival data analysis. It is the reason why standard machine learning technique cannot be used directly with survival data. Feature Selection is needed in medical field, however this process becomes complex with censoring. This paper proposed an approach to deal with the censoring problem in endovascular aortic repair survival data through Bayesian networks. It was merged and embedded with a hybrid feature Selection process that combines cox's univariate analysis with machine learning approaches such as ensemble artificial neural networks to Select the most relevant predictive variables. The proposed algorithm was compared with common survival variable Selection approaches such as least absolute shrinkage and Selection operator, and Akaike information criterion methods. The results showed that it was capable of dealing with high censoring in the datasets. Moreover, ensemble classifiers increased the area under the roc curves of the two datasets collected from two centers located in United Kingdom separately. Furthermore, ensembles constructed with center 1 enhanced the concordance index of center 2 prediction compared to the model built with a single neural network. Although the size of the final reduced model using the neural networks and its ensembles is greater than other methods, the model outperformed the others in both concordance index and sensitivity for center 2 prediction. This indicates the reduced model is more powerful for cross center prediction.