© APR 2025 | IRE Journals | Volume 8 Issue 10 | ISSN: 2456-8880

Adoption of Artificial Intelligence and Machine Learning Algorithms on Assessment and Prevalence of Cervical Cancer Risk Factor amongst Women in Umuagwo Ohaji/Egbema LGA Imo State, Nigeria

OGUOMA IKECHUKWU STANLEY[1], EBELE PRECIOUS OKEMBA[2], SHANICE KRISTY ARCHIBALD[3], DAVID OSONDU AKUCHIE[4], AJERO CHUKWUKA EVANS[5]

[1]*Department of Computer Science, University of Agriculture and Environmental Science Umuagwo, Imo State, Nigeria*

[2]*University of Kent United Kingdom*

[3]*University of the West Indies-St Augustine Campus*

[4]*University of Kairouan, Tunisia*

[5]*Department of Computer Science Imo State University of Owerri, Imo State*

Abstract- This study was able to ascertain clearly how cervical cancer could be managed in Umuagwo, Ohaji/Egbema L.G.A Imo State using a hybrid predictive model with outstanding focus on three different variables for each algorithm applied on the dataset which are (DT= motivation_willingness, Perception severity, empowerment knowledge and RF= empowerment ability, perception severity, motivation strength). The major objectives of the research includes to analyze the dataset using two major classification algorithms tool to predict or identify the risk factor of cervical cancer amongst women of Umuagwo Ohaji/Egbema LGA and to build and train a hybrid model that can analyze and predict the major cause of cervical cancer amongst women of childbearing age in Umuagwo Ohaji/Egbema LGA. The research was motivated in other to identify the strong believe the people of Umuagwo in Ohaji/Egbema has on their deity/juju as being responsible for the frequent death of their women. The study employed two machine learning classification algorithm methods which include: Decision Tree (DT) and Random Forest (RF) algorithms. The data was analyzed with R and JASP platform while the experiments are done on the dataset containing TRAIN, VALIDATION, TEST) for Random Forest (RF) with a percentage rate of 46%, 12% and 14% respectively and for Decision tree (DT) with percentage rate of (TRAIN 58% and TEST 14%) making up to the total of 72% for both algorithms. The results prediction accuracy was concluded by comparing the two developed models involving their different F1 scores, confusion matrix, Evaluation Metrics, Roc Curves, Precision (positive predictive value)/support and the RF out-of-bag result shown in Table 9 of this paper. From the predicted result shown by the hybrid model, it was observed that lack of willingness or motivation, empowerment ability and knowledge was the major risk factor causing the increase of cervical cancer amongst women of Umuagwo Ohaji/Egbema whereby leading them to frequent death and not their deity/juju as perceived by the villagers with percentage accuracy of DT= (0(NO) = 0.824 : 82% Accuracy and 1(YES) = 0.727 :73%) while RF= (0(NO) = 0.889: 89% Accuracy and 1(YES) = 0.800: 80% Accuracy) respectively.

Indexed Terms- Artificial Intelligence, Machine Learning, Decision Tree and Random Forest Classification Models, Cervical Cancer Risk Factor amongst Women, health diagnosis and treatment and cervical cancer Prevalence.

I. INTRODUCTION

Cancer of the cervix uteri is the fourth most common cancer among women worldwide and the leading cause of gynaecologic cancer death in the less developed regions (Binka *et al.*, 2019). Cervical cancer is cancer that starts in the cells of the cervix. The cervix is the lower, narrow end of the uterus (womb). The cervix connects the uterus to the vagina (birth canal). Cervical cancer usually develops slowly over time. Before cancer appears in the cervix, the cells of the cervix go through changes known as dysplasia, in which abnormal cells begin to appear in the cervical tissue. Over time, if not destroyed or removed, the abnormal cells may become cancer cells and start to grow and spread more deeply into the cervix and to surrounding areas (Bray, et al, 2018).

Globally, cervical cancer is the fourth most common cancer in women, with 604 000 new cases in 2020 (Binka, et al, 2019). About 90% of the 342 000 deaths caused by cervical cancer occurred in low- and middle-income countries (Binka, et al, 2019). The highest rates of cervical cancer incidence and mortality are in sub-Saharan Africa (SSA), Central America and South-East Asia (Basu, et al, 2018). Regional differences in the cervical cancer burden are related to inequalities in access to vaccination, screening and treatment services, risk factors including HIV prevalence, and social and economic determinants such as sex, gender biases and poverty (Basu, et al, 2018). Women living with HIV are 6 times more likely to develop cervical cancer compared to the general population, and an estimated 5% of all cervical cancer cases are attributable to HIV. The contribution of HIV to cervical cancer disproportionately affects younger women, and as a result, 20% of children who lose their mother to cancer do so due to cervical cancer (World Health Organization, 2023).

Highest incidence of cervical cancer related death occurs among middle age women about 30–40 years. Of the 273,505 deaths recorded, 80% occurred in low and middle income countries. In Sub-Saharan Africa (SSA), of the 78,897 women diagnosed with cervical cancer annually, 61,671 deaths were recorded which makes the disease one of the most prevailing cancers. In Uganda, Mali, Nigeria, and Zimbabwe, cervical cancer is the second most prevailing cancer among women aged 15–44 years. A major misconception lies in the treatment of cervical cancer which is viewed as the removal and reinsertion of the womb and believed to cause unavoidable death. In SSA, cervical cancer is yet to be acknowledged as an important public health problem. The low awareness of the disease in Africa which cuts across different literacy levels have been reported

The Ministry of Health and Child Welfare, Zimbabwe introduced the VIAC screening in 2011. In order to make cervical cancer screening affordable to the majority of women in the country, VIAC is offered for free in public hospitals (Kuguyo et al., 2017). Yet, despite these efforts, the uptake of cervical cancer screening in Zimbabwe is still quite low. The uptake of cervical cancer screening stands at 9.4% (all women aged 25–64 years) in the country (Bray et al., 2018). Of the 47,916 women aged 15–47 years in Chegutu Rural District, only 2.1% have been screened for cervical cancer through VIAC since 2014. There is low uptake of cervical cancer screening in Zimbabwe and other Southern African countries (e.g., Swaziland and Malawi), despite the fact that these countries have the highest age-standardized incidence rates globally of 62.3, 75.3 and 72.9 per 100,000, respectively (Bray et al., 2018; Msyamboza et al., 2016; Ngwenya & Huang, 2018).

