

ABSTRACT

Oxygen saturation and respiratory rate are two basic parameters used to assess the patient's condition, especially respiration. Heart failure and COVID-19 are some of the diseases related to these two parameters. Heart failure has specific respiratory symptoms such as sudden chest pains and shortness of breath, which are caused by an abnormality in oxygen saturation and respiratory rate. COVID-19 is a new disease found in 2019, and this disease also has a close relationship with respiratory. If infected, COVID-19 can cause acute respiratory distress syndrome (ARDS), pneumonia, and problems with other body organs, which can cause death for the sufferer. Therefore, these two parameters are very important to determine the patient's respiratory condition. This study aims to build a logistic regression model for classifying the patient's respiratory condition using oxygen saturation and respiratory rate as parameters. Logistic regression is used because of the suitability of the model's advantages with the data and this algorithm can explain the effect of the independent parameters used on the dependent parameter. Then this model will be evaluated using the F1-Macro method. This study uses the CRISP-DM methodology, and prepares data using the downsampling methodology and categorizing values of the variables to get a better model result. The accuracy of the testing model is 87.5%, while the evaluation accuracy using F1-Macro is 87%. This study's results are also already appropriate to existing medical theories regarding oxygen saturation and respiratory rate coefficient values.

Keywords— oxygen saturation, respiratory rate, logistic regression, classification of respiratory condition, downsampling, undersampling