

A Contemporary Machine Learning Method for Accurate Prediction of Cervical Cancer

Jesse Jeremiah Tanimu^{1,*}, Mohamed Hamada^{2,**}, Mohammed Hassan^{3,***}, and Saratu Yusuf Ilu^{3,****}

¹Department of Computer Science, Bayero University, Kano. Nigeria

²University of Aizu, Japan

³Department of Software Engineering, Bayero University, Kano. Nigeria

Abstract. With the advent of new technologies in the medical field, huge amounts of cancerous data have been collected and are readily accessible to the medical research community. Over the years, researchers have employed advanced data mining and machine learning techniques to develop better models that can analyze datasets to extract the conceived patterns, ideas, and hidden knowledge. The mined information can be used as a support in decision making for diagnostic processes. These techniques, while being able to predict future outcomes of certain diseases effectively, can discover and identify patterns and relationships between them from complex datasets. In this research, a predictive model for predicting the outcome of patients' cervical cancer results has been developed, given risk patterns from individual medical records and preliminary screening tests. This work presents a Decision tree (DT) classification algorithm and shows the advantage of feature selection approaches in the prediction of cervical cancer using recursive feature elimination technique for dimensionality reduction for improving the accuracy, sensitivity, and specificity of the model. The dataset employed here suffers from missing values and is highly imbalanced. Therefore, a combination of under and oversampling techniques called SMOTETomek was employed. A comparative analysis of the proposed model has been performed to show the effectiveness of feature selection and class imbalance based on the classifier's accuracy, sensitivity, and specificity. The DT with the selected features and SMOTETomek has better results with an accuracy of 98%, sensitivity of 100%, and specificity of 97%. Decision Tree classifier is shown to have excellent performance in handling classification assignment when the features are reduced, and the problem of imbalance class is addressed.

1 Introduction

We are awash with data. The amount of data in the world, in our daily lives, seems to continue increasing exponentially [1]. Technology now allows us to capture and store large amounts of data. Finding patterns, conceived ideas, trends, anomalies in these datasets, and summarizing them with simple quantitative models is one of the major challenges of the information age [2, 3].

Data mining is simply referred to as extracting or "mining" knowledge from these large amounts of data. It can also be seen as a "process of identifying and understanding conceived patterns and ideas in the data". Data mining is now among the trending area of research in the medical domain [4, 5]. The knowledge extracted from the mining can be used to support the treatment of fatal diseases. However, the accurate prediction of a disease is one of the most interesting and challenging tasks for physicians. As a result, machine learning techniques have become popular tools for medical researchers. These techniques, while being able to predict future outcomes

of a certain disease effectively, can discover and identify patterns and relationships between them from complex datasets [6, 7]. Predictive models developed using machine learning techniques can also be useful in speeding up the process of detecting and diagnosing numerous diseases such as Prostate cancer, Diabetes, cervical cancer, schistosomiasis, etc.[8, 9].

Cancer is the second leading cause of death globally; it is responsible for an estimated 9.6 million deaths in 2018. Globally, about 1 in 6 deaths is due to cancer (International Agency for Research on Cancer (IARC) Release, 2018) [10]. According to a guidance note by the World Health Organization (WHO), that women cancers, including breast, cervical and ovarian cancer, are the leading causes of premature mortality among women worldwide [11]. As per the statistics issued by WHO, every year more than 270,000 women die from cervical cancer, more than 85% of these deaths are in developing countries with estimated annual new cases of 444,500 annually(IARC release, 2018).

Developed countries like the US and England are also facing a significant increase in patients with cervical malignancy. About 13,420 new cases of cervical cancers are expected to be diagnosed per annum, with estimated

*e-mail: tanimujessej@gmail.com

**e-mail: hamada@u-aizu.ac.jp

***e-mail: mhassan.se@buk.edu.ng

****e-mail: syilu.cs@buk.edu.ng

death cases of 4,170 in the US (American Cancer Society (ACS), 2019) [12]. In England, the rate of cervical cancer in women has increased exponentially from 2.7% to 4.6% [13]. Emerging a developing country like Nigeria has an estimated population of 50.33 million women ages 15 years and above who are at risk of developing cervical cancer. The current estimate indicates that every year 14,943 women are diagnosed with this ailment (IARC Release, 2018). Cervical cancer is ranked as the second most frequent malignant tumor among women in Nigeria and the leading gynecological malignancy with high mortality among the afflicted [14].