The Papanicolaou (Pap) smear is one of the most essential screening tools for the early diagnosis of cervical cancer and has been found to be the most effective preventive measure (WHO, 2023). The value of cervical cancer screening in reducing the risk of cervical cancer and mortality has been established, and the risk of developing cervical cancer can be reduced by 80% through regular screening (Özgül, 2017). The benefits of Pap smear's wide availability and usage have been documented, resulting in lowering of mortality rates by up to 60 to 90% in some developed countries (Wong, 2019).

In Nigeria, only 22 (4.3%) among 500 attendees of a maternal and child health clinic in Lagos were found to be aware of cervical cancer disease. Over 80% of one hundred and thirty-nine patients with advanced cervical cancer said they have never heard of the disease whereas 10%–30% assumed that the symptoms they presented were related to lower genital infection, menstrual cycle, and irregular menses. Only 9% of the patients knew the disease was cancer related, a condition needing urgent medical attention. Poor awareness of the disease and attribution of symptoms to minor health condition led 98% of patients with advanced invasive cervical cancer to believe that their health problem was incurable.

Being aware, knowledgeable or educated on an issue or situation is very essential in the life of everyone since it enables individuals to take or make informed decisions on the issue at stake and follow it up with appropriate actions. This is in tandem with the popular adage which says that, “Knowledge is power and ignorance a deadly disease”. With knowledge on various issues readily available to us we can conquer the world (Rositch, et al. 2016).

Health they say is wealth, which implies that if one is healthy one can achieve anything in life; but where one is unhealthy he/she may spend all his/her resources on achieving good health and may never be wealthy (Ansink, 2018). Therefore, knowledge on health issues (such as cervical cancer and pap smear screening) enables us to be aware of the said health issues, their causes, risk factors, implications and the various alternatives or pro-health behaviors we can take in order to remain healthy and avert suffering, morbidity and mortality. Therefore the interest to assess cervical cancer incidence, Knowledge, attitudes and barriers to its screening in Umuagwo Ohaji/Egbema L.G.A. Imo State.

Furthermore, It is unarguable that the greater number of Umuagwo women in Ohaji/Egbema Local Government Area have little or no knowledge about cervical cancer, let alone how and what to do to detect it early (Udigwe, 2018). This is true, that these women are also more preoccupied with their peasant farming and their small scale business engagements that they use to earn a living. Many women have died from cervical cancer and they believed that their deity /juju were responsible for such death. Also, over adherence to belief, either Christianity or their worships, make it difficult to make this topic an issue for discussion, the gender inequality in our society further darkens the hope of our women. It may be so difficult for any man in the village, to release his wife for such knowledge exposure. Else, it will be termed a different negative thing, that may mare the women’s marital joy. This paper tends to discover incidence of cervical cancer and related risk factors among the women of Umuagwo, assess the knowledge of cervical cancer among women, identify the attitudes toward cervical cancer and its screening among women and identify the barriers to uptake of cervical cancer screening among women of Umuagwo in Ohaji/Egbema L.G.A Imo State.

II. LITERATURE REVIEW

According to Binka et al., (2019) on the topic Barriers to the Uptake of Cervical Cancer Screening and Treatment among Rural Women in Ghana Twenty-five in-depth interviews were conducted, while three focus group discussions were held among respondents. The data were analyzed with the R package for qualitative data analysis using a thematic analytical approach. Results shows Low level of knowledge about the disease and screening services, personal or psychological convictions, and cost of screening and treatment coupled with a low level of income were the barriers at the individual level. Perceived health personnel attitude, perceived lack of privacy, and misdiagnosis were the barriers at the institutional level while the sociocultural belief system of the communities about the etiology of the disease was the barrier at the community level. Inadequate education about the disease, lack of funding and access to screening facilities also constrained screening and treatment at the policy level. Conclusions. Cervical cancer screening and treatment are constrained at multiple levels in rural Ghana. This study underscores the need to address the low uptake of cervical cancer screening and treatment at the individual, community, institutional, and policy levels simultaneously.

Peterson et al, (2022) carried out a study on knowledge, altitudes and practice regarding cervical cancer screening among villages health volunteers. In Nakornnayok which have four sub districts, namely KlonaYai, Chompho, Buangsan and Suksara, Thailand. The respondent comprises of 128 village health volunteers from Nakornnayok. The questionnaire was designed to assess the knowledge and attitude of cervical cancer screen provided by the VHHS. In addition, cervical cancer screening coverage rates of each area were collected. The demographic data, scores of knowledge, attitude, practices and cervical screening coverage rates was analysed by one way ANOVA. The questionnaire reliability was assessed as 0.81, the total knowledge, attitude score were 10 and 15 points. The mean knowledge score of KlongYai, Chomphol and 9.0 points, respectively. The VHVs had a high level of overall knowledge about cervical cancer screening. The mean attitude score were 12.4, 13.2, 13.4 and 13.1 points. After analyzing it was found out that VHV have a positive attitude to the promotion of cervical cancer at the overall to level. The percentage of VHV promoting cervical cancer information in 50.0. However, the cervical screening coverage rate was 62.4%, 34.7%, 80.3% and 47.3% respectively. In conclusion, the knowledge, attitude and percentages of promoting information of cervical cancer screening among VHVs in the four sub-districts were high but did not correlate with the cervical screening coverage rates for each area. VHCs were recommended to understand the socio-cultural beliefs of the women in the target population and design suitable strategies to encourage high cervical screening coverage.

According to Peterson et al, (2022) on the topic Barriers to uptake of cervical cancer screening services in low-and-middle-income countries: a systematic review. This was a systematic review using Medical Subject Headings (MeSH) terms in Google Scholar, PubMed, Scopus, and Web of Science databases. We also contacted medical associations and universities for grey literature and checked reference lists of eligible articles for relevant literature published in English between 2010 and 2020. We summarized the findings using a descriptive narrative based on themes identified as levels of the social ecological model. Seventy-nine articles met the inclusion criteria. We identified individual, cultural/traditional and religious, societal, health system, and structural barriers to screening. Lack of knowledge and awareness of cervical cancer in general and of screening were the most frequent individual level barriers. Cultural/traditional and religious barriers included prohibition of screening and unsupportive partners and families, while social barriers were largely driven by community misconceptions. Health system barriers included policy and programmatic factors, and structural barriers were related to geography, education and cost. Underlying reasons for these barriers included limited information about cervical cancer and screening as a preventive strategy, poorly resourced health systems that lacked policies or implemented them poorly, generalized limited access to health services, and gender norms that deprioritize the health needs of women. In conclusion, A wide range of barriers to screening were identified across most LMICs. Urgent implementation of clear policies supported by health system capacity for implementation, community wide advocacy and information dissemination, strengthening of policies that support women's health and gender equality, and targeted further research are needed to effectively address the inequitable burden of cervical cancer in LMICs.

