Optimized Machine Learning based Classifications of Staging in Gynecological Cancers using Feature Subset through Fused Feature Selection Process

B Nithya[1]
VTU Research Scholar, Department of MCA
CMR Institute of Technology, Bangalore, India

V Ilango[2]
Professor, Department of MCA
CMR Institute of Technology, Bangalore, India

Abstract—**After diagnosing the cancer, the next step is to identify the staging of the cancer to start with the appropriate treatment plans. There are different kinds of gynaecological cancers and this research lays emphasis on cervical and ovarian cancer types with their staging classifications. The cervical and ovarian cancers data from SEER registry are used in this work. This work intends to propose an optimized classification method for staging prediction in gynaecological cancers through fused feature selection process that aimed to provide an optimal feature subset. The fused feature selection process includes the hybridization of relief filter approach with wrapper method of genetic algorithm to produce revised feature subset of data as an outcome. Accordingly, this work attained an improved feature subset through fused feature selection process for precise classification of cervical and ovarian cancer stages by identifying their significant features. The predictive models are established with 10-fold cross validation using major classification algorithms like C5.0, Random Forest and KNN. The classification results are attained for the respective types of cervical, ovarian cancer stages and the stage-wise classification based on patients age also obtained through this proposed method. The results portrayed that the women in the age group of 45 and above are more critical with the incidence of cervical and ovarian cancer types. Random Forest method has shown progressive accuracy rate with progressive percentage of other performance outcomes. Also, this work recognized that the best and optimal feature subset selection could condense the complexity of the predictive model.**

Keywords—*Ovarian cancer; cervical cancer; diagnosis; gynaecological cancers; staging; feature selection; machine learning; classification*

I. INTRODUCTION

Gynaecological cancer denotes five types of cancers which starts in the reproductive organs of women. Cervical cancer is a form of gynaecological cancer that originates in the cells that line the cervix. This cancer type is most identified in women between ages 35 and 44. The average age at diagnosis is 50 years. Also, there are higher chances of the patients risking the development of cervical cancer as they grow older. Ovarian cancer is a type of gynaecological cancer which is more perilous in recent times. Ovarian cancer is ranked fifth in cancer demises among women [1]. Early detection of ovarian cancer could have a huge influence on the cure rate and it is instantly needed [2], but only 20% are found at a primary stage. The study [3] to find the survival outcome in ovarian

cancer patients insisted that accurate estimation is essential for the reason that prognosis could be a determining factor of medication aggressiveness. Both cervical, ovarian cancers are critical but early detection of these cancers are erratic in most of the women. After diagnosing any type of cancer, it is obligatory to identify the staging to know about how much it has affected the other organs of the body. Staging procedure helps to decide better treatment plans and to know about survival information. Thus, it is essential to identify these conferred types of gynaecological cancer stages in an accurate manner to initiate with effective treatment procedures for the patients. The sections of this paper are structured as follows. In Section 2, the literature study is discussed, Section 3 shows the types of staging classes in cervical and ovarian cancers. In Section 4, the proposed methodology for staging classifications is discussed and the implementation procedure is discussed in Section 5. The classification outcomes of cervical, ovarian cancer stages are discussed in Section 6; in Section 7, the experimental results are shown, and the conclusion is conferred in Section 8.

II. LITERATURE STUDY

Machine Learning (ML) techniques are more effective in various types of cancer diagnosis and staging predictions. The study [4] to diagnose and classify the stages of an ovarian cancer used classification and clustering methods to train the cancer images with respect to the ovarian cancer stages and this work attained 94% of accuracy. This work aimed to improve the sensitivity measure. The work [5] which proposed the system for staging predictions in cervical cancer insisted that genetic algorithms are efficient in processing the huge quantities of information. But the comparative performance of classifiers is not deliberated using various performance metrics. The study [6] which used Gynecologic Cancer Society supported open dataset applied SVM technique and suggested that SVM method accomplishes better results in classifying the stages in cervical cancer. The dataset used in this work is moderately small. For staging predictions in cervical cancer, the study [7] designed CVSS dictionary learning framework by means of multi-view MR images. This work demonstrated the results of classification accuracy in identifying the stages of cervical cancer, however the accuracy is not reasonable. The comparative study [8] using various classifiers to identify the stages in cervical cancer suggested J48 as the suitable method for classifications of stages with the