With the advent of new technologies in the medical field, huge amounts of cancer data have been collected and are readily accessible to the medical research community [6]. Researchers are constantly aiming to develop better models that can analyze these data to extract hidden knowledge, so that mined information can be used as a support in decision making for diagnostic processes. There have been many attempts to use machine learning in tasks such as prediction to help in the early detection of cervical cancer. New technological developments combined with machine learning tools offer the potential to tackle cervical cancer in a more cyclopedic approach and build a healthier future for girls and women.

Machine learning researchers are constantly striving to develop better predictive models that can analyze huge amounts of data available in the cervical cancer domain to identify and extract hidden information and used it as a support in decision making for diagnostic processes. However, these models while being able to predict outcomes of cervical cancer still suffer from one or more of the following limitations: High computation costs [15], No feature selection technique for dimensionality reduction [9], [16], lack of technique that will resolve the problem of imbalanced data that may lead to model skewed to one side [17], poor model accuracy, sensitivity, and specificity [18], etc. Hence the need to further improve the existing models.

The aim of this research is to develop a model for predicting the outcomes of the patient's cervical cancer test, given their risk factors and preliminary screening results from individual medical records. The major significance of the study to the medical field is that the proposed predictive model will aid in the diagnostic process of cervical cancer using machine learning techniques to predict the results of patients with cervical cancer with minimal or no error. This research focuses mainly on predicting the biopsy results of patients with cervical cancer. It concentrates majorly on having the correct results of patients suffering from the ailment, that is, have a better positive value of the result of cervical cancer test; therefore it focuses more on the sensitivity rather than determining other metrics like accuracy for evaluating the model performance.

This paper is composed of five sections including this introductory section. Section 2 contains a reference to the literature to inform what is already known, the work that has been done, and some of the recent models that have been developed to predict cervical cancers. Section 3 gives an account of how the research has been carried out, when section 4 presents and describes the research findings sys-

tematically. Finally, section 5 summarizes and gives the final comments on the research.

2 Related Work

Much of the current literature on cervical cancer pays particular attention to developing a machine learning model that can accurately predict whether a patient has cervical cancer or not. Fatlawi [9] in his research applied an enhanced decision tree classifier (DT cost-sensitive) for the classification of cervical cancer patients based on the dataset on the UCI repository. In his research, the enhanced DT performed better than the typical DT with a true positive (TP) rate (0.429) comparing with (0.160) for typical decision tree in a binary classification task. In his work, there is no feature selection technique applied to select the best subset of features and to achieve dimensionality reduction. Data samples are the basic components when applying ML techniques. Each sample is described with several features and each feature consists of different types of values [6, 19]. ML techniques work effectively when the dimensionality is minimized [20]. Furthermore, reducing dimensionality can eliminate irrelevant features, reduce noise, and, due to fewer features, robust learning models can be developed. Kotsiantis [21] stated that "given an input set of features, dimensionality can be reduced by selecting, hopefully, the best subset of features of the input feature set."

Wu and Zhou [15] in their research, cervical cancer risk factors were analyzed and three SVM-based approaches including SVM, SVM-RFE, and SVM-PCA are applied to the classification of cervical cancer dataset from the UCI repository. Sensitivity, Specificity, and Accuracy were taken into consideration during the classification process. The authors concluded that SVM works well; however, SVM-RFE and SVM-PCA do better when the numbers of features are lower. SVM as a classifier has been used in cancer classification. Although they are extremely powerful classifiers, they do have some limitations as follows: Finding the best model requires testing of various combinations of kernels and model parameters, it can be slow to train, particularly if the input dataset has a large number of features or examples, and their inner workings can be hard to understand because the underlying models are based on complex mathematical systems and the results are difficult to interpret [22].

In another related research by Al-Wesabi, Choudhury, and Won [23] on the classification of Cervical Cancer Dataset. The research investigates different machine learning classifiers such as Gaussian Naive Bayes (GNB), Decision Tree (DT), Logistic Regression (LR), k-nearest neighbors (KNN), and Support Vector Machines (SVM). The study also applied SMOTETomek (combine method), under-sampling, and over-sampling method to address the problem of imbalanced data. The further investigation involves wrapper methods such as Sequential Feature Selector, both forward and backward version. The result of their work shows that DT and KNN classifiers are the most advantageous in handling classification assignments with excellent performance. Overall, it is noticeable that