Ngwenya P. and Huang, O. (2018) in their study carried out on knowledge, preventive measure and associated factors of female nurses towards cervical cancer in the selected government hospitals in Addis Ababa, Ethiopia. 275 nurses participated in the cross-sectional descriptive study by responding to a structured questionnaire about knowledge and preventive measures. Statistical analysis included both bivariate and regression analysis, while controlling for possible confounders. A result gotten was that a little over half (60.8%) of nurses had knowledge of cervical cancer but only 21.9% reported practicing prevention of cervical cancer. Marital status and training about cervical cancer screening had a strong and positive association on knowledge education, family history, unit of work education, family history, unit of work and ever cared patient with cervical cancer were also significantly associated with younger age, work experience, been diagnosed with cervical cancer and ever cared patient with cervical cancer and ever visited a health institution. In conclusion at least 60% of the respondent were knowledgeable but preventive measures among nurses were low. Consistent training is required on knowledge and preventive measures of cervical cancer to combat it's high morbidity and mortality in Ethiopia.

Msyamboza, Franco and Harper (2016) carried out a study on community based, qualitative study on women living in Kaduna state, Nigeria. It assessed women's knowledge about cervical cancer screening and identified barriers affecting screening for cervical cancer among women in Kaduna State. The population of the study includes women of reductive age group between 15-45 years who were purposively selected. Qualitative technique of data collection using focus group discussions were held for the women. The study took place within the three senatorial zones of the state. The results was transcribed themes and patterns emerged, systematically and critically analyzed and results presented verbatim in quotations. The findings shows that various factors influencing utilization of cervical cancer screening among a woman which was the fact that, majority of the respondents were not aware of the screening service. Among those who were aware of the screening, male health personnel screening the females was the major barrier. Major factors identified that can improve screening for the women and female health personnel's should be trained on cheap screening methods such as visual inspection using acetic acid which is a screening method suitable for low resource settings.

A study conducted by Rositch, Gatuguta and Choi (2016). carried out a study on knowledge, a hospital based cross-sectional study was conducted with a sample population of women aged 20 or more were interviewed using a structured questionnaire to evaluate the socio-demographic information on knowledge, attitude, practice and barriers to cervical cancers screening. Total of 360 participants were recruited, mean age was 30.13±10.4 years. More than 87% of participants had inadequately knowledge but around 72% had a favorable attitude towards cervical cancer screening. There was a significant portion of women (86.4%) who had never done any cervical cancer screening test. Despite being higher literature rate Brahmin and Chhetri ethnic group, they were less likely attend the cervical cancer screening than Dalit and Janajati and those who had a positive family history of cancer were more likely to attend the cervical cancer screening. Similarly, married women, who had adequate knowledge and or favourable attitude were, more likely to practice cervical cancer screening, though statistically not significant. In conclusion of the study adequate knowledge of preventive measures and practice of cervical cancer screening were meager among rural Nepalese women, but most of them had a favourable attitude. In recommendation there is an imperative need for related awareness programs to promote the uptake of cervical cancer screening tests, to assess knowledge of preventive measures of cervical cancer screening and to identify the actual barriers to cervical cancer screening.

III. METHODOLOGY

Two different classification algorithms was applied in this paper, Decision tree and random forest algorithms because of their ability to build a model with more accuracy in data prediction and for effective decision making in real life. With the application of the two algorithms, it means that the actual methodology used on this research was a hybrid-approach.

IV. DECISION TREE ALGORITHMS

Decision tree (Han and Kamber; 2001) could be seen as a type of tree structure typically in a form of flowchart design. These tree structures are used to carry out classification and prediction modeling of objects in a class in a form of nodes and internodes. Both root and the internal nodes are taken as the test cases in the modeling process which in terms used as a separator with different features(Han and Kamber; 2001). According to (Yan *et al*; 2020), these decision trees uses a classification or regression models to form a tree structure. The structure breaks down a particular data set into various smaller and smaller subsets as the associated decision tree development id in progress. The researcher further noted that the decision tree is build up with the nodes and leaf nodes, where the decision nodes has two or more different branches while leaf nodes shows the classification or decision results(Yan *et al;* 2020) stated. Figure 1: Illustrate the structure of a decision tree

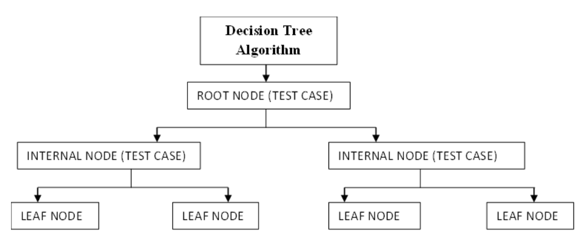


Figure 1: Illustrate the structure of a decision tree

Adoption of decision tree for this study did not just come but it was adopted because of its powerful technique for classification and prediction ability on a particular data set.

V. RANDOM FORESTS ALGORITHM

According to (Ho and Tin Kam 1995) stated that Random forests or random decision forests is an ensemble learning method for classification, regression and other tasks that works by creating a multitude of decision trees during training. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the output is the average of the predictions of the trees.

VI. THE PROPOSED SYSTEM DIAGRAM

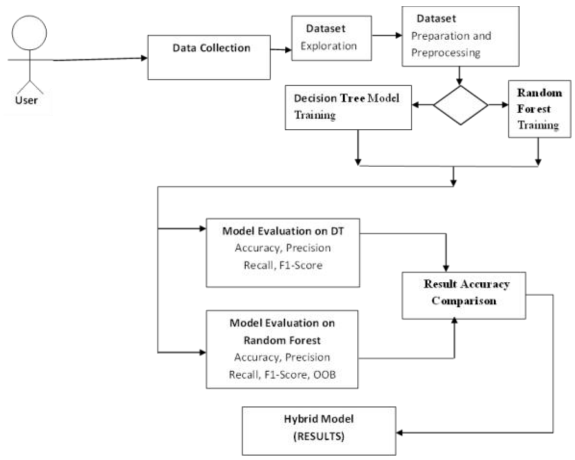


Figure 2: Diagram of the proposed Predictive model

The above diagram illustrates how the proposed system loads the dataset (cervical+cancer+behavior+risk.csv) dataset into the JASP machine learning platform. The dataset was first of all undergo data preprocessing/data cleaning stage, it is here that the researcher understood and discovered some missing values which was removed. Then extraction and feature process are performed. The dataset was now Split into three parts (TRAIN, VALIDATION, TEST) for Random Forest (RF) with a percentage rate of 46%, 12% and 14% respectively making up to the total of 72% while Decision Tree (DT) was split into two.