accuracy of 93%, still the score for sensitivity and specificity was trivial. A decision tree-based procedure applied in the study [9] used cervical cancer data from IGCS for staging classification. This method used correlation-based feature selection and C5 algorithm, the accuracy of this method is not consistent. The staging classification study which included ovarian cancer data combined the outcomes of ten feature selection methods to select the subset for eventual classification to improve the accuracy [10]. The literature study evidenced that the accuracy attained in staging classifications of cervical and ovarian cancer is not adequate. Also, the classifications of these cancer stages among women in various age groups is also mandatory to provide better treatment plans. Accordingly, a methodology has been proposed and the results are enhanced with broad staging classifications for cervical and ovarian cancers of women in various age groups.

III. TYPES OF STAGING CLASSES IN CERVICAL AND OVARIAN CANCERS

Staging is the procedure of identifying the amount of cancer in an individual's body and its setting in the body. This procedure helps to determine how severe the cervical or ovarian cancer is and deciding about the exact and best treatment for the same.

A. Cervical Cancer Staging Classes

Cervical cancer's most common staging classification is the FIGO system - International Federation of Gynaecology and Obstetrics. The cervical cancer stages are summarized in Table I.

TABLE I. CERVICAL CANCER STAGES

FIGO Stage	Stage Description
Stage I	Cancer has spread into the deeper tissue from the cervix liner.
IA	Tumour is < 5 mm deep, and < 7 mm wide.
IA1	Tumour's depth is not > 3 mm and < 7 mm wide in tissues.
IA2	Tumour depth is >3 mm, but < 5 mm and < 7 mm wide.
IB	Tumour is in the cervix, and the size is wider than in IA2.
IB1	Tumour is not > 4 cm at its widespread part.
IB2	Tumour is not < 4 cm at its widespread part.
Stage II	Cancer has spread outside of the cervix to closer parts.
IIA	Tumour has not spread into tissues adjacent the cervix and uterus.
IIA1	Tumour is not > 4 cm at its extensive part.
IIA2	Tumour is not < 4 cm at its extensive part.
IIB	Tumour has spread adjacent to the cervix and uterus.
Stage III	Tumour has spread to the pelvis walls.
IIIA	Tumour has grown into the lower part of the vagina.
IIIB	Tumour has grown into the pelvis walls and has blocked a ureter.
IVA	Tumour has grown into the bladder, rectum.
IVB	Cancer has spread to further parts of the body.

B. Ovarian Cancer Staging Classes

The FIGO (International Federation of Gynaecology and Obstetrics) system and the AJCC (American Joint Committee on Cancer) TNM staging system are the two main systems used for classification staging in ovarian cancer. The ovarian cancer stages are shown in Table II.
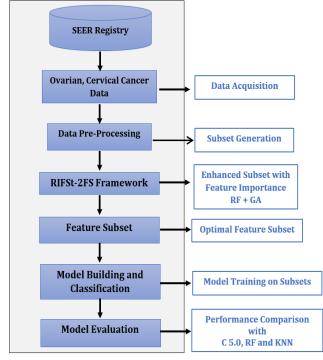
TABLE II. OVARIAN CANCER STAGES

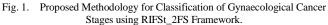| FIGO Stage	Stage Description
Stage I	Tumour curbed to ovaries.
IA	Tumour is limited to only one ovary.
IB	Tumour affects both the ovaries.
IC	Tumour covers one/both ovaries with subsequent consequences.
Stage II	Cancer has affected one/both ovaries, spreads to other pelvic areas.
IIA	Cancer has the extension and/or implant on uterus
IIB	Cancer has the extension to other pelvic intraperitoneal tissues.
IIC	This stage comprises IIA or IIB with positive washings/ascites.
Stage III	Tumour encompasses one/both ovaries with cytologically formed, spread to the peritoneum outside the pelvis.
IIIA	Cancer covers the pelvis only, but the cancer cells which are visible only through a microscope are spread to the out of the peritoneum.
IIIB	Cancer has moved to the peritoneum; its size is <= 2 cm.
IIIC	Cancer has moved to the peritoneum which is not < 2 cm and/or it has moved to the abdominal lymph nodes.
Stage IV	Cancer has affected the area outside the abdomen to other organs, such as the lungs or the tissue inside the liver.