DT is the recommended classifier model for the UCI cervical cancer data in their work. The great limitation of their work is model overfitting due to the technique they used in splitting the datasets with limited observations of just 858 using the holdout method to divide the data into training and testing. Hybrid ML algorithms such as SVM Linear, Random Forest, and GBM(Gradient Boosting) with the SMOTE (Synthetic Minority Oversampling Technique) were applied on the same dataset by [18]. The research applied Genetic Algorithm (GA) for feature selection and Bayesian optimization for hyperparameter tuning. Comparative analysis of all the models in the research was done on the basis of sensitivity and specificity, where, GBM has delivered more promising results with a sensitivity of 0.778 (77.8%) followed by SVM Linear with a sensitivity of 0.5558 (55.58%), and Random Forest with sensitivity = 0.44 (44.4%), thus, their result can be improved to achieve excellent model performance.

3 Experimental Methodology

The section contains a description of the methods employed in this research. It also described all the experimental methods used in this work. It is divided into four subsections.

3.1 Data Acquisition

The dataset used to conduct this research is collected from the Hospital 'Universitario de Caracas' in Caracas, Venezuela. It is published on the UCI (University of California, Irvine) machine learning repository.

This is a public dataset that contains data of 858 patients and 32 features as well as four test results (classes). This research focuses on studying the Biopsy target as it is recommended by the literature review.

3.1.1 Data Description

The dataset is collected from the UCI ML repository and the file contains a list of 32 risk factors and test results (Hinselmann, Schiller, Cytology, and Biopsy) for cervical cancer. It was originally collected at 'Hospital Universitario de Caracas' in Caracas, Venezuela. The features cover demographic data, habits, and medical records of 858 patients. Unfortunately, the dataset contains missing values since some of the patients decided not to answer some of the questions due to privacy concerns. The biopsy has been the goal standard for diagnosing patients with cervical cancer while the other 3 is considered preliminary tests. Our target in this research is to predict the biopsy accurately. Table. 1 illustrates all the features of the dataset and their corresponding data types respectively. The table contains a complete description of the attributes of our dataset and their corresponding data types. It shows all the attributes of the dataset used to conduct this research with their data types as collected from the UCI ML repository [24].

The attributes consist of just 2 kinds of data types, that is, an attribute must either be of an integer or a Boolean

Table 1. Features and data types of dataset of the research

Attribute	Type
Age	Integer
Number of sexual partners	Integer
First sexual intercourse (age)	Integer
Smokes	Boolean
Smokes (years)	Boolean
Smokes (packs/year)	Boolean
Hormonal Contraceptives	Boolean
Hormonal Contraceptives (years)	Integer
IUD	Boolean
IUD (years)	Integer
Schiller	Boolean
STDs	Boolean
STDs: HIV	Boolean
STDs (number)	Integer
SGTDs: Hepatitis B	Boolean
STDs: condylomatosis	Boolean
STDs: HPV	Boolean
STDs: cervical condylomatosis	Boolean
STDs: Number of diagnosis	Integer
STDs: vulvo-perineal condylomatosis	Boolean
STDs: Time since last diagnosis	Integer
STDs: syphilis	Boolean
Dx: Cancer	Boolean
STDs: pelvic inflammatory disease	Boolean
Dx: CIN	Boolean
STDs: genital herpes	Boolean
Dx: HPV	Boolean
STDs: molluscum contagiosum	Boolean
Dx	Boolean
STDs: AIDS	Boolean
Hinselmann	Boolean
Cytology	Boolean
Biopsy(Target)	Boolean

type. The integer data type basically represents whole numbers that have no fractional parts while the Boolean data type is a data type that has one of two possible values (usually denoted by Yes or No). A typical example of an integer data type in Table. 1 is the patient's age (which must be a whole number) while that of a Boolean data type is Smoke (which denotes whether a patient smokes or not).

3.2 Data Pre-processing

The preprocessing experiments involve data cleansing, missing value treatment using standard measurements like the mean for integer values and mode for Boolean values. As a result, 13% of total questions were missed. There are two features with 92% of missing values which are STDs: Time since first diagnosis and STDs: Time since the last diagnosis, so they have been dropped. Observations with up to 20 or more missing values were also omitted.