In summary, the researcher adopted the R and JASP analytical platform for the analysis while dataset was gotten from a survey conducted in Umuagwo Ohaji/Egbema LGA OF Imo State Nigeria.

Table 1: SYSTEM ALGORITHM

INPUT	cervical+cancer+behavior+risk.csv
OUTPUT	Artificial Intelligence Model to Predict Prevalence of Cervical Cancer Risk Factor amongst Women in Umuagwo Ohaji/Egbema LGA

VII. RESULTS EXPERIMENTS ON THE DATASET USING R

The first process was launching of the RStudio IDE after a successive launching, the following steps were done to design the model.

Step1: Loading packages to be used (that is libraries)
Step 2: Loading My Dataset to R Dataframe
Step 3: Exploring the data, at this stage, skimr::skim cervical cancer dataset (cervical+cancer+behavior+risk.csv)
Step 4: Converting outcome from numeric to factor and renaming them for easy understanding
Step 5: Changing some response by patients from alphabet to Numeric (YES = 1 and NO = 0)
Step 6: Plot Observations to view all the dataset
Step 7: Checking For Missing Values in Each Variable
Step 8: Replacing or removal of the missing values for each variable
Step 9: Normalizing the dataset

VIII. MODEL BUILDING (Random Forest)

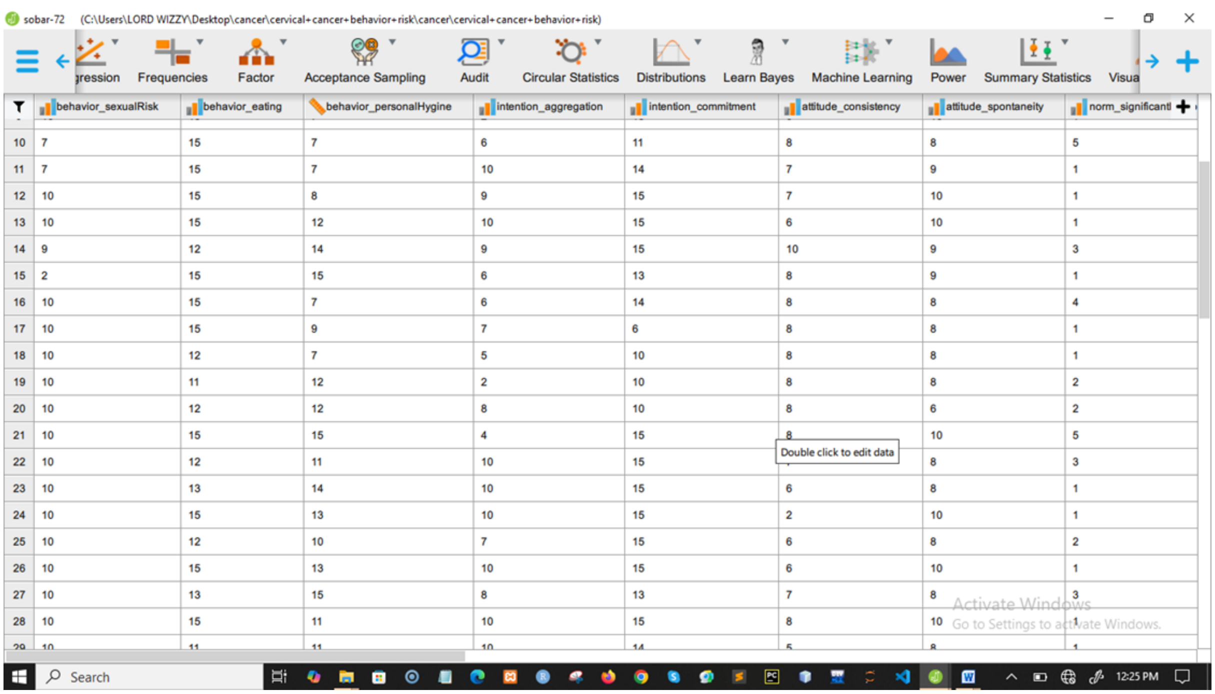
Step 10: The dataset was Split into two with the percentage of (46% = training, 12% = Validation and 14%=testing) respectively.
Step 11: applying Random Forest algorithm
Step 12: Using GINI MODEL to build the model.
Step 13: Analyzing the Prediction of the Model built With Gini Model
Step 14: Confusion Matrix
Hence the confusion matrix is used for more accuracy on the rate at which the model predicts user data.
Step 15: Validating the Model On the test Dataset
Step 11 was carried out again using Decision Tree algorithms on the same dataset in other to complete the hybrid use of the two algorithms and it is done following other steps below for a more accurate prediction of the model produced.

IX. EXPERIMENTS ON THE DATASET USING JASP PLATFORM

The first process was launching of the JASP PLATFORM after a successive launching, the following steps were done to build the model.

Step 1: Loading the dataset from the location cervical cancer dataset (cervical+cancer+behavior+risk.csv)
Step 2: Select machine learning packages
Step 3: Select first classification algorithm (Random Forest)
Step 4: Set the target and features (Class: 1 or 0)
Step 5: Click to start the analysis on the dataset
Step 6: Click on Data split
Step 7: Click Confusion Matrix
Step 8: Class Proportions
Step 9: Click Evaluation Metrics
Step 10: Out-of-Bag Accuracy Evaluation
Step 11: Click ROC Curve Plot
Step 12: click Andrews Plot
Step 13: click decision tree plot
Download visualizations results
Step 13: Start prediction accuracy by checking (F1 score, confusion matric, and Roc Curve and precision (positive predictive value) and OOB accuracy).
Step 3: was carried out again to use Decision Tree algorithms on the same dataset in other to complete the hybrid use of the two classification algorithms and it is done following other steps below for a more accurate prediction of the model produced which are predicted by looking at the output F1 score, ROC curve, confusion matric and decision tree models built.