IV. PROPOSED METHODOLOGY FOR STAGING CLASSIFICATIONS

It is evident through various inquiries that combined feature selection approaches are effectual in handling high dimensional data and proficient in achieving enhanced classification results [11][12]. This research aimed to provide an optimized classification method for staging predictions in cervical and ovarian cancers data with enhanced performance outcome. Consequently, a methodology is proposed here with Revised and Improved Feature Subset through Fused Feature Selection process (RIFSt_2FS) framework. The proposed methodology is depicted through Fig. 1.

After obtaining the data from the registry, the pre-processing of data is required to remove the missing values and formatting of the data. Initial feature set is generated with initial features. Inappropriate and superfluous features need to be removed to attain an effective classifier model [13]. To attain an enhanced feature subset the procedure mentioned in the RIFSt_2FS framework is implemented using Relief and Genetic Algorithm. After attaining the enhanced and revised feature subset the prominent classification algorithms in ML are applied for various types of staging classifications. The best and optimized classification approach is selected based on the evaluation of various models.

Revised and Improved Feature Subset through Fused Feature Selection process (RIFSt_2FS) framework is designed as shown in Fig. 2 which is intended at an optimal feature

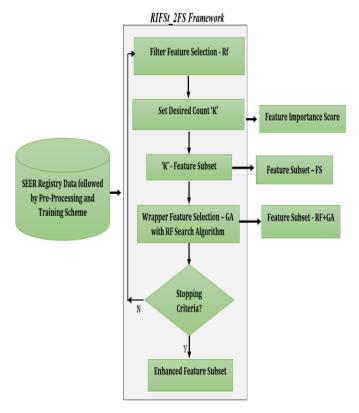subset for improved prediction performance for staging classifications in gynaecological cancers.

Fig. 1.   Proposed Methodology for Classification of Gynaecological Cancer Stages using RIFSt_2FS Framework.

Fig. 2.   RIFSt_2FS Framework.

V.   IMPLEMENTATION OF RIFST_2FS FRAMEWORK FOR STAGING CLASSIFICATION IN GYNAECOLOGICAL CANCER DATA

The implementation procedures of the proposed framework for gynaecological cancer staging prediction with fused feature selection process are conferred in this section.

A. Procedure for Classification of Cervical and Ovarian Cancer Stages with RIFSt_2FS Framework

The illustrative procedure to implement the proposed methodology is designed and the sequence of phases in the process are as follows.

Step 1: Begin
Step 2: Select gynaecological cancer data from SEER registry using cancer type, province, and year of diagnosis
Step 3: Import dataset as .csv file from SEER registry
Step 4: Apply Data Pre-Processing
Step 5: Plot preliminary feature importance
Step 6: Prepare training scheme
Step 7: Call the procedure RIFSt_2FS (Feature Importance using Relief and Genetic Algorithm)
Step 8: Generate Feature Subset FS
Step 9: Obtain classification result for initial set of features
  Repeat {
- *Train and build the models with preliminary features through repeated K-fold c.v using ML classifiers C5.0, RF and KNN*
- *Generate variable importance score*
     }

Until the specified number of required models' attainment
Step 10: Train and build the models with the feature subsets generated through RIFSt_2FS with repeated K-fold c.v using ML classifiers C5.0, RF and KNN
  Repeat {
- *Acquire Classification results with various stages of cervical, ovarian cancers*
- *Obtain age-wise classification of cervical, ovarian cancer stages*
     }
  Until the attainment of specified classifications of various cancer stages and ages
Step 11: Evaluate the performance of all the models with test data
Step 12: Designate the optimal feature subset and classification approach
Step 13. End

B. Dataset

The SEER [14] (Surveillance, Epidemiology, and End Results) is a database which contains largest and greatest comprehensive information on all the types of cancer incidences. This registry has the cancer patients' populations' data from America and from Asian/Pacific regions as it is the worldwide cancer data collection system. To select only relevant data, few conditions were made to decide the pertinent cases. The conditions are like region (Asian/Pacific Islander/Indian patients), cancer type as ovary (C56.9),

cervical (C53.9) with confirmed diagnosis and the year of diagnosis ranged from 2000 to 2017. The selected instances were organized as CSV files.

