International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307 (Online), Volume-13 Issue-2, May 2023

Survival Prediction of Cervical Cancer Patients using Genetic Algorithm-Based Data Value Metric and Recurrent Neural Network

Ojie, D.V, Akazue M, Omede E.U, Oboh E.O, Imianvan A

Abstract: ***Survival analysis and machine learning has been shown to be an indispensable aspect of disease management as it enables practitioners to understand and prioritize treatment mostly in terminal diseases. Cervical cancer is the most common malignant tumor of the female reproductive organ worldwide. Survival analysis which is a time –to –event analysis for survival prediction is therefore needed for cervical cancer patients. Data Value Metric (DVM) is an information theoretic measure which uses the concept of mutual information and has shown to be a good metric for quantifying the quality and utility of data as well as feature selection. This study proposed the hybrid of Genetic Algorithm and Data Value Metric for feature selection while Recurrent Neural Network and Cox Proportionality Hazard ratio was used to build the survival prediction model in managing cervical cancer patients. Dataset of 107 patients of cervical cancer patients were collected from University of Benin Teaching Hospital, Benin, Edo State and was used in building the proposed model (RNN+GA-DVM). The proposed system outperform the existing system as the existing system had accuracy of 70% and ROC score of 0.6041 while the proposed model gave an accuracy of 75.16% and ROC score of 0.7120 respectively. From this study, It was observed using the GA_DVM features selection that the variables highly associated with cervical cancer mortality are age_at_diagnosis, Chemotherapy, Chemoradiation, Histology, Comorbidity, Menopause, and MENO_Post. Thus, with early diagnosis and proper health management of cervical cancer, the age of survival of cervical cancer patients can be prolonged.***

Keywords**: Cervical cancer, Cox Proportional Hazard, Machine Learning, Survival Model.***

I. INTRODUCTION

Computing and information technologies in recent times have shown to be an invaluable asset in all walks of life owing to its disruptive and yet positive tendencies in complementing and making life easier for mankind.

Manuscript received on 12 April 2023 | Revised Manuscript received on 20 April 2023 | Manuscript Accepted on 15 May 2023 | Manuscript published on 30 May 2023.

*Correspondence Author(s)
Ojie Deborah Voke*, Department of Software Engineering, University Delta, C Agbor, Nigeria. Email: Deborah.ojie@unidel,edu.ng, ojiedv7@gmail.com, ORCID ID: https://orcid.org/0009-0007-9579-7444
Dr. Akazue M, Department of Computer Science, Delta State University, Abraka, Nigeria. Email: akazue@delsu.edu.ng, ORCID ID: https://orcid.org/0000-0003-2518-3889
Dr. Omede E. U, Department of Computer Science, Delta State University, Abraka, Nigeria. Email: edithomede@delsu.edu.ng, ORCID ID: https://orcid.org/0000-0002-9627-0552
Dr. Oboh E.O, Department of Radiotherapy/ Clinical Oncology, University of Benin Teaching Hospital, Edo State. Email: Oseva2good@yahoo.com, ORCID ID: https://orcid.org/0000-0003-2636-1333
Prof. Imianvan A., Department of Computer Science, University of Benin, Benin, Edo Nigeria. Email: tonyvanni@uniben.edu.ng

It has progressively crept and weaves itself into the fabric of human existence making it nearly indistinguishable from human being as they can hardly stay away from the computing devices [8]. Medical care and its associated processes are not left out in this new wave of technology-induced environment and has since become an essential ingredient in the twenty first century medical care. [8];[6] Medical diagnostic models built on the concept of Soft-Computing (SC), Artificial Intelligence (AI) and Machine learning (ML) has been widely accepted and implemented in different problem domain [1]. These techniques have been used independently or in collaboration with certain tools in creating models which enhances existing capabilities. ML, extract hidden patterns from synthesized, structured and unstructured clinical data in conceptualizing intelligent models with the ability to predict varied illnesses [3].These predictions are obtained through training instances visualized from medical dataset. Pattern Recognition (PR) and Machine Learning (ML) algorithms, perhaps are the avalanche of data prediction (James et al., 2001. The varied nature of ML algorithms – supervised, unsupervised and reinforced – has provided a unique trail in data classification and prediction [2] Cervical cancer is experienced by women were the cancer cells 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. According to (Joanne and Mark, 2013), Cervical cancer is the most common major cause of death in women worldwide from the last few decades. It is one of the main types of cancer after the lungs and breast cancer among women and is prone to higher medical burden on the patients and their families. (WHO, 2019). The disparity in survival rate between developing and developed nations is pathetic, 33 – 77% is highly unfathomable and must be reduced (WHO, 2019). Cervical cancer is the fourth most common cancer among women globally, with an estimated 604 000 new cases and 342 000 deaths in 2020. About 90% of the new cases and deaths worldwide in 2020 occurred in low- and middle-income countries (WHO, 2020). According to (WHO, 2021) the mortality to incidence ratio of cervical cancer in Nigeria is at 0.66 which is very high[13].

These statistics highlights the need for a coordinated effort toward improving the extent and quality of services for cervical diagnoses. [5] opines that survival rate was greater in postmenopausal cervical cancer patients than premenopausal cervical cancer patients. It is therefore imperative to have a system that can predict the overall survivability of a person based on her characteristics or covariates in order to ensure proper management of the cervical cancer by helping the doctors decides which treatment provides the most benefit. This research propose a system that predicts the survivability of cervical cancer patient's as the risk of failure or death to ensure appropriate diagnosis and management of cervical cancer patients. The research used Genetic Algorithm and Data Value Metric for the feature selection while Neural Network in combination with Cox Proportionality Hazard model was used to build the survival prediction model. and 342 000 deaths in 2020. About 90% of the new cases and deaths worldwide in 2020 occurred in low- and middle-income countries (WHO, 2020). According to (WHO, 2021) the mortality to incidence ratio of cervical cancer in Nigeria is at 0.66 which is very high[13]. These statistics highlights the need for a coordinated effort toward improving the extent and quality of services for cervical diagnoses. [5] opines that survival rate was greater in postmenopausal cervical cancer patients than premenopausal cervical cancer patients. It is therefore imperative to have a system that can predict the overall survivability of a person based on her characteristics or covariates in order to ensure proper management of the cervical cancer by helping the doctors decides which treatment provides the most benefit. This research propose a system that predicts the survivability of cervical cancer patient's as the risk of failure or death to ensure appropriate diagnosis and management of cervical cancer patients. The research used Genetic Algorithm and Data Value Metric for the feature selection while Neural Network in combination with Cox Proportionality Hazard model was used to build the survival prediction model.

II. RELATED WORKS

Rocky et al.(2005) proposed a system as an early warning for cervical cancer diagnosis. The proposed system used a hybridized ridge polynomial neural network and chaos optimization algorithm. Their system showed self learning ability due to the use of machine learning but devoid of malignant cancer recognition and survival analysis. [4] design a hybrid decision support system for detecting the different stages of cervical cancer. They used rough set theory, genetic algorithm and neural network which had a good performance however over fitting and under fitting was an imminent problems in addition to the fact that there was no survival analysis. [5] presented new feature of cervical cancer that is suitable and can be used as inputs for neural networks in cervical cell classification system using hierarchical hybrid multilayered perception.[7] presented a survey of soft computing to improve the accuracy of predicting cancer susceptibilities, recurrence and portability and implemented it using artificial neural network. Their system though was able to address implicit relationship, large data and non linear data yet was prone to overfitting and consequent performed poor. Jim et al. (2015) developed a cervical cancer progress prediction tool for human papillomavirus using a support vector machine but was computationally intensive[9] developed a system to classify cervical cancer using different types of artificial neural network architectures. Chari et al(2013) presented a magnetic resonance imaging (MRI) appearance of cervical carcinoma and cross sectional imaging and introduce subjectivity in diagnosis through picture recognition. [11] proposed a novel methodology for screening cervical cancer using artificial neural network. [12] present an improvement of Multilayer perception classification on cervical pap smear data with feature extraction using multilayered perceptron. [14] presented a cervical cancer detection and classification system using texture analysis as a proper classification technique to obtain the staging of cervical cancer patients using texture analysis of magnetic resonance imaging. Abdullah et al.(2017)[15] proposed an efficient model for feature selection and classification of cell in cervical smeared images using fuzzy K-nearest neigbours algorithm and Particle swarm algorithm. Mohammed et al.(2017) [16] proposed a system for determining high risk patient with cervical cancer through machine learning using multilayered perceptron, Bayesian network and k-nearest neigbour. Their proposed system yielded a high level of classification accuracy as well as reduction of overfitting through Bayesian network and k-nearest neigbour.[17] investigated the efficacy of using multilabel classification techniques for diagnosing cervical cancer at early stage. They used learning algorithms (multi-label classification) , Naïve bayes, J48 decision tree, sequel minimization optimization and random forest method[18] presented a model for the prediction and diagnosis of cervical cancer using Adaptive Neuro-Fuzzy Inference system.[19] developed an expert system for predicting the cervical cancer using data mining technique (Genetic Algorithm and Artificial Neural Network). [19] also presented an innovative system for classifying cervical cancer using Adaptive Neuro Fuzzy Inference System [20] presented a supervised deep learning embeddings for the prediction of cervical cancer by predicting the outcome of a patient biopsy given risk patterns from individual medical records.[21] designed a system to predict 10 year overall survival in patients with operable cervical cancer using probabilistic neural network model.[21][10] also presented a system to predict 5-year overall survival in cervical cancer patient treated with radical hysterectomy using computational intelligence of probabilistic neural network, multilayered perceptron, gene expression programming classification, k means algorithm and radial basic function based support vector machine. Their system utilized lesser time in determining the network architecture and in training. Sushruta et al(2017) proposed Genetic algorithm as an effective tool for global optimization. Mercy et al(2019)[24] developed an algorithm for automated detection of cervical pre-cancers with low-cost point of care, pocket colposcope which used automatic feature extraction and classification for VIA and VILI cervigram and combining features of VIA and VILI feature to train Support Vector Machine.