X. EXPERIMENT OUTPUT



	behavior_sexualRisk	behavior_eating	behavior_personalHygine	intention_aggregation	intention_commitment	attitude_consistency	attitude_spontaneity	norm_significant
10	7	15	7	6	11	8	8	5
11	7	15	7	10	14	7	9	1
12	10	15	8	9	15	7	10	1
13	10	15	12	10	15	6	10	1
14	9	12	14	9	15	10	9	3
15	2	15	15	6	13	8	9	1
16	10	15	7	6	14	8	8	4
17	10	15	9	7	6	8	8	1
18	10	12	7	5	10	8	8	1
19	10	11	12	2	10	8	8	2
20	10	12	12	8	10	8	6	2
21	10	15	15	4	15	8	10	5
22	10	12	11	10	15		8	3
23	10	13	14	10	15	6	8	1
24	10	15	13	10	15	2	10	1
25	10	12	10	7	15	6	8	2
26	10	15	13	10	15	6	10	1
27	10	13	15	8	13	7	8	3
28	10	15	11	10	15	8	10	1

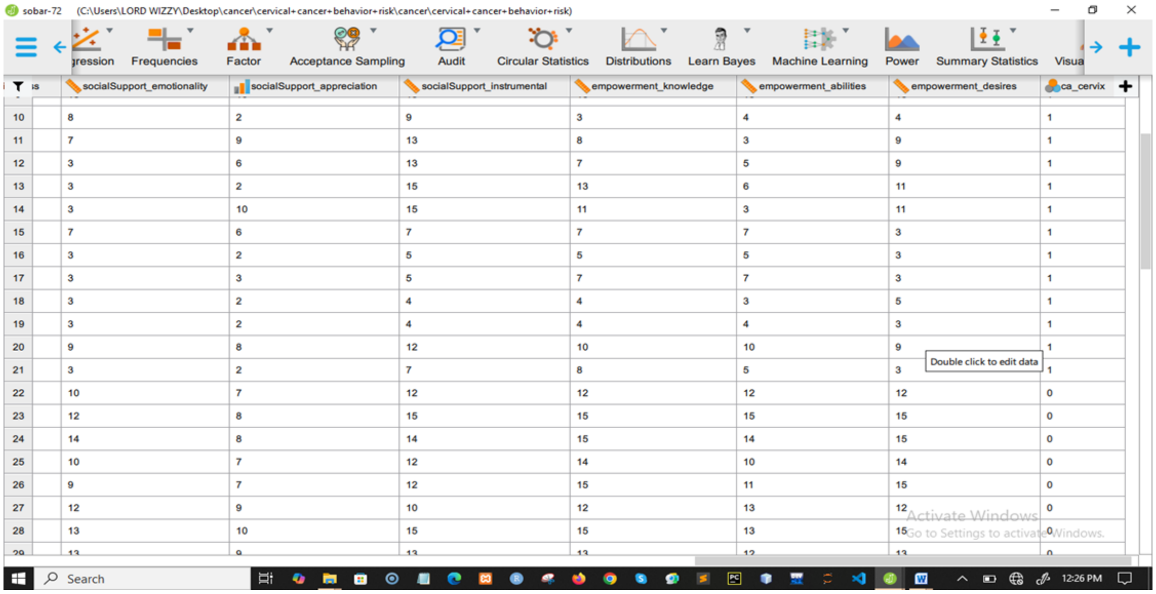


Figure 3: JASP View on the dataset (Fieldwork 2025)

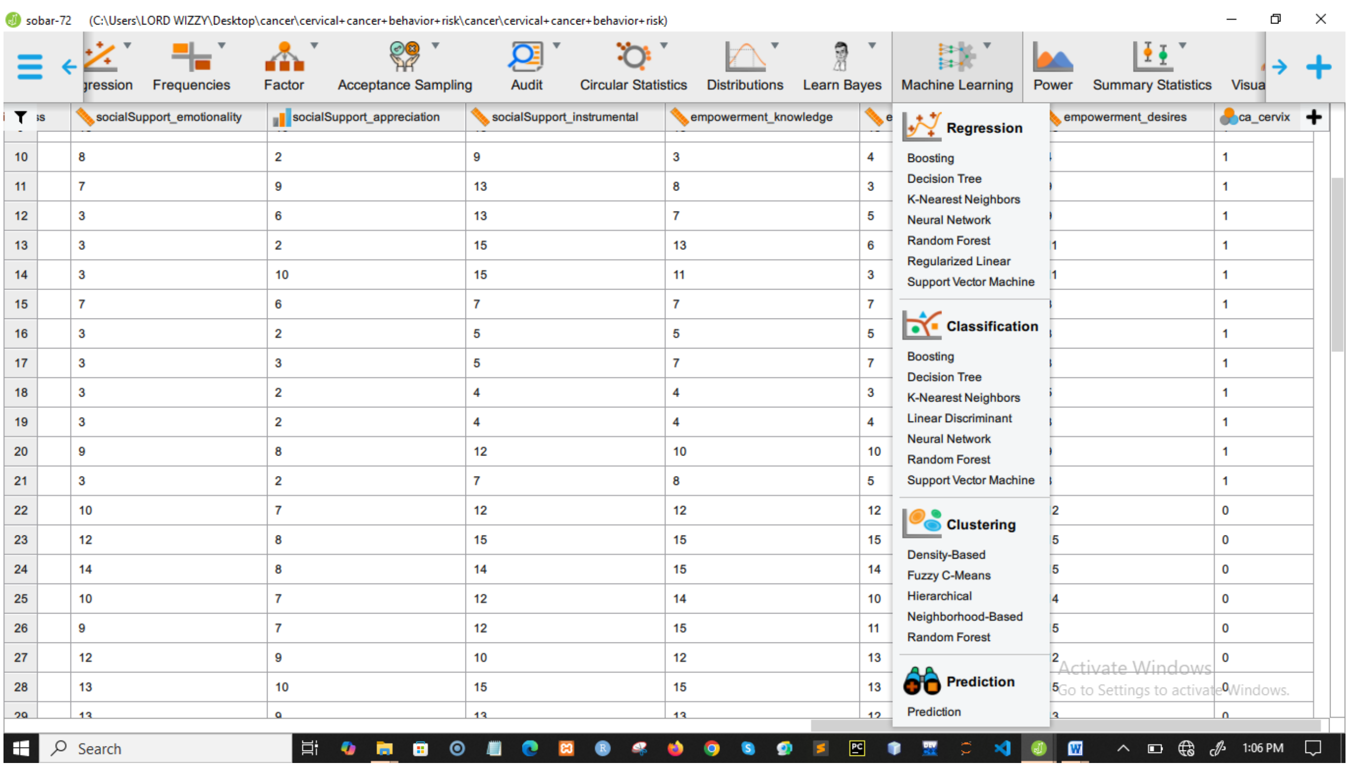


Figure 4a: JASP Platform Selection of the Machine Learning Algorithms (Fieldwork 2025)