C. Data Pre-Processing

The selected instances had some missing values for few columns. Missing values were replaced with mean or mode based on the attribute types and few rows were removed which had major number of missing values in their fields. After pre-processing, the cervical cancer dataset comprises 4062 instances and the ovarian cancer dataset had 5843 instances. The cervical and ovarian cancer data instances are segregated as training and test data for training and performance evaluation of prediction models.

D. Feature Selection

Feature selection is the vital stage in obtaining optimized prediction models with advanced classification accuracy through reduced number of features of the dataset. As our objective is to obtain reduced and an improved feature subset, we intended to apply fused feature selection process by integrating filter and wrapper methods. The pre-processed dataset had more than 30 attributes which represented various test results of the patients along with their age and marital status. The integrated approach of feature selection process aimed to reduce these attributes and thereby getting an enhanced and improved feature subset which can be used for classifying the cancer stages. The expected outcome of this phase is to obtain an optimal feature subset which contains major significant features which are essential for staging classification of cervical, ovarian cancer data.

E. Incorporation of Filter and Wrapper Feature Selection Methods

There are several methods available for feature selection processes using filter and wrapper approaches. It is evident through various researches that the Relief algorithm which uses filter approach and Genetic algorithm with wrapper approaches are more effective in gaining optimal feature subsets by assigning precise rankings and selecting relevant features of the dataset. Based on our prior findings, the filter feature selection approach of relief method is fused with wrapper approach using genetic algorithm to attain a fused performance by fabricating an optimal feature subset of mentioned cancer datasets. The phases of the fused feature selection process anticipated for the generation of the feasible combinations of improved feature subsets is described below:

1) Procedure: RIFSt_2FS – Revised and Improved Feature Subset through Fused Feature Selection Process.

The feature subset obtained through this integrated feature selection process is expected to have an enhanced and revised feature which could be used for the succeeding phase of classification of discussed gynaecological cancer stages.

Input: Cervical, Ovarian Cancer training dataset with primary features
Output: Feasible groupings of enhanced feature subsets which are ideal for staging prediction
Phase 1 - Filter Phase - Obtaining refined datasets through filter-based FS process
Determine the Number of Desired Features
Apply Relief Algorithm
Obtain Dataset (Rf) - Features nominated through Relief
Phase 2 - Wrapper Phase - Wrapper-based FS using Random Forest (RF) for the refined datasets attained in Phase 1
- *Initialise this phase with the features selected in Phase 1*
- *Repeat this phase until termination criteria of Genetic Algorithm's*

Dataset obtained from phase 1 (Rf) applied to wrapper approach with Genetic Algorithm (GA) by assessing the fitness with Random Forest (RF) to accomplish a compact feature subset as Rf_GA.

F. Gynaecological Cancer Stages Classification with K-Fold Cross Validation

For classifying all the stages of cervical and ovarian cancers, prominent ML classification algorithms are executed on the datasets accomplished with optimal features. Repeated K-fold cross validation technique is recommended for model training and in the process of building various classifiers through ML classification algorithms for determining an optimized classification technique [15]. The prevalent classification algorithms used for model training and validation are Random Forest, C5.0 and K-Nearest Neighbor. The classifier models are initially constructed with conferred ML algorithms using all the preliminary features and with the feature subset gained through fused feature selection process by combing Relief and Genetic Algorithm. The classification results are conferred under Results Section with detailed analysis of stagewise and age-wise groupings.