Sowjanya et. al(2019)[14] presented a machine aided identification of risk factors of cervical cancer among individuals who are likely to get the disease using feature selection methods and C45 classification algorithm. [22] developed a microarray based cancer prediction using soft computing approach of rough set theory which performed comparatively well. [23] proposed a data value metric for quantifying the value of data in a big data ecosystem and its suitability for feature selection even in traditional structured data. Analysis of Existing System Analysis is an important phase in system development life cycle where factual data are collected in view of understanding the processes involved, identifying problems and recommending solutions in order to improve the system functioning. It is an attempt to give birth to new ideas that satisfy the current needs of the user and provide a basis for future improvements.

Algorithm of the Existing System

The algorithm of the existing system is heavily reliant on the probabilistic neural network (PNN) and is shown in Fig. 1

Input: D: a training data set. T: a test data set. PNN: Probabilistic Neural Network. **Output:** L: label of classification results
Procedure: **Step 1**: Read in the file of exemplar vectors and class numbers **Step 2**: Sort these into k sets where each set contains one class of vectors **Step 3**: for each k Define a Gaussian function centered on each exemplar vector in set k Define the summed Gaussian output function Once the PNN is defined , then we can feed vectors into it and classify them as follows **Step(i)** read the input vector and feed it to each Gaussian function in each class **Step(ii)** for each group of hidden nodes, compute all Gaussian functional values at the hidden nodes **Step (iii):** for each group of hidden nodes, feed all its Gaussian functional values to the single output node for that group **Step (iv):** at each output node, sum all of the inputs and multiply by constant **Step(v):** find the maximum value of summed functional values at the output nodes
Prediction/ Classification L ′= predict Label(T); L= convertLabel(L ′)

Fig. 1: Algorithm of the Existing system using PNN

The algorithm of the existing system is heavily reliant on the probabilistic neural network (PNN) and is shown in Figure 1.

Weakness of the Existing System

The existing system of Bodgan et al. (2019) has some limitations which prevent it from having a practically good performance as needed. The limitations are as follows:

i. The model did not consider the use of feature selection as with the high number of variables or features could lead to degradation of performance and make the model more complex more so when PNN consumes a lot of memory, Their model was shown to prone to lengthy training time, over-fitting and under-fitting problem which might be a consequence of not using feature selection.
ii. The existing systems used Probabilistic Neural Network. PNN alone is very poor when it comes to temporal data for classification task; it is unable to perform well on time based aspect of the problem under review.

III. ARCHITECTURE OF THE PROPOSED SYSTEM

The architecture of the proposed system is shown in Fig. 2 while the feature selection sub component is expanded in Fig. 3

