

Prediction of Breast Cancer Survival using Machine Learning Techniques

Ye Yao, Zhilin Zhang, Hengde Ouyang

12/17/2022

Introduction

Breast cancer is the most common cancer among women in the United States. According to the survey, the average risk of a woman in the United States developing breast cancer is about 13%[1]. The survival prediction, especially 5-year survival analysis, was largely used to help doctors understand the prognosis, predict treatment efficiency and develop individualized treatment plans. In previous studies, machine learning was largely applied for its high accuracy and fewer assumption requirements. In this study, the survival prediction model was built based on data with 300998 cases and 9 variables extracted from the SEER research database.

Data Description

Data Preprocessing