Experiment on dataset using Random Forest Classification

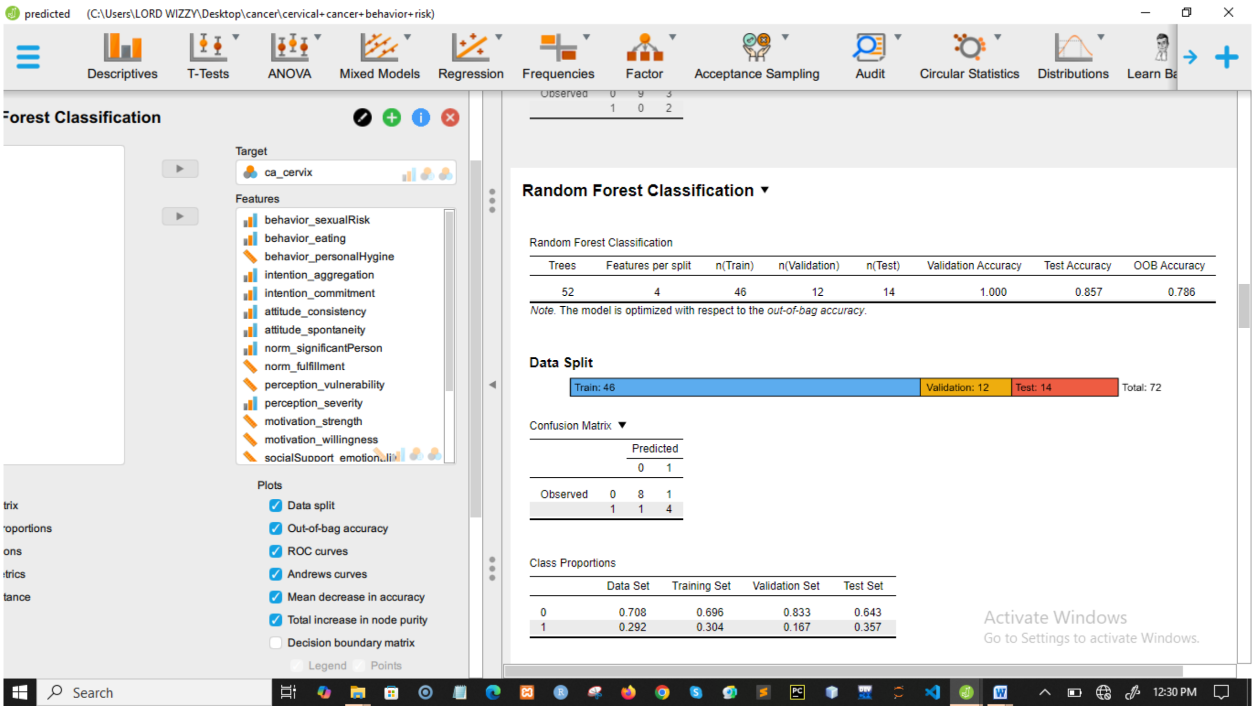


Figure 4b: JASP Platform View on the prediction of dataset using Random Forest Classification (Fieldwork 2025)

Table 2: Random Forest Classification

Trees	Features per split	n(Train)	n(Validation)	n(Test)	Validation Accuracy	Test Accuracy	OOB Accuracy
52	4	46	12	14	1.000	0.857	0.786

Note. The model is optimized with respect to the *out-of-bag accuracy*.

Data Split

Figure 5: Data Split

Table 3: Confusion Matrix

		Predicted	
		0	1
Observed	0	8	1
	1	1	4

Table 4: Class Proportions				
	Data Set	Training Set	Validation Set	Test Set
0	0.708	0.696	0.833	0.643
1	0.292	0.304	0.167	0.357

Table 5: Evaluation Metrics			
	0	1	Average / Total
Support	9	5	14
Accuracy	0.857	0.857	0.857
Precision (Positive Predictive Value)	0.889	0.800	0.857
Recall (True Positive Rate)	0.889	0.800	0.857
False Positive Rate	0.200	0.111	0.156
False Discovery Rate	0.111	0.200	0.156
F1 Score	0.889	0.800	0.857
Matthews Correlation Coefficient	0.689	0.689	0.689
Area Under Curve (AUC)	0.944	0.956	0.950
Negative Predictive Value	0.800	0.889	0.844
True Negative Rate	0.800	0.889	0.844
False Negative Rate	0.111	0.200	0.156
False Omission Rate	0.200	0.111	0.156
Threat Score	2.667	1.333	2.000
Statistical Parity	0.643	0.357	1.000
Note. All metrics are calculated for every class against all other classes.			

