ISSN: 1391-8796
Proceedings of 10th Ruhuna International Science & Technology Conference
University of Ruhuna, Matara, Sri Lanka
January 18, 2023

Oral potential malignant disorder prediction using supervised machine learning techniques

Zainab M.N.F.[1*], Gamage C.Y.[1], Liyanage C.R.[1] and Jayasinghe R.D.[2]

[1]*Department of Information Communication Technology, Faculty of Technology, University of Ruhuna, Kamburupitiya, Matara, Sri Lanka.*
[2]*Department of Oral Periodontology, University of Peradeniya, Peradeniya, Sri Lanka.*

Precursor lesions, pre-cancer, intra-epithelial neoplasia, and pre-malignant, are all terms used to refer to Oral Potential Malignant Disorder (OPMD) in the literature to characterize clinical manifestations that have the possibility to develop into Oral Cancer. This is the most common malignancy among Sri Lankan men and has the utmost death rate of all types of malignancies. The main problem with this OPMD which is known as a pre-cancer stage is the risk factors of it could vary according to the region and country because of its nature. Hence, this research will expose the risk factors of OPMD in Sri Lanka, the impact of risk factors, and the transformative potential of OPMD. Around 1000 data samples available at the Faculty of Dental Science, University Peradeniya was collected, preprocessed, and analyzed by handling null values. The Inter Quartile Range method is used for the outlier handling for this identification. Machine learning techniques, such as Random Forest classification, Support Vector Machine, Logistic Regression, and Variance Analysis were used for the identification of risk factors. Moreover, Random Forest, Support Vector Machine, Decision Tree, Logistic Regression, K Nearest Neighbors, Gradient Boosting Tree, and XGP classification algorithms were used to predict the Malignant Transformation Potential. The most important risk factors, the habits of patients and medical history attributes were identified, and the Gradient Boost Classifier, Logistic Regression, and K Nearest Neighbors were identified as best models to predict transformation potential with 97% of accuracy for each.

Keywords: Machine learning, Malignant transformation potential, Oral Potential Malignant Disorder, Pre Cancer, Risk factors.

Acknowledgements: *The authors would like to express appreciation for the support of the research supervisor and wish to thank Dental Surgeon Dr. W.A. Jayantha Karunaratne for help in finding a suitable supervisor.*

Corresponding author: Zainab.nizhan97@gmail.com