3.3 Modeling

The machine learning algorithm that is used for training this dataset is Decision Tree (DT) with the Recursive Feature Elimination (RFE) method for feature selection and SMOTomek for balancing the data. Figure. 1 shows the basic architectural framework of the proposed model. De-

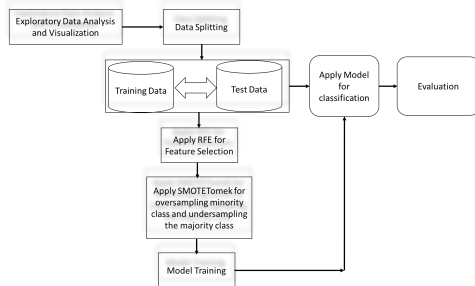


Figure 1. Architecture of the Proposed Model

cision trees (DTs) are intuitive models that make their decisions based on a branching sequence of the Boolean test, that is, the question asked about feature values. A decision tree can be described as a series of yes/no questions asked about our data leading to predicted class. DTs follow a tree-structured classification scheme where the nodes represent the input variables, and the leaves correspond to decision outcomes [25]

3.4 Validation and Execution (testing)

The model will be executed, and outcomes are validated for model Accuracy, Sensitivity, and Specificity based on K-fold cross-validation. To evaluate the performance of the developed machine learning model, the evaluation metrics used are sensitivity, specificity, and accuracy.

4 Result and Discussion

While training the decision tree classifier, 10-fold repeated cross-validation was applied with 10 repeats because observations are limited and may lead to overfitting, which is a major drawback of the decision tree. This will also allow all the data to participate in the training process.

4.1 Basic Notations

Some notations are used to present the results of our findings. These notations are defined as follows:

- **Accuracy:** It is the number of correct predictions by the model out of the total number of observations as shown in equation (1).

$$Accuracy = TP + \frac{TN}{TP + TN + FP + FN} \quad (1)$$

- **Sensitivity:** Is the ability of the classifier to correctly identify the positive class (cancerous). It is defined mathematically in equation (2).

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

- **Specificity:** Is the percentage of times the classifier predicted a Negative class out of all the times the class was is Negative (non-cancerous). Equation (3) gives the definition of specificity

$$Specificity = \frac{TN}{FP + TN} \quad (3)$$

4.2 The basic classification

Table 2. Results of classifier I without applying techniques for feature selection and class balancing

Metrics	Classifier I Performance
Accuracy	0.96
Sensitivity	0.86
Specificity	0.97

The first experiment was mainly to examine the impact of not applying feature selection techniques in predicting cancerous cases. This experiment was therefore conducted without extracting the most relevant features from the dataset. Equations (1), (2), and (3) give the mathematical representations of the metrics used for experimental evaluation of the model. Table. 2 provides the breakdown of the first experimental results based on the metrics used in evaluating the model. The resulting metrics were computed using the confusion matrix. The confusion matrix is a table representing the performance of a classifier to classify labels correctly.

Table. 2 above shows the result of using a decision tree classifier with no feature selection and data balancing techniques applied. It has a significant performance in terms of the overall accuracy with an accuracy of 96% and also performed very good in predicting patients with negative cervical cancer results with a specificity of 97%. But the result delivered for true positive cases does not seem promising. The single most striking observation to emerge from the results presented was the sensitivity value, which was found to be the least (86%). The sensitivity is the metric that can best identify positive cases. Therefore, there is a need to further improve the classifier by applying feature selection techniques for dimensionality reduction.

4.3 Improved classifier with feature selection technique

When we applied the selected features with the highest ranking in RFE, we got a different as shown in Table. 3. The selected features as shown in the table are the patient's age, number of sexual partners, first sexual intercourse, number of pregnancies, etc., and they are based on their significance in determining whether the patience is cancerous or not.

As seen in Table. 4, the first column shows the evaluation metrics. The model performance based on the metrics defined in the second column, and finally, the improvement from the last classifier is shown in column three. The classifier has shown some levels of improvement in terms

Table 3. Features with the top ranking in RFE

S/N	Attribute
1	Age
2	Number of sexual partners
3	First sexual intercourse (age)
4	Number of pregnancies
5	Smokes
6	Smokes (packs/year)
7	Hormonal Contraceptives
8	Hormonal Contraceptives (years)
9	IUD
10	IUD (years)
11	STDs
12	STDs: molluscum contagiosum
13	STDs: HIV
14	STDs: Hepatitis B
15	STDs: HPV
16	STDs: AIDS
17	Dx: Cancer
18	Dx: CIN
19	Dx: HPV
20	Dx
21	STDs (number)
22	STDs: Number of diagnosis

Table 4. Results of classifier II with the selected features

Metrics	Performance	Improvement (II - I)
Accuracy	0.98	0.98-0.96 = 0.02
Sensitivity	0.86	0.86-0.86 = 0.00
Specificity	0.97	0.97-0.97 = 0.00

of the overall accuracy with an increase of 2%. However, there was no improvement in terms of sensitivity was observed, which is our main target. Consequently, there is a further need to also improve the current classifier.

