Abstract

In the healthcare system, consider having a system that could provide information regarding the survival time for a particular patient. Such as, in the case of cancer, if we have the risk prediction of a particular patient surviving for X days, suitable measured can be taken into account to improve the patient's health. This information can be obtained using Survival analysis. However, we need a complete dataset to have accurate results to perform the survival analysis. Unfortunately, in the healthcare system, obtaining the result for each patient undergoing study is very difficult because individuals tend to leave the study for many reasons. Such individuals for whom the exact outcome for a disease is not available are called censored data, accurately right censored data. Hence one of the biggest challenges while obtaining the survival analysis is dealing with censored data.

In this research, we will explore the methods to deal with censored data. Furthermore, obtain their comparative analysis. We will work with IPCW weights and weighted censored instances to check which methods work best to deal with censored data. Along with this, we will also look into the parameter and structure learning for the Bayesian network and see which Bayesian network algorithm is better when modelling such data.

We will present the comparative analysis of both methods to deal with censored data and the Bayesian network algorithms. we will compare the results obtained from each model.

Keywords : Bayesian network, Survival analysis, Inverse probability censoring weight (IPCW), weighted censored instances.