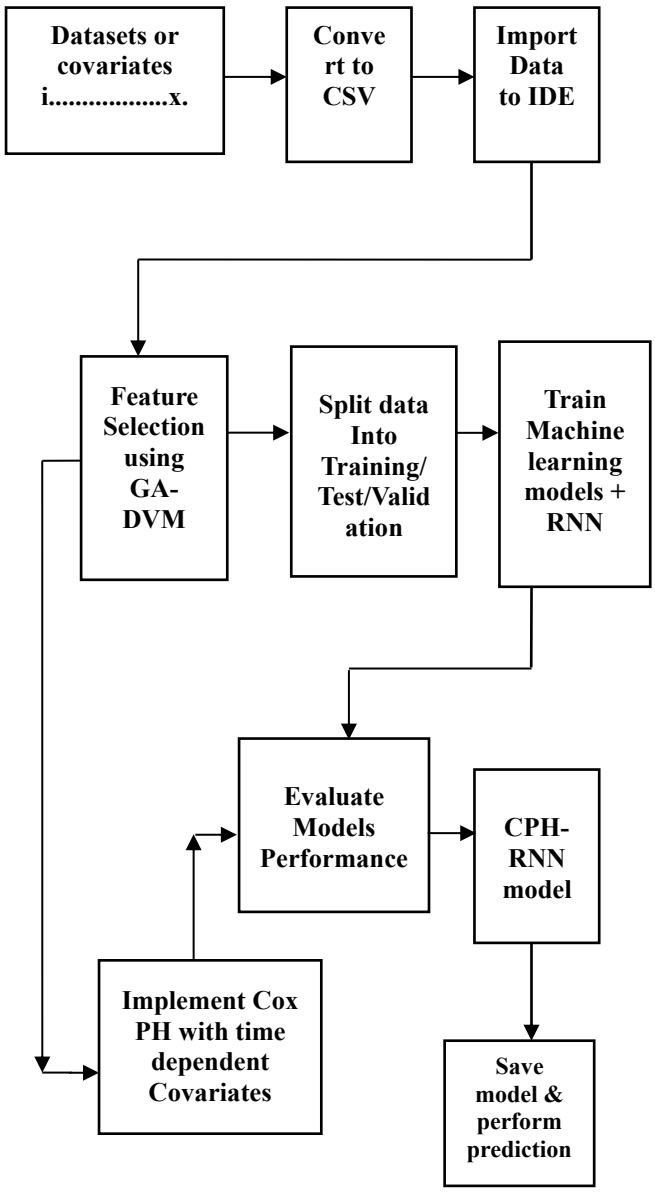

Figure 2: The Architecture of the proposed System

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

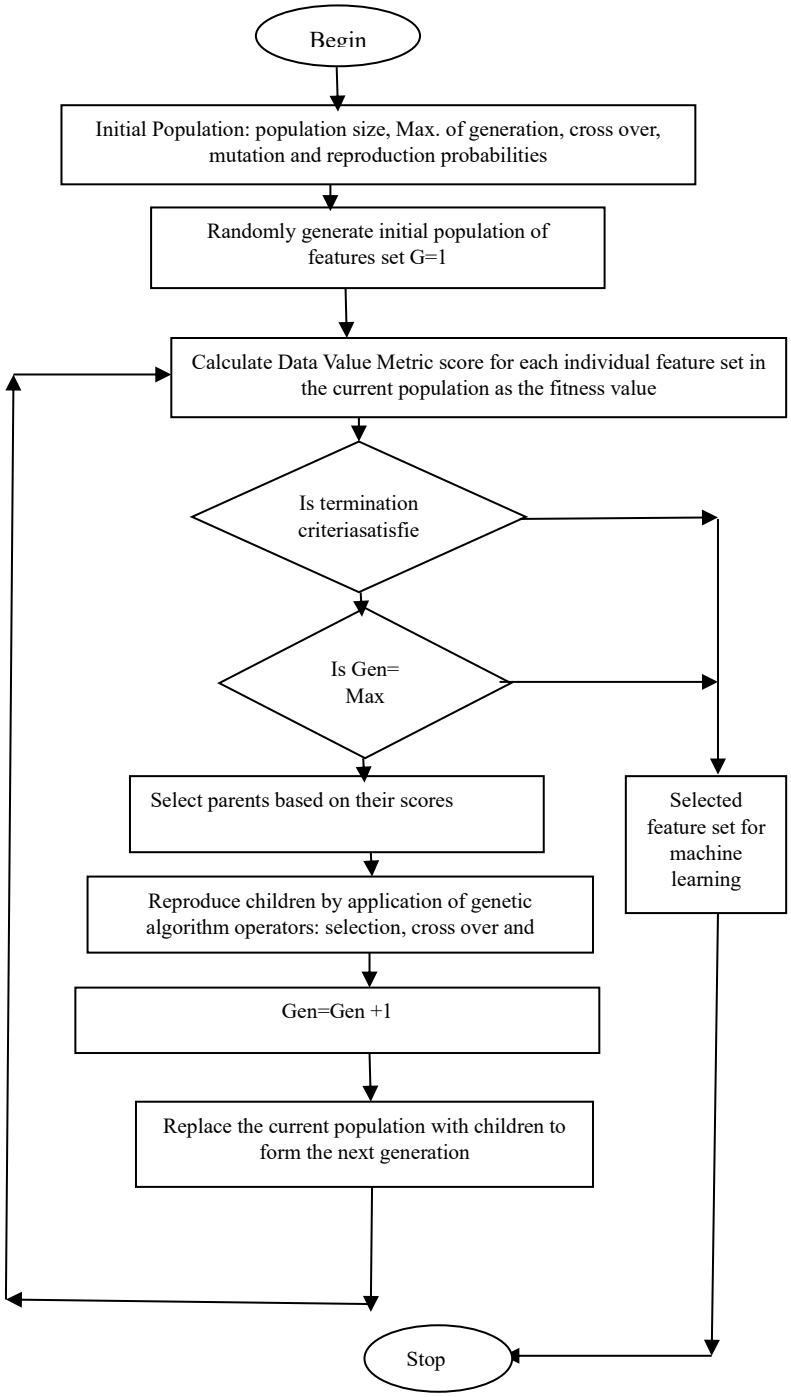

Figure 3: Proposed Genetic Algorithm based Data Value Metric feature selection

A. Brief Description of the Components of the Proposed System

This section gives a brief description of the various components in the proposed system architecture indicated in Fig. 2. and Fig. 3

1. Patient Symptoms and Signs Subcomponent

Disease symptoms are the biological indicators which are associated with the clinical presentation of diseases as learnt from medical literature and expert physicians. George *et al* (2000) opined that a symptom is a visible or even a measurable condition indicating the presence of a disease and thus can be regarded as an aid towards diagnosis. It is based on this clinical presentation that a doctor or physician makes a tentative judgment about the state of the patient of either being positive or negative to the disease and consequently a test for confirmation

2. Feature Selection Sub-Component

It is important to note that the essence of feature selection in this model is feature reduction. The feature selection process points out all the input features relevant for the survival prediction of the cervical cancer and it is an indispensable data pre-processing step. The difficulty of extracting the most relevant and informative variables is due mainly to the large dimension of the original feature set.

The feature selection is done using Genetic algorithm and data value metric where \data value metric is used as the fitness function in the genetic algorithm. The subcomponent of the feature selection is:

B. Justification for the hybridizing the two algorithms for feature selection (Genetic Algorithm and Data Value Metric GA-DVM)

Noshad et al., (2021), proposed the use of Data Value Metric for feature selection in both supervised and unsupervised machine tasks. These methods used a sequential forward selection search strategy, which is prone to local minima and has a negative impact on the machine algorithm's performance when outputted features are used. On one hand, there is a need to avoid the local minima that is inherent in the Data Value Metric algorithm for feature selection, while on the other hand, there is a need to address the problem of some redundant features inherent in Mutual information, as the features chosen may not be guaranteed to be non-redundant (Chandrashekar and Sahin, 2014). The Genetic algorithm is well-suited to dealing with the dual problem as genetic algorithm can be used to avoid local minima and also deal with redundant features.

C. Data Value Metric

Noshad et al., (2021) proposed a new information theoretical measure that quantifies the useful information content of large heterogeneous and traditional datasets. Data analytical value (utility) and model complexity are used by the DVM. It can be used to determine whether appending, expanding, or augmenting a dataset will benefit specific application domains. DVM quantifies the information boost or degradation associated with increasing the data size or the richness of its features, depending on the data analytic, inferential, or forecasting techniques used to interrogate the data. DVM is a combination of fidelity and regularization terms. The fidelity measures the utility of the sample data in the context of the inferential task. The computational complexity of the corresponding inferential method is represented by the regularization term. Inspired by the concept of information bottleneck in deep learning, the fidelity term depends on the performance of the corresponding supervised or unsupervised model. DVM captures effectively the balance between analytical-value and algorithmic-complexity. Changes to the DVM highlight the tradeoffs between algorithmic complexity and data analytical value in terms of sample size and dataset feature richness. DVM values can be used to optimize the relative utility of various supervised or unsupervised algorithms by determining the size and characteristics of the data.

D. Genetic Algorithm (GA)

A GA is a heuristic search algorithm that is based on natural selection and genetics. To evolve a solution to a problem, the idea is to mimic biological processes such as survival of the fittest. GA is a method of evolving chromosome populations to new populations by combining selection with operations such as crossover and mutation (Mitchel, 1998). Each chromosome contains genes. Selection operators select the fittest individuals from the population, whereas crossover and mutation mimic biological processes that introduce diversity into the population. Crossover and mutation are exploration processes, whereas selection is an exploitation process. Evolutionary algorithms are best suited for problems with a large search space, or a large number of possible solutions. Other problems necessitate the creation of new solutions at each stage in order to investigate new options, or they involve complex solutions that cannot be processed by hand (Mitchel, 1998). GAs, like the evolutionary process, relies on the fittest organisms/solutions to survive. The fitness of an organism/solution is determined by the problem at hand, and it is a factor that evolves over time.

E. Recurrent Neural Network (RNN)

This is a type of artificial neural network that is suitable for processing temporal information or data and learn sequences. It contains at least one feedback connection; therefore the activations can flow in a loop. The survival data used in this research is a temporal data hence the suitability of using RNN for it.

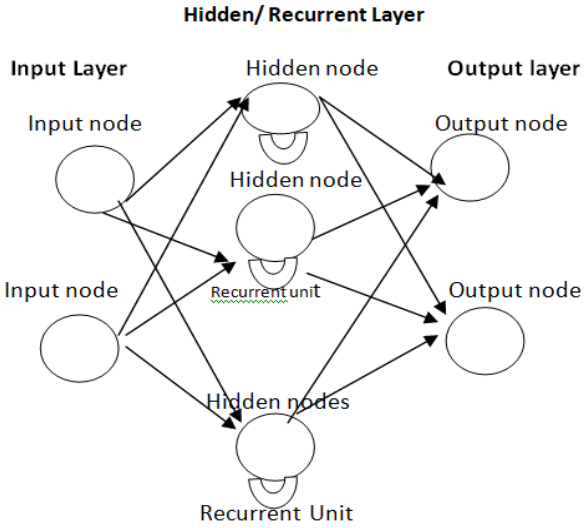

Figure 4: A simple architecture of RNN

IV. METHODS AND MATERIALS

To carry out survival analysis and develop a predictive model on cancer using a machine learning algorithm, the following steps were adopted (i) Data collection, (ii) Data preparation, (iii) Feature selection (iv) Implementing the proposed model, and (iv) Evaluation.

A. Data Collection

Collection of quality data is an indispensable aspect of machine learning. This research work made use of cervical cancer data collected from the University of Benin Teaching Hospital, Edo State, Nigeria which was well approved by the Ethical committee of the hospital. An Oncologist who is specialized in diagnosis and treatment of cervical cancer guided the collection of the cervical cancer data. Records of about 120 patients were collected using excel sheets to enter the data. About 13 records were dropped because of incomplete information leaving a total of 107 records. A follow up on those whose files were around and were not coming for treatment again was done through a well-established procedure to assay whether they are still alive or not. Cervical cancer Data was collected from 2012 till date which was subsequently used for the system dataset during the implementation.

The cervical cancer data were recorded using a spreadsheet with the assistance of the health workers in the unit. Most of the features relating to cervical cancer that needed to be collected for the study were outlined by medical personnel (Oncologists) that the researcher contacted Features relating to the survival or mortality of diabetes mellitus patients were collected from the Hospital. Table 1 below is a description of the variables collected and used for the proposed system.

