APAMI2020 General Oral Presentation Session | APAMI 2020 | General Oral Presentation Session

Artificial Intelligence in Healthcare and Clinical Practice (3)

Mon. Nov 23, 2020 10:30 AM - 12:00 PM Room E-1 (Congress center 5F - Conference Room 52)

[AP3-E1-2-02] Cervical Cancer Screening using Machine Learning Approach

*Sari Rahmawati Kusuma Dewi[1], Jakir Hossain Bhuiyan Masud[1], Emily Chia-Yu Su[1], Ming-Chin Lin[1,2]　(1. College of Medical Science and Technology, Taipei Medical University, Taiwan, 2. Taipei Medical University-Shuang Ho Hospital, Taiwan)

Keywords: Cervical Cancer, Machine Learning, Resample, Imbalance Data, Artificial Intelligence

Cervical cancer (ca cervix) is one of most common cancer in women and will be highly treatable because of screening. It implied that ca cervix should be diagnosed as early as possible. Therefore, the prediction model is needed to support the early detection and screening of ca cervix. This present study is aimed to make and evaluate predictive model of ca cervix by machine learning approach using open imbalance data. The machine learning approach was implemented by four steps, consists of data pre-processing, feature selection, predictive model development, and model performance evaluation. We used features in Waikato Environment for Knowledge Analysis (WEKA) version 3.8.4 for all of those steps, including imputation, interquartile range, Synthetic Minority Oversampling Technique (SMOTE), resample, and experimenter. The result of this study showed that all of predictive models (RandomForest, Bagging, LogitBoost, ClassificationViaRegression (CVR), and RandomCommittee) were acceptable for ca cervix prediction. The highest accuracy of all predictive model was 99.34%. Our proposed predictive model with highest performance is LogitBoost. The performance of ca cervix prediction using imbalance data will improve if we can handle the missing value, outlier, and data imbalance. We concluded that our proposed predictive model showed the good performance and may be visible for noninvasive ca cervix screening.

Cervical Cancer Screening using Machine Learning Approach

Sari Rahmawati Kusuma Dewi[a], Jakir Hossain Bhuiyan Masud[a], Emily Chia-Yu Su[a]
and **Ming-Chin Lin[a, b]**

[a] College of Medical Science and Technology, Taipei Medical University, Taiwan
[b] Taipei Medical University-Shuang Ho Hospital, Taiwan

Abstract

Cervical cancer (ca cervix) is one of most common cancer in women and will be highly treatable because of screening. It implied that ca cervix should be diagnosed as early as possible. Therefore, the prediction model is needed to support the early detection and screening of ca cervix. This present study is aimed to make and evaluate predictive model of ca cervix by machine learning approach using open imbalance data. The machine learning approach was implemented by four steps, consists of data pre-processing, feature selection, predictive model development, and model performance evaluation. We used features in Waikato Environment for Knowledge Analysis (WEKA) version 3.8.4 for all of those steps, including imputation, interquartile range, Synthetic Minority Over-sampling Technique (SMOTE), resample, and experimenter. The result of this study showed that all of predictive models (Random-Forest, Bagging, LogitBoost, ClassificationViaRegression (CVR), and RandomCommittee) were acceptable for ca cervix prediction. The highest accuracy of all predictive model was 99.34%. Our proposed predictive model with highest performance is LogitBoost. The performance of ca cervix prediction using imbalance data will improve if we can handle the missing value, outlier, and data imbalance. We concluded that our proposed predictive model showed the good performance and may be visible for noninvasive ca cervix screening.

Keywords:

Cervical Cancer, Machine Learning, Resample, Imbalance Data, Artificial Intelligence

Address for correspondence

Sari Rahmawati Kusuma Dewi
Graduate Institute of Biomedical Informatics, College of Medical Science and Technology,
Taipei Medical University
Taiwan
Email: kusuma.rahmawati@gmail.com

Note: This paper has been reviewed and recommended by AMAPI2020 Scientific Program Committee to be submitted to the International Journal of Computer Methods and Programs in Biomedicine (CMPB) in its "APAMI2020 Special Edition" as a full-paper scheduled to be published in 2021. Therefore, APAMI2020 Conference Proceeding has published only the abstract of this papers here to avoid any issue of double publication.