Interestingly, since the dataset has a problem of imbalanced data where the data collected are not equally represented, we applied a data balancing technique called the SMOTETOMEK. It has some advantages over other techniques, where it undersamples the majority classes and at the same oversamples the minority classes. Often real-world datasets are predominately composed of "normal" examples with only a small percentage of "abnormal" or "interesting" examples. This imbalance gives rise to the "class imbalance" problem [26, 27] which is the problem of learning a concept from the class that has a small number of observations. In the real world, numerous studies have shown that better prediction performance can be achieved by having balanced data; therefore, a number of well-known methods have been developed and used in machine learning to tackle this issue for improving the prediction models' performance [28, 29].

Table 5. Results of classifier III with the selected features

Metrics	Performance	Improvement (III - II)
Accuracy	0.98	0.98-0.98 = 0.00
Sensitivity	1.00	1.00-0.86 = 0.14
Specificity	0.97	0.97-0.97 = 0.00

4.4 Improved classifier with selected features and balanced classes

As shown in Table. 5, the decision tree classifier with the combination of selected features shown in Table. 3 has outperformed all other classifiers in terms of sensitivity. This is because the research work aims to develop a better predictive model that can identify majorly the positive cases. The result in the Table. 5 revealed that the classifier has improved significantly when the data balancing technique was applied. There was a significant sensitivity difference between the models (the difference of 14%). The difference between sensitivities is interesting because it means that the model can best identify positive cases compared to models presented in sections 4.2 and 4.3. The main aim of the research is achieved since all the positive cases were identified and correctly classified based on the

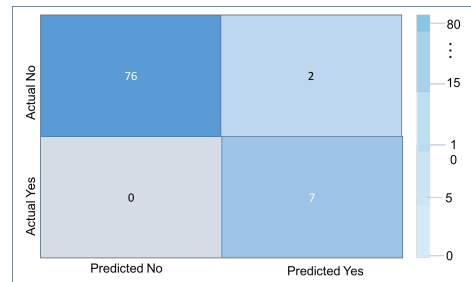


Figure 2. Confusion Matrix of our Classifier

confusion matrix shown in Figure. 2. From the data in the figure, it is apparent that no real or actual case was wrongly classified.

The study has shown the potential for DT learning for predicting cervical cancer. The DT is easy to understand and interpret since it can be visualized. The decision tree is able to handle both numerical and categorical data. This work was able to overcome the problem of overfitting (a problem where decision-tree learners create over-complex trees that do not generalize the data very well) and data imbalance. Mechanisms such as setting the minimum number of samples required at a leaf node and setting the maximum depth of the tree were applied to overcome overfitting. While we applied SMOTETomek, to address the data imbalance problem. The cost of developing a decision tree is inherently lesser when compare to the work of Wu & Zhou [15]. Finally, feature selection techniques for dimensionality reduction were applied in our work to fill the gap found in Fatlawi [9].

Altogether, our result shows significant improvement when compared with other approaches. We have described a new process of predicting patients with cervical cancer

by selecting the most relevant subset of features through which we achieve better results when compared to previous works. In recent works using the same dataset, the metrics used are Accuracy, Specificity, and sensitivity. Our model has better sensitivity (metric of interest) over the previously developed models.

5 Conclusion and Recommendation

The main goal of the current study was to propose a model for the prediction of cervical cancer named a novel Machine Learning model for classification of cervical cancer using DT with a Recursive Feature Elimination, which could be used to accurately predict patients with this fatal disease. A ML model was implemented to predict the biopsy results of cervical cancer. One of the more significant findings to emerge from this study is that the model using a DT, SMOTETomek with Recursive Feature Elimination technique has delivered a more reliable predictive model to classify cervical cancer by using patient's data with their preliminary screening test. The experimental results indicate significant improvement in the proposed model when compared with other recent models. The study has gone some way towards enhancing our understanding of applying SMOTETomek with the Recursive Feature Elimination technique. The issue of predicting cervical at an early stage is an intriguing one which could be usefully explored in further research. Further research regarding the role of other feature selection methods such as LASSO would be worthwhile. It is recommended that further research be undertaken with a very large volume of the dataset so that in-depth analysis and understanding can be performed and a better predictive model can be developed for the same problem.