Table 1. Identified Variables for determining Cervical Cancer Data

S/N	Names of Variables	Labels
1	Age at first diagnosis (in years)	Numeric
2	Present age	Numeric
3	Highest Education	Primary, Secondary, Tertiary, Others, Nil
4	Occupation	Business, Civil Servant, Teacher, Electrician, Trader, Carpenter, Farmer, Cleaner, Nil
5	Marital Status	Single, Married, Widow, Widower, Divorced
6	Ethnicity	Urhobo, Yoruba, Igbo, Hausa, Itsekiri, Ijaw
7	History of Smoking	Yes, No
8	Stage	Early stage cervical cancer, Locally Advance cervical cancer, Advance Stage cervical cancer
9	Stage _level	
10	Treatment options	Chemotherapy-1,Brachytherapy-2,Radiotherapy-3,Chemoradiation-4
		Numeric
11	Menopause	Adrenocasinoma- 1, Squamous cell 2, Adrenosquamous -3
12	Histology	Bleeding, Diabetes, Hypertension (Any 2 --- 2, Any 1 --- 1 , Non --- 0)
13	commobidity	
14	Mortality (Dead or Alive)	Numeric

B. Data Preparation

After the collection of data, the data was prepared and cleaned. This means that data necessary for use in the various machine learning algorithms were processed. The process involved the following:

1. The ethnicity (Tribe) column was dropped.
2. Yes / No type of columns were respectively converted and cleaned.
3. Single observation on target variable value was found to be missing and hence dropped.
4. Missing observations on a few categorical columns were detected and thus filled with their respective column mode value.
5. All the categorical features were one-hot encoded.
6. Treatment options were splitted into chemotherapy, radiotherapy, radiology, Brachytherapy, chemo radiation
7. Commodity column was splitted to CM1, CM2 and CM3
8. All the values in the dataset were scaled between 0 and 1 as a standardization technique.
9. Dataset was split into training and testing sets

C. Feature selection using Genetic Algorithm-Data Value Metric (GA-DVM)

Building a model requires feature selections to ensure right number of features as well as non-redundant and relevant features is selected which are representative of the domain under discussion (in this case, cervical cancer patients). The dataset of the cervical cancer of 107 patients and 15 attributes or features shown in Fig. 5 are fed into the feature selection using GA-DVM. The histogram of the age_at_diagnosisis shown in Fig. 6. The pair plot of some selected features is shown in Fig. 7. The output of the model showing the selected features and a plot of the fitness value (Data Value Metric) plotted against the generation is shown in Fig. 8 and Fig. 9 respectively.

	years_after_diagnosis	age_at_diagnosis	stage level	chemotherapy	brachtherapy	chemoradiation	radiotherapy	radiation	menopause	MENO_post
0	5.0	54	6	1	0	0.0	0	0	0	0
1	8.0	41	6	0	0	1.0	0	0	0	0
2	7.0	65	5	0	1	0.0	0	0	1	7
3	7.0	85	5	1	0	0.0	0	0	1	14
4	8.0	60	10	0	0	1.0	0	0	1	7
...	...	...	...	...	...	...	...	...	...	...
103	1.0	49	6	0	0	0.0	1	0	0	0
104	8.0	58	4	1	0	0.0	0	0	1	21
105	1.0	47	5	0	1	0.0	0	0	1	12
106	2.0	73	8	0	0	1.0	0	0	1	27
107	6.0	63	8	0	0	0.0	1	0	1	9