G. Performance Evaluation

The performance of the predictive models obtained through ML classification algorithms are assessed based on performance metrics accuracy. Initially, the predictive performance of algorithms such as Random Forest, C5.0 and K-Nearest Neighbour are evaluated on training datasets with preliminary features by means of applying test data for validation. Subsequently the model generated through fused feature selection process is assessed by means of test data with the stated classification algorithms based on their proficiency in classifying the data to the appropriate cervical, ovarian cancer stages. The performance metrics used in this work s accuracy, which is calculated as shown below.

Accuracy = (Total No. of correct predictions)/(Total No. of instances)

The optimal feature subset generated through RIFSt_2FS method is effectual for appropriate staging classifications of cervical, ovarian cancer with extreme accuracy and improved performance outcomes. This approach has shown prominent results as compared with the existing techniques with image classifications [16], [17] The outcomes are conferred in the subsequent section.

VI. DISCUSSION

The proposed framework is efficient in identifying the important and relevant features which are to be designated for cervical and ovarian cancer staging classifications.

A. Variable Importance

The significant features derived through RIFSt_2FS method are termed as an optimal feature subset for staging classification of gynaecological cancers data. The overall variable importance in cervical and ovarian staging classifications is obtained and the chart of C5.0 method for combined staging prediction in ovarian cancer is depicted in Fig. 3.

B. Classification of Cervical and Ovarian Cancers Test Data

Based on the dataset retrieved from SEER Registry, we were able to predict the stages for 1078 cervical cancer patients and 1446 ovarian cancer patients which are specified as test data in this work. This work aimed to classify the stages of cervical and ovarian cancers with their sub types and through the patients' age-wise aspects through each stage of cervical and ovarian cancers.

1) Comprehensive Stagewise Classification Results of Cervical and Ovarian Cancers. The inclusive classification of all the stages of cervical cancer is shown in Table III and Fig. 4. It is obvious through the results that incidence of Stage IIIB and Stage IB1 cervical cancer are higher and Stages III and IV are considered as more critical.
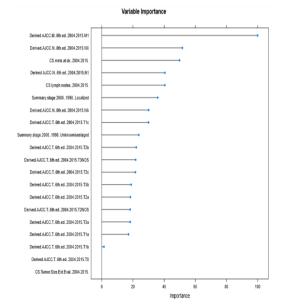

Fig. 3.   Variable Importance in C5.0 Algorithm for all the Stages of Ovarian Cancer.

TABLE III.   CLASSIFICATION OF STAGE I TO IV OF CERVICAL CANCER USING RIFST_2FS FRAMEWORK

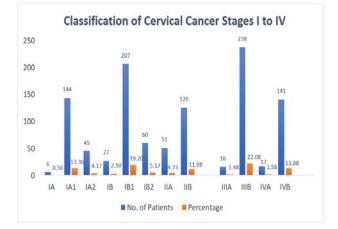| STAGES	Type	Patients	Percentage
STAGE I N=489, Ratio=45.36	IA	6	0.56
	IA1	144	13.36
	IA2	45	4.17
	IB	27	2.50
	IB1	207	19.20
	IB2	60	5.57
STAGE II N=177, Ratio=16.42	IIA	51	4.73
	IIB	126	11.69
STAGE III N=254, Ratio=23.56	IIIA	16	1.48
	IIIB	238	22.08
STAGE IV N=158, Ratio=14.66	IVA	17	1.58
	IVB	141	13.08

Fig. 4.   Classification of Stage I to IV of Cervical Cancer using RIFSt_2FS Framework.

In a similar way, the inclusive classification of all the stages of ovarian cancer is shown in Table IV and Fig. 5. It is obvious through the results that incidence of Stage IIIC and Stage IV ovarian cancers are higher besides considered as more critical. The study [18] to know the implication among diagnostic patterns and stages in ovarian cancer using medical indicative features and symptoms insisted that self-attention is vital for all women.

2) Age-Wise Classification Aspects with Specific Stages of Cervical and Ovarian Cancer. The comprehensive classification of all the stages of cervical cancer based on the age groups is depicted in Fig. 6. It is obvious through the outcomes that the women in the age group 35-44 and 45-54 are more critical to be affected with all the types of cervical cancer stages of 1 to 4. The stages IB1 and IIIB is higher. The women in the age group 35-44 are also having more percentage of occurrences in the stages IA1, IB1 and IIIB. So consistent follow-up and timely treatment could diminish the critical and life-threatening situations for those women. Consequently, the comprehensive classification of all the stages of ovarian cancer based on the age groups is depicted in Fig. 7.

