

Abstract No: SO-03

Estimating determinants of healthcare expenditure in OECD countries

W. A. P. S. Wickramarachchi^{1*}

¹Software Engineering Teaching Unit, Faculty of Science, University of Kelaniya, Sri Lanka
panchaniwickramarachchi@gmail.com*

The COVID-19 pandemic has had a direct impact on increasing the healthcare expenditure of countries across the world. Thus, it is essential to determine the factors that contributed to the expenditure in the healthcare sector previously in order to face the current and future health risks successfully. This research aimed to identify the major variables in the estimation of Healthcare Expenditure (HE) as a share of gross domestic product in Organization for Economic Co-operation and Development (OECD) member countries with the decision tree method and random forest method. 2018 data of 37 OECD countries was comprised as the study population in this study. The HE product was considered the dependent variable. Further, 11 independent variables were defined based on previous studies, such as the Gross Domestic Product (GDP) per capita, percentage of the total population covered by Public and Private Insurance (PPI), Out-of-Pocket (OOP) health expenditure as a percentage of total expenditure on health, Age Dependency Ratio (ADR), Life Expectancy at Birth (LEB), Number of Hospitals (NOH) per million population, Number of Physicians (NOP) per 1000 population/head counts, Pharmaceutical Sales (PS) in USD/per capita (using economy-wide PPPs) and Perceived Health Status categorized into good, bad, and fair (PHSG/F/B). The data were taken from OECD health data and World Bank data repositories. Similar studies done to identify major variables in HE, had used the decision tree method by using different types of algorithms. Furthermore, the previous studies that were conducted considering the variables GDP, HE, public financing, NOP, number of hospital beds per 1000 population, tobacco consumption, life expectancy, and population above 65 years old, have shown that gross domestic product, the population aged 65 years and life expectancy as the most important determinants in health expenditure of OECD countries. In this study according to the fitted decision tree model GDP, PS, NOP, LEB, NOH, and PHSB were identified as the major variables in the estimation of HE. GDP and ADR were identified as the major variables in the random forest method. The precision, recall, F1 – score, accuracy, and the ROC Area Under the Curve (AUC) values were used to compare the performance of the two methods. This study indicates that the random forest model performs better in determining the HE and identified GDP as a major variable from both methods. With the current pandemic situation all around the world, we believe that this evidence-based information will be useful in making the decision for policymakers. However, the factors that affect healthcare expenditure are not limited. Those factors change from time to time, especially with the pandemic situations such as COVID-19. Thus, investigating those using different models would provide more comprehensive information.

Keywords: Data mining, Decision tree method, Health expenditure, OECD, Random-Forest Method