108 rows × 15 columns

Fig. 5: snapshot of the cervical cancer dataset

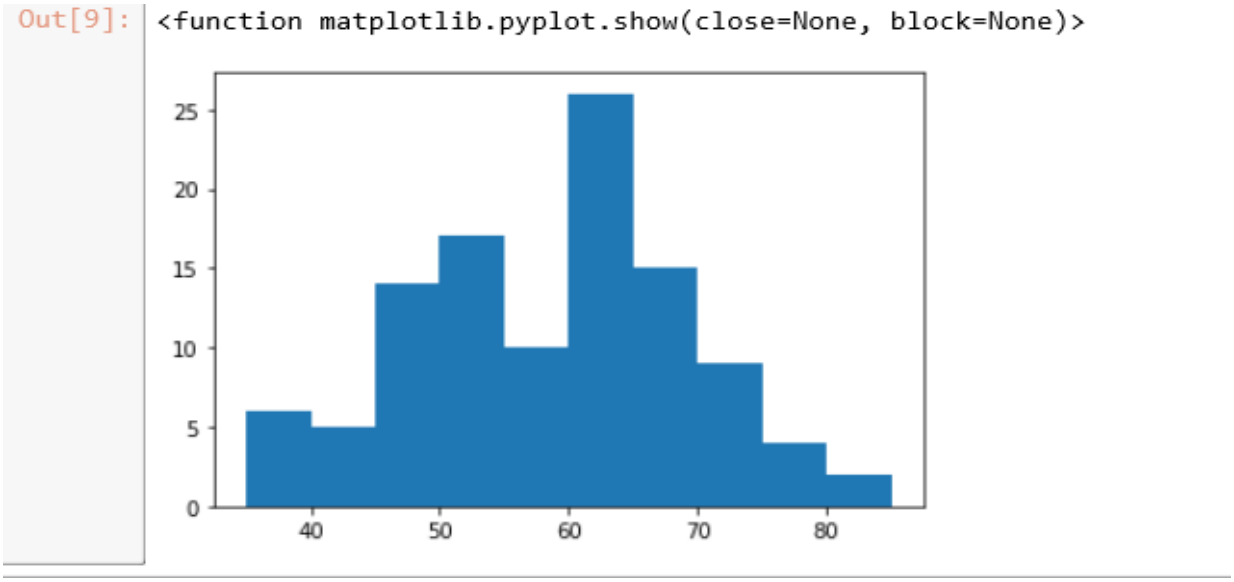

Fig. 6: The histogram showing the age_at_diagnosis

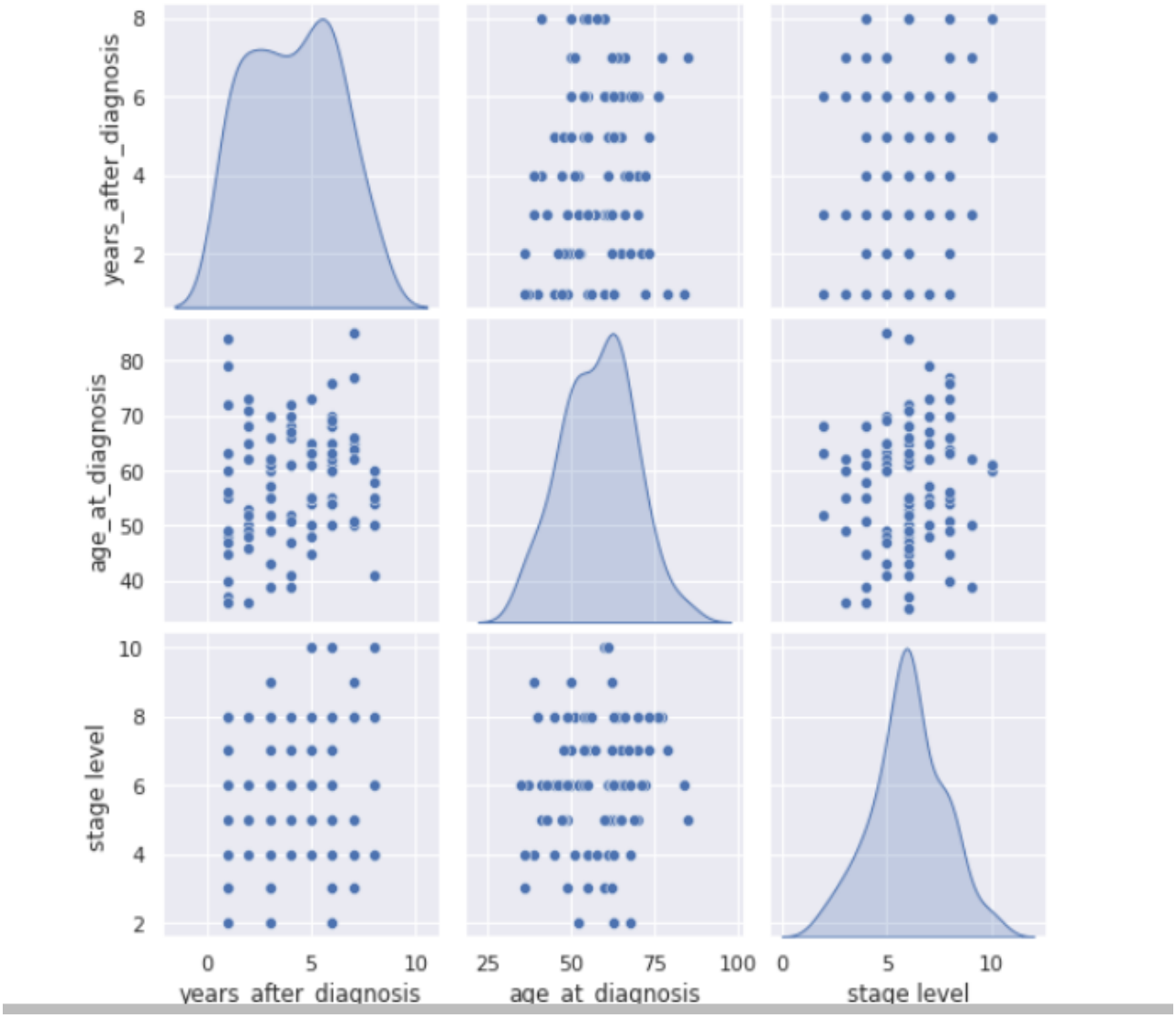

Fig. 7: A pair plot of stage level,age_at_diagnosis, years_after_diagnosis

Retrieval Number: 100.1/ijsce.B36080513223
DOI: 10.35940/ijsce.B3608.0513223
Journal Website: www.ijsce.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

gen	nevals	fitness	fitness_std	fitness_max	fitness_min
0	30	0.81713	0.0080853	0.833333	0.805556
1	60	0.822685	0.00587434	0.833333	0.819444
2	60	0.823148	0.0061419	0.833333	0.819444
3	60	0.824074	0.00654729	0.833333	0.819444
4	60	0.825	0.00680414	0.833333	0.819444
5	60	0.827315	0.00688244	0.833333	0.819444
6	60	0.826852	0.006929	0.833333	0.819444
7	60	0.827778	0.00680414	0.833333	0.819444
8	60	0.826389	0.00694444	0.833333	0.819444
9	60	0.829167	0.00636469	0.833333	0.819444
10	60	0.830556	0.00555556	0.833333	0.819444
11	60	0.82963	0.0061419	0.833333	0.819444
12	60	0.830556	0.00555556	0.833333	0.819444
13	60	0.829167	0.00636469	0.833333	0.819444
14	60	0.828704	0.00654729	0.833333	0.819444
15	60	0.826389	0.00694444	0.833333	0.819444
16	60	0.827315	0.00776067	0.833333	0.805556
17	60	0.827778	0.00680414	0.833333	0.819444
18	60	0.824074	0.00654729	0.833333	0.819444
19	60	0.827778	0.00680414	0.833333	0.819444
20	60	0.826852	0.006929	0.833333	0.819444

Fig. 8: Output of the feature selection using GA-DVM

[False,True,False,True,False,True,False,False,True,True,True,True,False,False]

[age_at_diagnosis, Chemotherapy, Chemoradiation, Menopause, MENO_Post, Histology, CM_2]

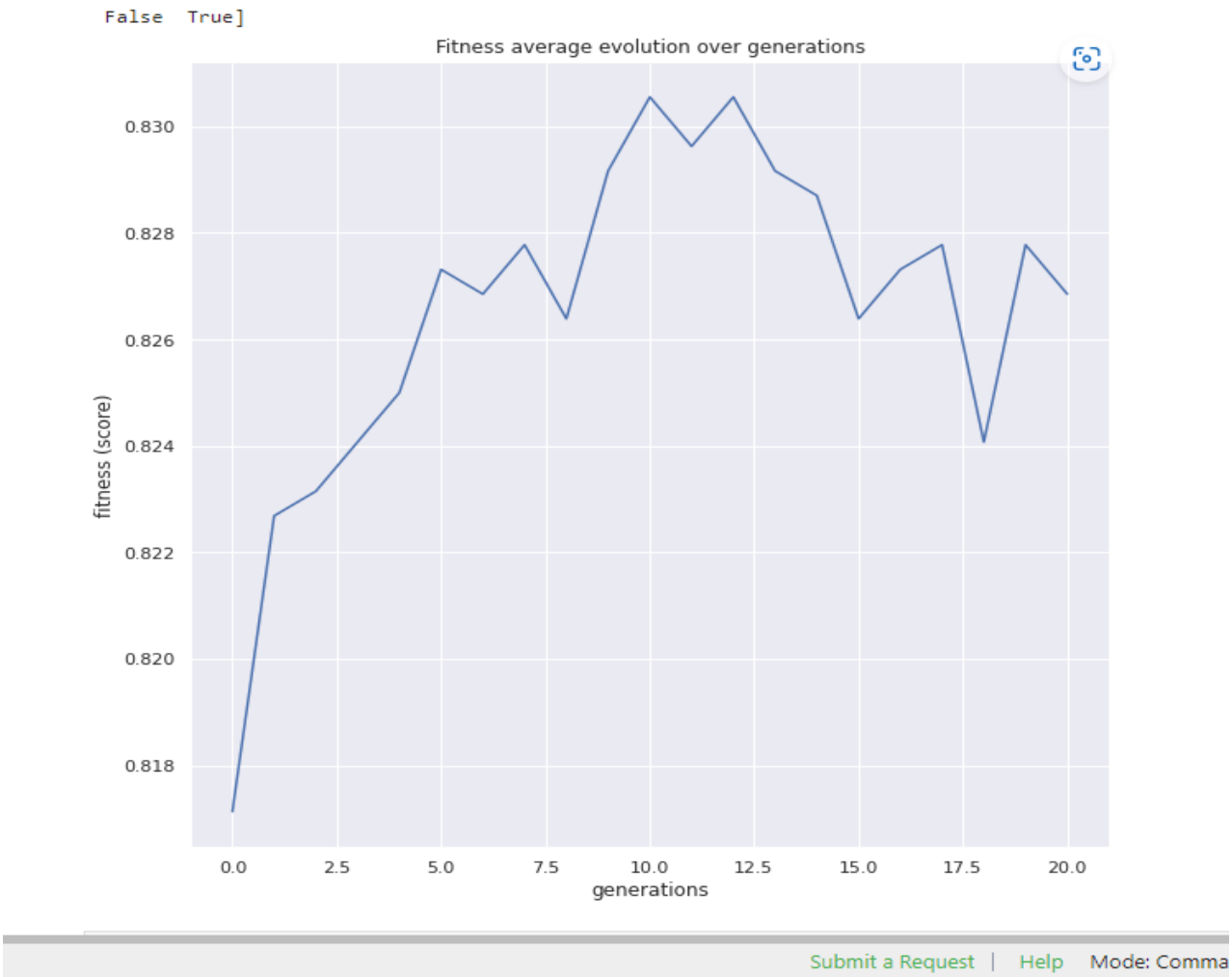

Fig. 9: Genetic Algorithm Fitness Score (DVM Score) Versus Generations

The GA-DVM selected some features such as: age_at_diagnosis, Chemotherapy, Chemoradiation, MENO_post, Histology,CM_2

Implementing the proposed model

D. Data Training and Learning using Recurrent Neural Network

The cervical cancer datasets are prepared for training and learning of the dataset features using Recurrent Neural Network. During the training process, the model makes an effort to comprehend the properties and instance representation of the cervical cancer patients' dataset that is used as input.The programmer chose to divide the datasets into three (3) partsof training, validation and testing respectively. The training data is represented using the attributes after the instance and attributes have been chosen.
In Recurrent Neural Networks (RNN), outputs from the preceding states are fed as input to the current state. The hidden layers in RNN can remember information. The hidden state is updated based on the output generated in the previous state. However, in this present study, RNN going through 200 epochs. Fig. 10 shows that the training accuracy of the proposed model is 88.53%andthe validation accuracy is 73.85%. This clearly shows an acceptable generalization capability. The training epoch output showing the training performance in terms of validation loss, training loss, training accuracies and validation accuracies through the iterations is shown in Fig. 11

```
The accuracy of the training model:
0.8852777850627899
The accuracy of the validation model:
0.7383333298563958
```