TABLE IV.   CLASSIFICATION OF STAGE I TO IV OF OVARIAN CANCER USING RIFST_2FS FRAMEWORK

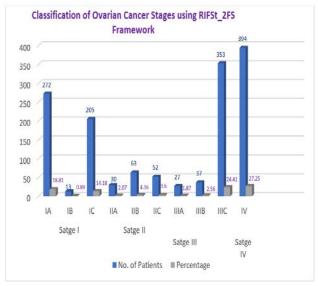| STAGES	Type	Patients	Percentage
STAGE I N=490, Ratio=33.88	IA	272	18.81
	IB	13	0.89
	IC	205	14.18
STAGE II N=145, Ratio=10.03	IIA	30	2.07
	IIB	63	4.36
	IIC	52	3.6
STAGE III N=417, Ratio=28.84	IIIA	27	1.87
	IIIB	37	2.56
	IIIC	353	24.41
STAGE IV N=394, Ratio=27.25	IV	394	27.25

Fig. 5.   Classification of Stage I to IV of Ovarian Cancer using RIFSt_2FS Framework.
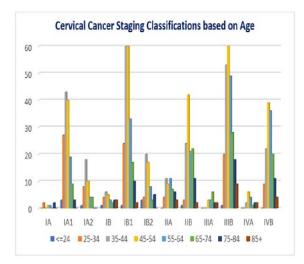

Fig. 6.   Classification of Cervical Cancer Stages I to IV based on Age.
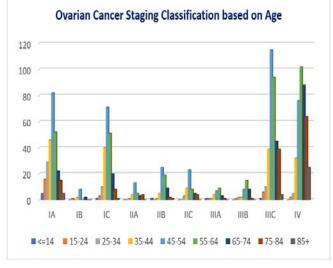

Fig. 7.   Classification of Ovarian Cancer Stages I to IV through Age.

It is apparent through the findings that the women in the age group 45-54 are more critical to be affected with all the types of ovarian cancer stages of 1 to 4. The women in the age group 55 and above are more affected with stage 4. Correspondingly, this classification shows that the ovarian cancer stage IIIC is having more incidence in the women with age group 45-54.

VII. RESULTS

Random Forest and C5.0 classifiers have tremendously performed well in categorizing all the stages of cervical and ovarian cancer types using RIFSt_2FS Framework with precise results. The results proved that RF and C5.0 are the finest classifiers and the results attained through KNN are not satisfactory. Consequently, we have attained an optimized classification results using Random Forest classifier. The performance results of the classifiers are shown in Table V and Fig. 8.

The performance of proposed approach with RF classifier is aggregated and compared with some of the existing studies and the findings are shown in Table VI and Fig. 9.

TABLE V.   PERFORMANCE OF ML CLASSIFIERS IN CERVICAL AND OVARIAN CANCER STAGING CLASSIFICATIONS

ML Classifiers with Performance Measure		Accuracy		
		C5.0	RF	KNN
Cervical Cancer Staging Types	Stage I	97.4	97.6	93.9
	Stage II	95.2	96.3	87.6
	Stage III	96.4	96.8	83.5
	Stage IV	97.2	97.2	89.2
Ovarian Cancer Staging Types	Stage I	97.6	97.6	91.8
	Stage II	95.9	97.3	86.9
	Stage III	97.3	97.5	85.4
	Stage IV	97.2	97.2	88.6

Fig. 8. Performance Analysis of ML Classifiers in Cervical and Ovarian Cancer Staging Classifications.