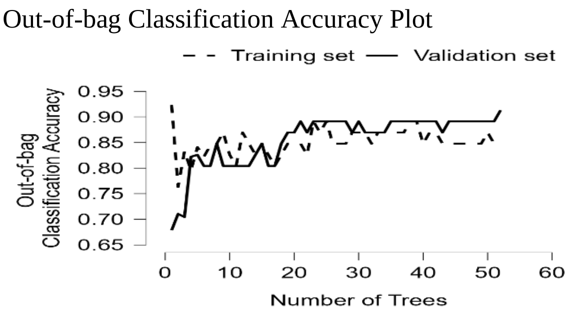


Figure 6: Out-of-bag Classification Accuracy Plot of the model

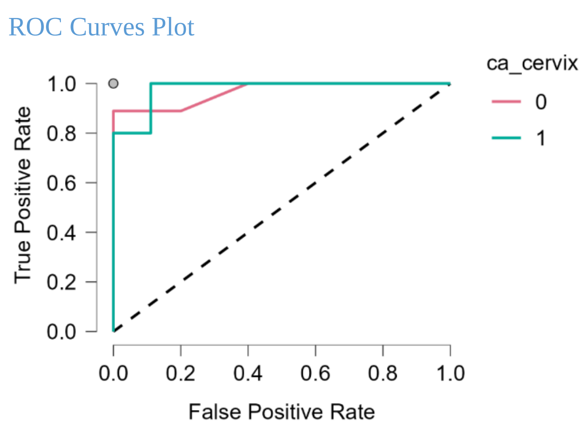


Figure 7: ROC curves Accuracy Plot of the model

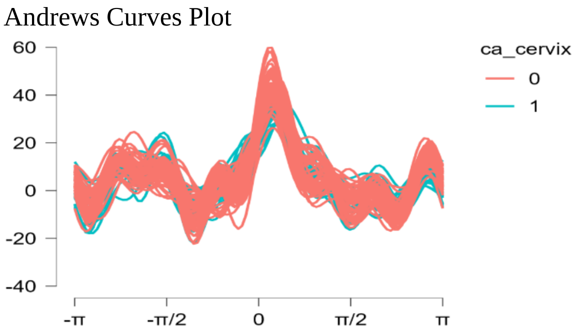


Figure 8: Andrews curves Accuracy Plot of the model

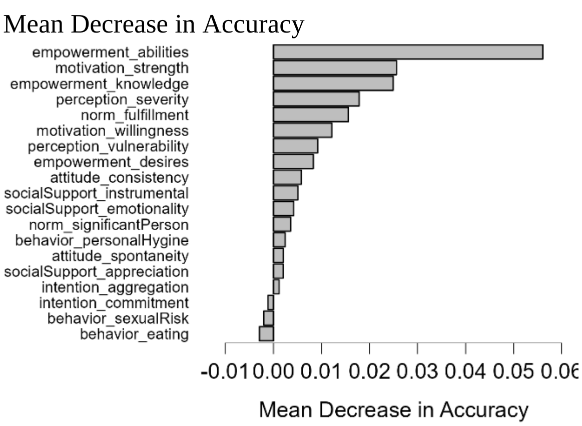


Figure 9: Mean Decrease in Accuracy Plot of the model

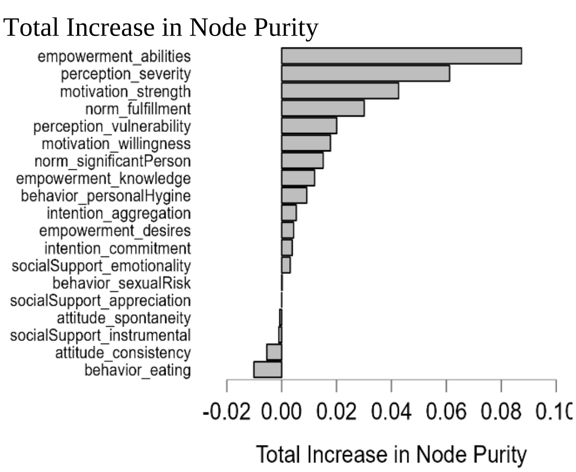


Figure 10: Total Increase in Node Purity Accuracy Plot of the model

Experiment on dataset using Decision Tree Classification Algorithm

Table 6: Decision Tree Classification			
Splits	n(Train)	n(Test)	Test Accuracy
30	58	14	0.786

Figure 11: Decision tree algorithm Data split

Table 7: Confusion Matrix			
		Predicted	
		0	1
Observed	0	7	2
	1	1	4

Table 8: Class Proportions			
	Data Set	Training Set	Test Set
0	0.708	0.724	0.643
1	0.292	0.276	0.357

Table 8: Evaluation Metrics			
	0	1	Average / Total
Support	9	5	14
Accuracy	0.786	0.786	0.786
Precision (Positive Predictive Value)	0.875	0.667	0.801
Recall (True Positive Rate)	0.778	0.800	0.786
False Positive Rate	0.200	0.222	0.211
False Discovery Rate	0.125	0.333	0.229
F1 Score	0.824	0.727	0.789
Matthews Correlation Coefficient	0.559	0.559	0.559
Area Under Curve (AUC)	0.789	0.789	0.789
Negative Predictive Value	0.667	0.875	0.771
True Negative Rate	0.800	0.778	0.789
False Negative Rate	0.222	0.200	0.211
False Omission Rate	0.333	0.125	0.229
Threat Score	1.750	0.800	1.275
Statistical Parity	0.571	0.429	1.000

Note. All metrics are calculated for every class against all other classes.

ROC Curves Plot

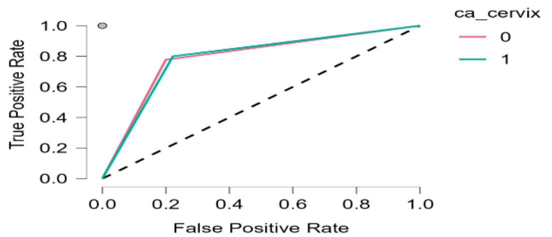


Figure 12: ROC Curves Accuracy Plot

Andrews Curves Plot

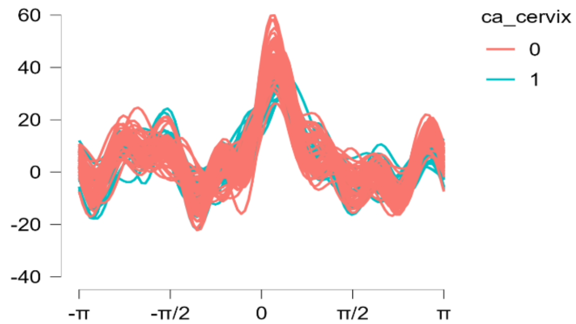


Figure 13: Andrews Curves Accuracy Plot

Decision Tree Plot

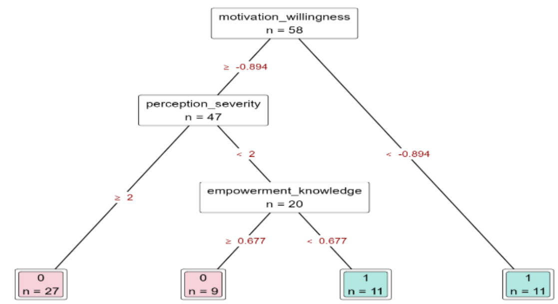


Figure 14: Decision Tree Accuracy Plot

Table 9: RESULT ACCURACY COMPARISON FOR THE TWO DEVELOPED MODELS

ALGORITHMS	F1 SCORE	PRECISION/SUPPORT	ROC CURVE	CONFUSION MATRIX
DECISION TREE	0(NO) = 0.824 : 82% Accuracy 1(YES) = 0.727 : 73% Accuracy	PRECISION (positive predictive value) 0= 0.875 : 88% 1 = 0.667 : 67% SUPPORT 0(NO) = 9 persons 1(YES) = 5 persons	True Positive Rate 0= 0.692 1= 0.791 False Positive Rate 0= 0.200 1= 0.221	0 = 7 1 = 2 AND 0 = 1 1 = 4 The Difference of 5 for 0(N0) and 2 for 1(YES)