Fig. 10: output of the model in terms of accuracies

```
3/3 [==============================] - 0s 14ms/step - loss: 0.3541 - accuracy: 0.8750 - val_loss: 0.5392 - val_accuracy: 0.7500
Epoch 186/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3514 - accuracy: 0.8750 - val_loss: 0.5258 - val_accuracy: 0.7500
Epoch 187/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3512 - accuracy: 0.8750 - val_loss: 0.5120 - val_accuracy: 0.7778
Epoch 188/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3515 - accuracy: 0.8750 - val_loss: 0.5181 - val_accuracy: 0.7500
Epoch 189/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3494 - accuracy: 0.8750 - val_loss: 0.5369 - val_accuracy: 0.7500
Epoch 190/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3497 - accuracy: 0.8750 - val_loss: 0.5577 - val_accuracy: 0.7500
Epoch 191/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3521 - accuracy: 0.8889 - val_loss: 0.5641 - val_accuracy: 0.7500
Epoch 192/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3516 - accuracy: 0.9028 - val_loss: 0.5489 - val_accuracy: 0.7500
Epoch 193/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3484 - accuracy: 0.8750 - val_loss: 0.5297 - val_accuracy: 0.7500
Epoch 194/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3488 - accuracy: 0.8750 - val_loss: 0.5152 - val_accuracy: 0.7500
Epoch 195/200
3/3 [==============================] - 0s 14ms/step - loss: 0.3485 - accuracy: 0.8750 - val_loss: 0.5194 - val_accuracy: 0.7500
Epoch 196/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3470 - accuracy: 0.8750 - val_loss: 0.5336 - val_accuracy: 0.7500
Epoch 197/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3470 - accuracy: 0.8750 - val_loss: 0.5410 - val_accuracy: 0.7500
Epoch 198/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3472 - accuracy: 0.8750 - val_loss: 0.5476 - val_accuracy: 0.7500
Epoch 199/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3468 - accuracy: 0.8889 - val_loss: 0.5619 - val_accuracy: 0.7500
Epoch 200/200
3/3 [==============================] - 0s 13ms/step - loss: 0.3484 - accuracy: 0.9028 - val_loss: 0.5672 - val_accuracy: 0.7500
<keras.callbacks.History object at 0x7fab824cf250>
```

Fig. 11: Output of the training in terms of epoch

E. Learning Curve of the Proposed Model

A learning curve is a plot of model learning performance over experience or time. The learning curve of model performance on the train and validation datasets can be used to diagnose an underfit, overfit or w well-fit as well as used to diagnose the representative of the training and validation datasets of the problem domain. There are two types to be considered here, Optimization learning curve and performance learning curve.

F. Optimization Learning Curve

The optimization learning curve is calculated based on the metric by which the parameters of the model are being optimized, in this case, the loss. The optimization learning curve gives an idea of how well the model is learning. The optimization learning curve is shown in Fig. 12.

G. Performance Learning Curve

This is a learning curve calculated based on the metric by which the model will be evaluated and select. In this case, we used accuracy. The performance learning curve is shown in Fig. 9.

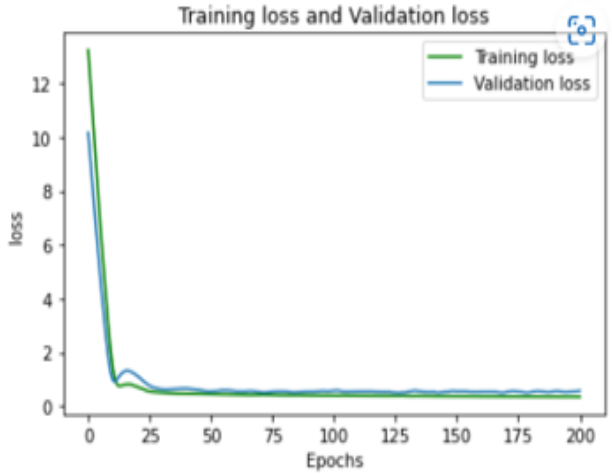

Fig. 12: Optimizing Learning curve (Training Loss versus Validation Loss)

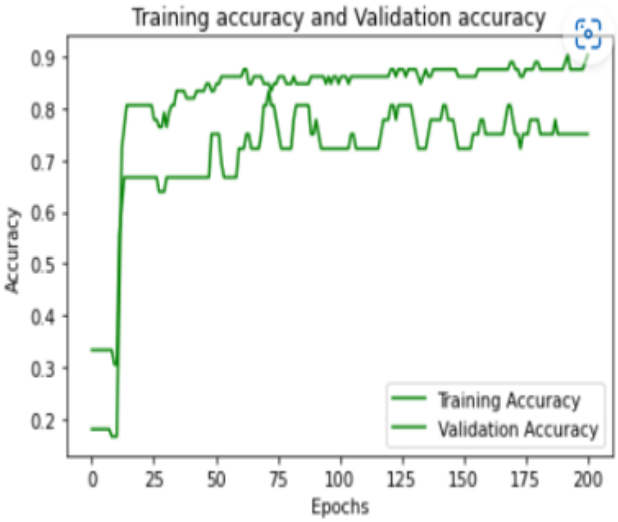

Fig. 13: Performance Learning Curve for the training

From Fig.12, it could be seen that the proposed model had a good fit as the plot of training loss and validation loss decreased to a point of stability and has a small gap between them. It could be seen that both the validation and training datasets are adequately representative and provide enough information to learn the problem and evaluate the ability of the model to generalize. The proposed model as shown in Fig. 7 is representative because the validation loss is always higher than the training loss and the gap between them is small.

V. EVALUATION

The performance metrics were used to evaluate the performance of the various algorithms in this study. This section compares the output of the training datasets with that of the testing datasets in order to check all conceivable combinations and evaluate how effectively a model will predict the intended or expected results. If the expected result is far different from the output result, input can be adjusted and the model will be fine-tuned based on the results of the test data set. This is accomplished by comparing the attributes of the training and testing datasets, computing the probability for each hypothesis based on the attributes, and categorizing the attributes that are most similar to the outcome.

A. Result from Existing of Bodgan et al (2019) using PNN Algorithm

Table 2: Performance measure of our existing system using PNN (Bodgan et al, 2019)

Metrics	std=0.2	std=0.4	std=0.6	std=0.8	std=1.0
F-measure	0.75	0.8	0.8	0.8	0.78
Accuracy	0.64	0.69	0.7	0.7	0.67
ROC score	0.5625	0.5	0.6041	0.6041	0.5625
Precision	0.7	0.71	0.72	0.72	0.7
Recall	0.79	0.92	0.96	0.96	0.88

A Probabilistic Neural network (PNN) algorithm used by our existing system of Bodgan et al. (2019) was implemented using the same cervical data used in the proposed system to compare their performance. Its output in terms of Precision, Recall, Accuracy, F-measure and ROC AUC is shown in Table 2 using standard deviation of 0.2, 0.4, 0.6, 0.8 and 1.0 respectively while the performance measures are graphically displayed as shown in Fig.14. The highest Area under Curve of the standard deviation configurations, its classification report and confusion matrix is shown in Fig. 15, Fig.16 and Fig. 17 respectively.

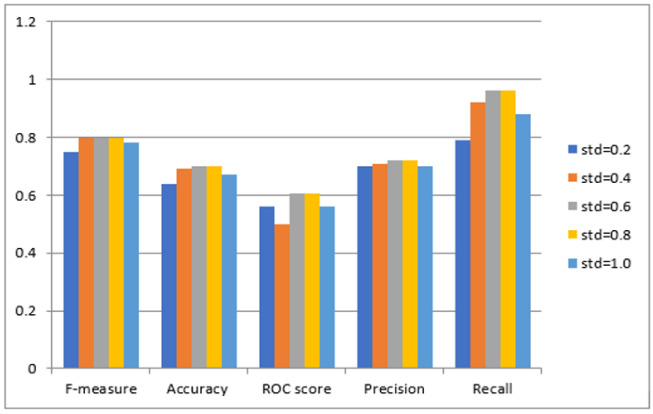

Fig. 14: performance measures on Cervical dataset using existing system PNN (Bodgan et al, 2019)