TABLE VI. PERFORMANCE ANALYSIS OF PROPOSED APPROACH RIFST-2FS

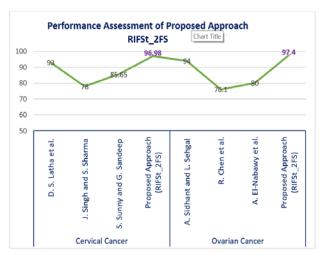| Staging Classification	Research Studies	Accuracy (%)
Cervical Cancer	D. S. Latha et al. [8]	93
	J. Singh and S. Sharma [19]	78
	S. Sunny and G. Sandeep [20]	85.65
	Proposed Approach (**RIFSt_2FS**)	**96.98**
Ovarian Cancer	A. Sidhant and L. Sehgal [4]	94
	R. Chen et al. [21]	76.1
	A. El-Nabawy et al. [22]	80
	Proposed Approach (**RIFSt_2FS**)	**97.4**

Fig. 9. Performance Assessment of Proposed Approach RIFSt_2FS.

VIII. CONCLUSION AND FUTURE WORK

This work is intended to attain for an improved feature subset through fused feature selection process for precise classification of cervical and ovarian cancer stages by identifying the significant features. The integration of feature

selection methods through fused approach enhances the performance of staging classifications through its positive, negative predicted values of the results with uppermost accuracy and improved performance measures. The predictive models are established with 10-fold cross validation using major classification algorithms like C5.0, Random Forest and KNN procedures. The classification results are attained for the respective types of cervical and ovarian cancer stages and the stage-wise classification based on patients age also obtained through this proposed method.

This proposed method has shown improved performance outcomes than the studies discussed in the literature. The results portrayed that the women in the age group of 45 and above more critical with the incidence of cervical and ovarian cancer types. It is mandatory for all the women to have a regular follow-up and timely treatments to reduce the complications in the advanced stages. Random Forest method has shown progressive accuracy rate with 97 percentage of combined performance outcomes. C5.0 algorithms has also shown improved accuracy in all the types of staging classifications of cervical and ovarian cancers. But the performance of KNN algorithm is comparatively less than RF and C5.0 methods. The experiments revealed that through enactment of fused feature selection approach an optimal and reduced feature subset is appropriate for the improvement of classification accuracy with a reduced computational cost. Also, this work recognized that the best and optimal feature subset selection could condense the complexity of the predictive model.

In future work, the staging classifications for other types of gynaecological cancers like uterine, vaginal, vulvar cancers will be analyzed using further types of ML classifiers with other performance metrics like sensitivity, specificity, precision, and F-score values.

REFERENCES

[1] The American Cancer Society Medical and Editorial content Team, "Key Statistics for Ovarian Cancer," Www.Cancer.Org, 2018.

[2] S. Marcišauskas, B. Ulfenborg, B. Kristjansdottir, S. Waldemarson, and K. Sundfeldt, "Univariate and classification analysis reveals potential diagnostic biomarkers for early stage ovarian cancer Type 1 and Type 2," J. Proteomics, vol. 196, no. January, pp. 57–68, 2019, doi: 10.1016/j.jprot.2019.01.017.

[3] E. S. Paik et al., "Prediction of survival outcomes in patients with epithelial ovarian cancer using machine learning methods," J. Gynecol. Oncol., vol. 30, no. 4, pp. 1–13, 2019, doi: 10.3802/jgo.2019.30.e65.

[4] A. Sidhant and L. Sehgal, "Ovarian Cancer Detection Using K-Svm Algorithm," Int. J. Comput. Sci. Eng., vol. 6, no. 12, pp. 182–188, 2018, doi: 10.26438/ijcse/v6i12.182188.

[5] S. R. P. Singh, "Genetic Algorithms for Staging Cervical Cancer," Genet. Algorithms Staging Cerv. Cancer, vol. 3, no. Ii, pp. 39–43, 2013.

[6] C. J, N. M, and A. S Pillai, "Predicting Cervical Carcinoma Stages Identification using SVM Classifier," Int. J. Comput. Trends Technol.,

vol. 22, no. 3, pp. 122–125, 2015, doi: 10.14445/22312803/ijctt-v22p125.

[7]   N. Bnouni, I. Rekik, M. S. Rhim, and N. Essoukri Ben Amara, "Cross-view self-similarity using shared dictionary learning for cervical cancer staging," IEEE Access, vol. 7, pp. 30079–30088, 2019, doi: 10.1109/ACCESS.2019.2902654.