References

- [1] R.A. Guzzo, A.A. Fink, E. King, S. Tonidandel, R.S. Landis, *Industrial and Organizational Psychology* **8**, 491 (2015)
- [2] T. Ngo, *ACM SIGSOFT Software Engineering Notes* **36**, 51 (2011)
- [3] M. Hassan, Ph.D. thesis, The University of Aizu (2018)
- [4] T. Anand, R. Pal, S.K. Dubey, *International Journal of Business Information Systems* **31**, 372 (2019)
- [5] M. Hassan, M. Hamada, *International Journal of Computational Intelligence Systems* **11**, 146 (2018)
- [6] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, *Computational and structural biotechnology journal* **13**, 8 (2015)
- [7] K. Haruna, M.A. Ismail, A. Qazi, H.A. Kakudi, M. Hassan, S.A. Muaz, H. Chiroma, *Scientometrics* **125**, 101 (2020)
- [8] L.H. Xiao, P.R. Chen, Z.P. Gou, Y.Z. Li, M. Li, L.C. Xiang, P. Feng, *Asian journal of andrology* **19**, 586 (2017)
- [9] H.K. Fatlawi, *International Journal of Computer Techniques* **4**, 115 (2017)
- [10] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D. Parkin, M. Piñeros, A. Znaor, F. Bray, *International journal of cancer* **144**, 1941 (2019)
- [11] W.H. Organization et al. (2013)
- [12] R.A. Smith, K.S. Andrews, D. Brooks, S.A. Fedewa, D. Manassaram-Baptiste, D. Saslow, R.C. Wender, *CA: a cancer journal for clinicians* **69**, 184 (2019)
- [13] A. Castanon, P. Sasieni, *Preventive medicine* **107**, 21 (2018)
- [14] E. Oluwole, A. Mohammed, M. Akinyinka, O. Salako, *Journal of Community Medicine and Primary Health Care* **29**, 81 (2017)
- [15] W. Wu, H. Zhou, *IEEE Access* **5**, 25189 (2017)
- [16] T.M. Alam, M.M.A. Khan, M.A. Iqbal, W. Abdul, M. Mushtaq, *IJACSA International Journal of Advanced Computer Science and Applications* **10** (2019)
- [17] D.N. Punjani, K. Atkotiya, *International Journal of Research and Analytical Reviews* (2019)
- [18] H.D. Singh, Ph.D. thesis, Dublin, National College of Ireland (2018)
- [19] A.H. Abba, M. Hassan, *Design and Implementation of a CSV Validation System*, in *Proceedings of the 3rd International Conference on Applications in Information Technology* (2018), pp. 111–116
- [20] C.C. Aggarwal, *Data classification*, in *Data Mining* (Springer, 2015), pp. 285–344
- [21] S. Kotsiantis, *Artificial Intelligence Review* **42**, 157 (2011)
- [22] S. Huang, N. Cai, P.P. Pacheco, S. Narrandes, Y. Wang, W. Xu, *Cancer Genomics-Proteomics* **15**, 41 (2018)
- [23] Y. Al-Wesabi, A. Choudhury, D. Won, *Classification of cervical cancer dataset*, in *Avishek Choudhury, Wesabi, Classification of Cervical Cancer Dataset, Proceedings of the 2018 IISE Annual Conference, Orlando* (2018), pp. 1456–1461
- [24] A. Asuncion, D. Newman, *Uci machine learning repository* (2007)
- [25] A. Navlani, *Data Camp* (2018)
- [26] R.C. Prati, G.E. Batista, M.C. Monard, *Data mining with imbalanced class distributions: concepts and methods.*, in *IICAI* (2009), pp. 359–376
- [27] M. Hassan, M. Hamada, *Smart media-based context-aware recommender systems for learning: A conceptual framework*, in *2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET)* (IEEE, 2017), pp. 1–4
- [28] M. Alghamdi, M. Al-Mallah, S. Keteyian, C. Brawner, J. Ehrman, S. Sakr, *PloS one* **12**, e0179805 (2017)
- [29] M. Hamada, N.B. Odu, M. Hassan, *A fuzzy-based approach for modelling preferences of users in multi-criteria recommender systems*, in *2018 IEEE 12th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MC-SoC)* (IEEE, 2018), pp. 87–94