Title: Survival Analysis of Thyroid Cancer Patients Using Machine Learning Algorithms.

Author Name: Saadat M. Alhashmi, Md. Shohidul Islam Polash, Aminul Haque, Fazley Rabbe, Shazzad Hossen, Nuruzzaman Faruqui.

Abstract: The medical community strives continually to improve the quality-of-care patients receive. Predictions of prognosis are essential for doctors and patients to choose a course of treatment. Recent years have witnessed the development of numerous new cancer survival prediction models. Most attempts to predict the prognosis of people with malignant growth rely on classification techniques. We could experiment with significantly different results using only a subset of SEER (Surveillance, Epidemiology, and End Results) data. These models were created using machine learning techniques by selecting univariate features and calculating correlations. We illustrated the variation in results and discrepancy of impurity that can result from varying data quantities and critical factors. Seventeen crucial factors were identified, and a group of classification algorithms were trained to evaluate the effectiveness of an estimation technique. In the display mode, the accuracy of these computations' ranges from 97% to 99% A˙ long with accuracy, the models are further evaluated regarding the F1 score, precision, recall, and the AUC score. Compared to earlier studies, a more accurate model has been developed, and, to the best of our knowledge, our prediction model is superior to the models studied in the previous works.

Keyword: Logistic Regression, Machine Learning, Random Forest, Thyroid Survivability.

DOI: 10.1109/ACCESS.2024.3392275