[8]   D. S. Latha, P. V Lakshmi, and S. Fathima, "Staging Prediction in Cervical Cancer Patients – A Machine Learning Approach," vol. 2, no. 2, pp. 14–23, 2014.

[9]   S. Sharma, "Cervical Cancer stage prediction using Decision Tree approach of Machine Learning," Int. J. Adv. Res. Comput. Sci. Commun. Eng., vol. 2, no. 2, pp. 345–348, 2016, doi: 10.17148/IJARCCE.2016.5488.

[10]   E. Saghapour, S. Kermani, and M. Sehhati, "A novel feature ranking method for prediction of cancer stages using proteomics data," PLoS One, vol. 12, no. 9, pp. 1–17, 2017, doi: 10.1371/journal.pone.0184203.

[11]   A. Ç. Pehlivanlı, "A novel feature selection scheme for high-dimensional data sets: four-Staged Feature Selection," J. Appl. Stat., vol. 43, no. 6, pp. 1140–1154, 2016, doi: 10.1080/02664763.2015.1092112.

[12]   O. A. Alomari, A. T. Khader, M. A. Al-Betar, and Z. A. Alkareem Alyasseri, "A hybrid filter-wrapper gene selection method for cancer classification," 2nd Int. Conf. BioSignal Anal. Process. Syst. ICBAPS 2018, pp. 113–118, 2018, doi: 10.1109/ICBAPS.2018.8527392.

[13]   B. Nithya and V. Ilango, "An Optimized Data Classifier Model Based Diagnosis and Recurrence Predictions of Gynecological Cancer for Clinical Decision Support System," in Lecture Notes on Data Engineering and Communications Technologies, 2020.

[14]   SEER, "Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) Research data (1973-2011).," National Cancer institutes, DCCPS, Surveillance Research Program, 2013.

[15]   B. Nithya and V. Ilango, "Machine Learning Aided Fused Feature Selection based Classification Framework for Diagnosing Cervical Cancer," in Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020, 2020, doi: 10.1109/ICCMC48092.2020.ICCMC-00011.

[16]   B. Kaur, K. S. Mann, and M. K. Grewal, "Ovarian cancer stage based detection on convolutional neural network," Proc. 2nd Int. Conf. Commun. Electron. Syst. ICCES 2017, vol. 2018-Janua, no. Icces, pp. 855–859, 2018, doi: 10.1109/CESYS.2017.8321206.

[17]   A. Bentaieb, M. S. Nosrati, H. Li-Chang, D. Huntsman, and G. Hamarneh, "Clinically-inspired automatic classification of ovarian carcinoma subtypes," J. Pathol. Inform., vol. 7, no. 1, 2016, doi: 10.4103/2153-3539.186899.

[18]   K. Bogie et al., "Associations between diagnostic patterns and stages in ovarian cancer," Model Assist. Stat. Appl., vol. 12, no. 3, pp. 275–285, 2017, doi: 10.3233/MAS-170402.

[19]   J. Singh and S. Sharma, "Prediction of Cervical Cancer Using Machine Learning Techniques," Int. J. Appl. Eng. Res. ISSN, vol. 14, no. 7, pp. 150–153, 2019, doi: 10.2174/2213275912666190617160834.

[20]   S. Sunny and G. Sandeep, "Decision Tree approach in Machine Learning for Prediction of Cervical Cancer Stages using WEKA," Int. J. Recent Trends Eng. Res., vol. 02, no. 08, pp. 74–83, 2016.

[21]   R. Chen, A. M. Rosado, and J. Zhang, "Machine learning for ovarian cancer: lasso regression-based predictive model of early mortality in patients with stage I and stage II ovarian cancer," medRxiv, Cold Spring Harb. Lab. Press, 2020, doi: 10.1101/2020.05.01.20088294.

[22]   A. El-Nabawy, N. El-Bendary, and N. A. Belal, "Epithelial Ovarian cancer stage subtype classification using clinical and gene expression integrative approach," Procedia Comput. Sci., vol. 131, pp. 23–30, 2018, doi: 10.1016/j.procs.2018.04.181.