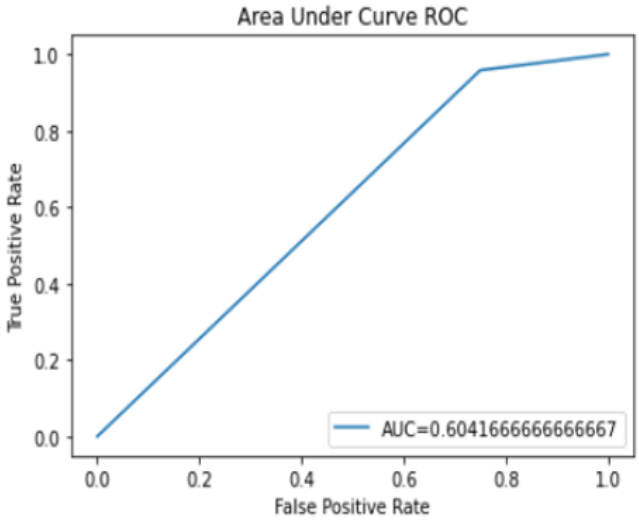

Fig. 15: ROC curve for existing system PNN (Bodganet al., 2019)

Retrieval Number: 100.1/ijsce.B36080513223
DOI: 10.35940/ijsce.B3608.0513223
Journal Website: www.ijsce.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Confusion Matrix
[[4 8]
 [5 19]]

	precision	recall	f1-score	support
0	0.44	0.33	0.38	12
1	0.70	0.79	0.75	24
accuracy			0.64	36
macro avg	0.57	0.56	0.56	36
weighted avg	0.62	0.64	0.62	36

Fig. 16: Classification report of the existing system

B. Result from our Proposed Model

A recurrent neural network algorithm used by our proposed system was implemented using the same cervical dataset and was tested. Two variants of the model was built, our proposed system hybridizing the feature selection algorithm of GA-DVM while the other one is without the feature selection component and as such feature selection was not done.30% of the dataset was kept as a testing datasets in order to assay the performance of the model. The classification report of the proposed system is shown in Fig.18; the ROC-Curve is shown in Fig.19. A comparative analysis of the proposed system (RNN+GA-DVM) and the other system (RNN without feature selection) in terms of Precision, Recall, Accuracy, F-measure and ROC AUC is shown in Table 3 while the comparative performance measures are graphically displayed shown in Fig. 15 and Fig.16.

Confusion Matrix
[[3 9]
 [1 23]]

	precision	recall	f1-score	support
0	0.75	0.25	0.38	12
1	0.72	0.96	0.82	24
accuracy			0.72	36
macro avg	0.73	0.60	0.60	36
weighted avg	0.73	0.72	0.67	36

Fig. 17: Classification report of the proposed system

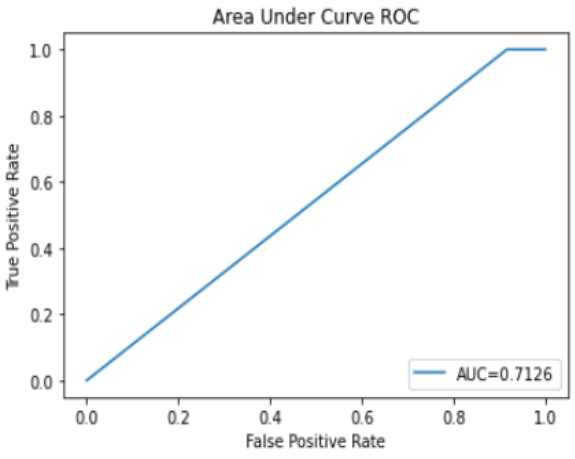

Fig. 18: ROC Curve of the proposed system

Table 3: Performance comparison of proposed system (RNN+GA-DVM) and RNN without GA_DVM

Metrics	RNN+GA-DVM (Proposed system)	RNN without GA_DVM
F-measure	0.82	0.781
Accuracy	0.7516	0.7
ROC score	0.7126	0.5948
Precision	0.72	0.7
Recall	0.96	0.93

Table 4: Performance comparison of proposed system (RNN+GA-DVM), RNN without GA_DVM and Existing System (PNN model)

Metrics	RNN+GA-DVM (Proposed system)	RNN without GA_DVM	Existing Model (PNN using std=0.8)
F-measure	0.8200	0.7810	0.8000
Accuracy	0.7516	0.7000	0.7000
ROC score	0.7126	0.5948	0.6041
Precision	0.7200	0.7000	0.7200
Recall	0.9600	0.9300	0.9600

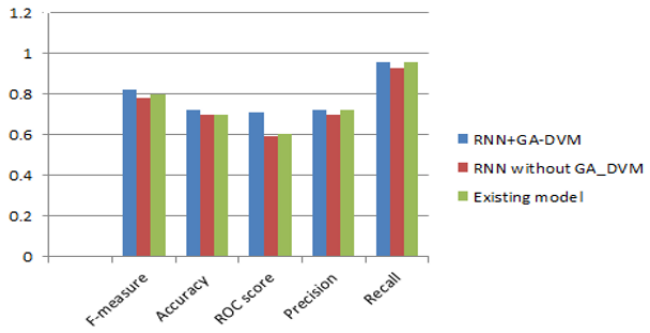

Fig. 19: Performance comparison of proposed system (RNN+GA-DVM) and RNN without GA_DVM

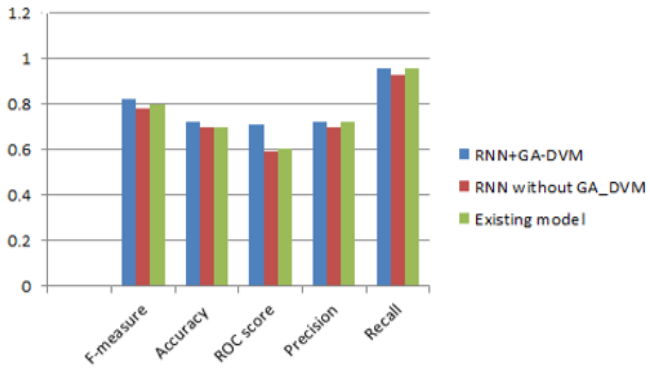

Fig. 20: Performance comparison of proposed system (RNN+GA-DVM) and RNN without GA_DVM

VI. RESULT DISCUSSION

To perform the prediction of cervical cancer survival, this study has developed a hybrid model of implementing Genetic Algorithm based Data Value Metric for the feature selection and the Recurrent Neural Network model for prediction. From the performance and evaluation of the designed model, this study has shown that the integration of feature selection into the recurrent neural network in survival prediction of cervical cancer prediction is achievable.

Retrieval Number: 100.1/ijsce.B36080513223
DOI: 10.35940/ijsce.B3608.0513223
Journal Website: www.ijsce.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The result of the Genetic Algorithm based Data value Metric for feature selection on the dataset showed that the variables that are associated with cervical cancer mortality are age at diagnosis, Chemotherapy, Chemo-radiation, Histology, Menopause, Comorbidity, and MENO_Post. Thus, with early diagnosis and proper health management of cervical cancer, the age of survival of cervical cancer patients can be prolonged. The study ascertained that our proposed system (RNN+GA-DVM) performs better than the existing system (PNN, presented by Bodgan et al, 2019) with an accuracy of 75.16% and 70.00%.

VII. CONCLUSION

The study did a survival analysis of cervical cancer patients using cervical cancer dataset from University of Benin Teaching Hospital, Benin, Delta State to train a recurrent neural network armed with feature selection capability of genetic algorithm and Data Value Metric algorithm for the survival prediction of cervical cancer patients. The genetic Algorithm-Data Value Metric selected seven (7) features for prediction of the survivability of any patient and was used to train the recurrent neural network which had an accuracy of 75.16% which is better performance than the existing system of Bodgan et al.(2019).The study has been able to establish that a genetic algorithm based data value metric can be used for feature selection which could improve the performance metrics of machine learning model for survival prediction of cancer patients can be implemented. This is because machine learning tools and algorithms and efficient in building prediction and analysis models.

DECLARATION

Funding/Grants/ Financial Support	No, I did not receive.
Conflict of Interest/ Competing Interests	No conflict of interest to the best of our knowledge
Ethical approval and Consent to participate	Yes, University of Benin Teaching Hospital Benin City, Edo State.
Availability of data and material/ Data Access Statement	Yes, the dataset was collected from University of Benin Teaching Hospital After careful approval from the ethical committee of the Teaching hospital.
Authors Contribution	First three authors have equal participation in this article and fourth author provided the need expertise in gathering the data as well as during the interpretation of results.

