EMPIRICAL ANALYSIS OF PROSTATE CANCER PREDICTION USING MACHINE LEARNING ALGORITHM

MARIZEL B. VILLANUEVA
http://orcid.org/0000-0003-0125-0629
mmvillanueva1@tip.edu.ph
Technological Institute of the Philippines
Manila, Philippines

ABSTRACT

Machine Learning is a branch of Artificial Intelligence (AI) that uses numerous techniques to complete tasks, improving itself after every iteration. A variety of these techniques, including Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs), and Decision Trees (DTs), have been widely applied in cancer research for the development of predictive models, resulting in effective and accurate decision making. According to some research, sixty percent is the accuracy rate of predicting cancer outcomes by pathologists. The main objective of this study is to enhance the accuracy rate of detecting prostate cancer using a machine learning algorithm under the supervised learning category. The methodology of this study consists of data acquisition, data preprocessing, training of data, and validation of test results. Kaggle is used to acquire train data that pertains to prostate cancer. It is considered the world's largest repository of data that and applied in data science researches. Among the supervised algorithms such as J48, Random Forest, Multilayer Perceptron, and Support Vector Machine Algorithm used in the experimental testing, the SVM algorithm garners the highest accuracy rate in diagnosing prostate cancer using limited historical data. The main conclusion of this study is that machine learning algorithms perform better than the accuracy rate compared to physicians and can be used to diagnose prostate cancer.

Keywords: Artificial Intelligence, Machine Learning algorithm, Supervised Learning, Support Vector Machine