RANDOM FOREST	0(NO) = 0.889: 89% Accuracy 1(YES) = 0.800: 80% Accuracy	PRECISION (positive predictive value) 0 = 0.889 : 89% 1 = 0.080 : 80% SUPPORT 0(NO) = 9 persons 1(YES) = 5 persons OOB Accuracy of (0.93) 93%	True Positive Rate 0= 1.000 1= 0.888 False Positive Rate 0= 0.200 1= 0.111	0 = 4 1 = 1 AND 0 = 8 1 = 1 The Difference of 4 for 0(N0) and there is no difference for 1(YES)

THE MODEL RESULT SUMMARY

The experiment conducted on the dataset with 58 trains and 14 tests produced a significant result within the two applied machine learning algorithms producing a new model that helped in predicting the risk factor and presents of cervical cancer amongst women of Umuagwo Ohaji/Egbama LGA of Imo State Nigeria with accuracy status shown in table 9 above.

CONCLUSION

As earlier stated, that the focus of this work is to develop a model using Machine Learning algorithms as one of the branches of artificial intelligence that can help to predict the prevalence of cervical Cancer Risk Factor Amongst Women in Umuagwo Ohaji/Egbema LGA. This research was able to show clearly how cervical Cancer could be managed using hybrid predictive models with outstanding focus on three different variables for both Decision Tree and Random Forest algorithms. DECISION TREE (motivation_willingness at n = 58 accuracy of 0.894, Perception_severity at n= 47 accuracy of 0.200 and empowerment_knowledge at n = 20 accuracy of 0.677) while RANDOM FOREST (empowerment_ability = 0.088, perception_severity = 0.067 and motivation_strength = 0.042),

Due to this major factors as identified problems such as women/men perception, lack of willingness or motivation, empowerment ability/knowledge have been identified by the new developed model as the major risk factor causing the increase of cervical cancer amongst women of Umuagwo Ohaji/Egbema. Also a critical view as done by the scholars further x-ray the stated problem which lead the villagers to believe that “their deity /juju was responsible for the death of their women”

RECOMMENDATION

The researcher therefore recommends the following:

1. An enlightenment program should be conducted for knowledge awareness creation amongst the women of the community
2. There is need for health empowerment programme for Umuagwo Ohaji/Egbema women
3. Different models showed be employed and then compared against each other for accurate data prediction.
4. Full adoption of machine learning tools should be used in solving real life challenging problems more especially in health related problems more especially in rural areas.
5. Other researcher should be carryout a study towards screening of women on cervical cancer outcome.

REFERENCES

[1] Abraham P. and Ziad, M. (2022). Cervical cancer and Pap smear screening in Botswana: knowledge and perceptions. *International Nursing Review*. 2022;50(3):167–175. doi: 10.1046/j.1466-7657.2003.00195.x.

[2] Ansink, C. M. (2018). Knowledge, practice, and barriers toward cervical cancer screening in Elmina, southern Ghana. *International Journal of Women's Health*. 2018;7:31–39. doi: 10.2147/ijwh.s71797.

[3] Basu M. H., Solomon, D and Lawson (2018). Preventing cervical cancer. *Continuing Medical Education*. 2018;27(10):444–449.

[4] Binka, F., Pett M. and Menon U. (2019). Cervical cancer beliefs and pap test screening practices among Chinese American immigrants. *Oncology Nursing Forum*. 2007;34(6):1203–1209. doi: 10.1188/07.onf.1203-1209.

[5] Ho, Tin Kam (1995) *Random Decision Forests (PDF)*. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. *278*–282. Archived from *the original (PDF)* on 17 April 2016. *Retrieved 3 April 2025*.

[6] Jiawei Han and Micheline Kamber, (2001) "Data Mining Concepts and Techniques", Morgan Kauffman Publishers.

[7] Kuguyo, P., Saslow D. and Lawson H. W., (2017). American cancer society, American society for colposcopy and cervical pathology, and American society for clinical pathology

[8] Msyamboza, G., Franco E., Harper D. (2016). Vaccination against human papillomavirus infection: a new paradigm in cervical cancer control. *Vaccine*. 2016;23(17-18):2388–2394. doi: 10.1016/j.vaccine.2016.01.016.

[9] Ngwenya P. and Huang, O. (2018) Cervical cancer prevention: safety, acceptability, and feasibility of a single-visit approach in Accra, Ghana. *American Journal of Obstetrics and Gynaecology*. 2018;196(4):407.e1–407.e9. doi: 10.1016/j.ajog.2006.12.031.

[10] Ozgul, C. (2017). Cervical cancer knowledge, perceptions and screening behaviour among female university students in Ghana. *Journal of Cancer Education*. 2017;31(2):322–327. doi: 10.1007/s13187-015-0852-x.

[11] Rositch R. K. M.,Gatuguta. D. and Choi, R., (2016). Clinical visits and cervical cancer screening in Accra. *Ghana Medical Journal*. 2016;44(2):59–63. doi: 10.4314/gmj.v44i2.68885.

[12] Udigwe, O. (2018). Examining barriers to cervical cancer screening and treatment in Florida through a socio-ecological lens. *Journal of Community Health*. 2018;36(1):121–131. doi: 10.1007/s10900-010-9289-7.

[13] Wong, X. (2019). Knowledge, and beliefs about cervical cancer screening among men in Kumasi, Ghana. *Ghana Medical Journal*. 2019;46(3):147–151.

[14] World Health Organization. (2023). *Comprehensive Cervical Cancer Control: A Guide to Essential Practice*. 2nd. Geneva, Switzerland: WHO; 2023.

[15] Yan, Y., Omisore, O. M., Xue, Y. C., Li, H. H., Liu, Q. H., Nie, Z. D., ... & Wang, L. (2020). Classification of neurodegenerative diseases via topological motion analysis—A comparison study for multiple gait fluctuations. Ieee Access, 8, 96363-96377.