REFERENCES

1. I.A Oludare., Aman J., Abiodun E. O., Kemi V. D., Nachaat A. M. and Humaira A. (2018), State-ofthe- art in artificial neural Network Application: A Survey, Heliyon, Vol. 4, Iss. 11 [CrossRef]
2. R.Sathya. and A. Annamma. (2013), Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification, (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 2, No. 2, pp. 34 – 38 [CrossRef]
3. Y.D, Rocky., C.M Gloria, & L.L Jans (2005) . Early Warning system for cervical cancer Diagnosis using Ridge Polynomial neural network and chaos optimization M. Journal of theoretical and Applied information technology 96(7) April.
4. M. Pabitra, M.Sushmita, & K.P Sankar, (2000).Staging of cervices cancer with soft computing. IEEE Transactions on biological Engineerng. 47(7), July. [CrossRef]
5. M. Mazahah, M. NorAshidi, Y.M Mohd, (2008). Capability of new features of cervical cells for cervical cancer diagnostic system using hierarchical neural network. /JSSST, 2(2) May.
6. B.Ighoyata, ,P. Suyatha & I.A, Maureen (2017); Fuzzy based multi-fever symptom classifier Diagnosis model, International journal of Information and computer science/MECS Press, vol 9 (10 pages; 13 – [CrossRef]
7. C.Tanupriya, , K. Vwek, , & N. Darshika (2013). Cancer research through the help of soft computing techniques; A survey international journal of computer science and mobile computing (IJCSCME) 2 April PP. 467-477.
8. A.E Okpako & P.O Asagba (2017) An Improved Frame work for Diagnosing confusability Disease using Neutrosophic Based Neural Newwork, Neutrosophic sets and systems. Vol 16 /iss 1/7
9. M. Anousouya, S. Devi, J. Ravi, ,V. Vaidhnav & S. Putnitha,. (2018). Classification of cervical cancer using artificial neural networks retrieved at www.sciencedirect .com on 10/06/2019.
10. B. Charis , C Konstantinos , & A.M Lia (2016) Current Imaging Strategies for the Evaluation of Uterine Cervical Cancer. https://www.researchgate:net/publication/ 301483 937, Retrieved 14/5/2019.
11. P. Elayaraja, & M Sugarithi, (2018). Automatic Approach foprCervical cancer detection and segmentation using neutral network classify. Asian pacific journal of cancer prevention (19)12 pp 3571-3580. [CrossRef]
12. K. Hernalatha, & K, Usah Rani, (2016). Improvement of Multi layer perception classification on cervical pop smear data with feature extraction. International Journal of Innovative Research in Science, engineering and technology. 5(12).December.
13. M.K Soumya, K, Sneha., & C ,Arunvinodh . (2016) cervical cancer Deletion and classification using texture Analysis Biomed Pharmacol J. 9(2) http://biomedpharma journal.org/?p=7750. [CrossRef]
14. M.I Abdullah ., & F. Chastine . (2017) A Quantum Hybrid PSO Combined with fuzzy K. NN Approach to feature Selection and Cell classification in Cervical cancer detection. F.U Muhammed, .(2017). Determining cervical cancer possibility by using machine learning methods.International journal of latest research in Engineering and technology 3(12) December.www.ijret.compp 65-71.
15. C.,Zeynep., & P. Ebru (2017) Comparison of Multi-Label classification methods for Prediagnosis of cervical cancer international journal of intelligent system and Applications in Engineering IJISAE 5(4), 232-236. [CrossRef]
16. A.O, Akinrotim. & Olugbebi, MuyiawaAseolu (2018). Modeling and Diagnosis of Cervical cancer using adaptive neuron fuzzy inference system.World journal of research and review. 6(5) May pp1-3
17. B. Bargana, & A. Nagarajan, . (2018).An expert system for predicting the cervical cancer using data mining techniques. International Journal of pure and applied Mathematics. http://www.ijpam. edu. Retrieved 28/06/2019.
18. F. Kelvin , C. David, , S.Jaime . & F.Jessica . (2018). Supervised deep learning embeddings for the prediction of Cervical cancer diagnosis:Peer computer science 4:e154. https://doic/org/10.771/peer j-cs. 154
19. O. Bogdan , K.Maciej ., S. Andrzej , O. Marzanna , & K.Jacek K (2019). Prediction of 10 years overall survival in patients with Operable cervical cancer using a probabilistic Neural Network: J. cancer 10(18): 4189-4195 doi: 10. 7150/Jca.33945 Retrieved 10/6/19. [CrossRef]
20. W. Xiaosheng , & Osamu G., (2016) Microarray-Based Cancer Prediction using soft computing Approach.
21. M. Noshad., Y. Choi., Sun, , D hero, A & Ivo, D. (2021) A data value metric for quantifying information. Journal of big data [CrossRef]
22. N.A Mercy , S. Anish , C, Usamah , L.M Jenna ., Christopher T.L., John W.S., Gino V. Guillermo S. &Nimmi R (2018). Algorithms for automated detection of cervical pre-cancers with a low –cost, point-of-care, pocket colpo scope http://dx.doi.org/10.1101/32541 Retrieved 10/4/19.

23. D.V Ojie , M. Akazue , A. Imianvan . A Framework for Feature Selection using Data Value Metric and Genetic Algorithm, International Journal of Computer Applications, 2023 Volume 184, Issue 43, p14-21, doi:10.5120/ijca2023922533 [CrossRef]
24. B. Ojeme, M Akazue, E Nwelih. Automatic Diagnosis of Depressive Disorders using Ensemble Techniques. African Journal of Computing & ICT, 2016

AUTHORS PROFILE

Ojie Deborah Voke She is currently a lecturer in the department of Software Engineering, University of Delta, Agbor, Delta State. She had her first degree in Computer Science from Ambrose Alli University in 2001, PGDE in Delta State University Abraka, 2007, M.Sc Computer Science from Benson Idahosa University Benin City, Edo State 2011. She is a registered member of Computer Professional Registration council of Nigeria 2015, and currently on her last lap of her PhD programme in computer science Delta State University, Abraka. She has published so many journals both national and international. Her area of interest is software engineering, machine learning and IOT.

Dr. Akazue Maureen Ifeanyi she is a Lecturer in the Computer Science Department, Delta State University, Abraka, Delta State, Nigeria. She received a Master of Information Science degree in 2001 from the University of Ibadan, Oyo State, Nigeria, M.Sc. Computer Science in 2008 and Ph.D. Computer Science in 2014, both from the University of Benin, Edo State, Nigeria. She has a journey of almost 17 years in academics and is consistently striving to create a challenging and engaging learning environment where students become life-long scholars and learners. She is a dedicated teacher, researcher, and mentor that imparts lectures using different teaching strategies. She has several publications in reputed national and international journals accompanied by participation in conferences. Her research interests are HCI, Online fraud prevention Modeling, IoT, Trust model, and E-commerce. She is a member of the Nigerian Computer Society and Computer Professionals of Nigeria profile which contains their education details, their publications, research work, membership, achievements, with photo that will be maximum 200-400 words.

Dr. OMEDE Edith Ugochi Mary is a lecturer in the Computer Science Department, Delta State University, Abraka, Delta State, Nigeria. She obtained B.Sc. in Computer Science in 1997 from Enugu State University of Science and Technology, Enugu, Nigeria, M.Sc. Computer Science in 2005 from Nnamdi Azikiwe University, Anambra State, Nigeria and PhD in 2016 from the University of Benin, Edo State, Nigeria. She started academics race as lecturer in 2020, though has been in the University system for almost 20 years, 17 years of which she worked System Analyst/Programmer where she rose to position of Principal System Analyst/Programmer before passion for research and knowledge sharing drove her to academics. She is a dedicated teacher, researcher, and mentor that imparts lectures using varied pedagogical strategies. She has several publications in reputed national and international journals accompanied by participation in conferences. Her research interests are Software Engineering, Artificial Intelligent and IoT

Dr. Oboh E.O, He is currently the Head of Department of Radiotherapy/ Clinical Oncology, University of Benin Teaching Hospital, Benin City, Edo State, Nigeria. A Member, medical and DENTAL Consultant Association of Nigeria, Fellow, West African College of Surgeons (FWACS) member Nigeria Medical Association (NMA), Association of Radiologist of West Africa (ARAWA), Member, American Society of Clinical Oncologist (ASCO), Secretary Non – Communicable diseases Committee of the World Medical Association.

Prof. Imianvan Anthony Agboizebeta is presently Professor of Computer Science at the University of Benin, Benin City, Nigeria. His research interest includes Artificial Intelligence, Knowledge Engineering, and BioMedical Computing. Professor Imianvan Anthony Agboizebeta obtained PhD. (Computer Science) from Federal University of Technology Akure in 2009. He had previously received M.Sc and B.Sc in Computer Science from University of Benin, Benin City, Nigeria in 1998 and 1992/93 respectively.