

# Improved Classification Accuracy for Identification of Cervical Cancer

D. Merlin<sup>1</sup>, Dr. J. G. R. Sathiseelan<sup>2</sup>

<sup>1</sup>Department of Computer Science, Bishop Heber College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, India

<sup>2</sup>Head, Department of Computer Science, Bishop Heber College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, India

## ABSTRACT

### Article Info

Volume 7, Issue 6

Page Number: 245-258

### Publication Issue :

November-December-2021

### Article History

Accepted : 02 Dec 2021

Published : 10 Dec 2021

The major purpose of this research is to forecast cervical cancer, compare which algorithms perform well, and then choose the best algorithm to predict cervical cancer at an early stage. Cervical cancer classification can be automated using a machine learning system. This study evaluates multiple machine learning techniques for cervical cancer classification. For this classification, algorithms such as Decision Tree, Naive Bayes, KNN, SVM, and MLP are proposed and evaluated. The cervical cancer Dataset, which was retrieved from the UCI machine learning data repository, was used to test these methods. With the help of Sciklit-learn, the algorithms' results were compared in terms of Accuracy, Sensitivity, and Specificity. Sciklit-learn is a Python-based machine learning package that is available for free. Finally, the best model for predicting cervical cancer is developed.

**Keywords :** Decision Tree, Naive Bayes, KNN, SVM, MLP

## I. INTRODUCTION

Cancer is defined as the uncontrolled proliferation of body cells that can occur anyplace in the human body. Cervical cancer, often known as cancer cervix, is cancer that occurs in the female cervix. The cervix is the bottom section of the uterus that connects the uterus's body to the vaginal canal (birth canal). The cervix's lower third (ectocervix) is located within the vagina, while the upper two-thirds (endocervix) are located above the vagina. The area where the

endocervix and ectocervix meet is where the majority of cervical cancers begin. The major purpose of this research is to forecast cervical cancer, compare which algorithms perform well, and then choose the best algorithm to predict cervical cancer at an early stage.

### Cervical Cancer

Cervical cancer begins in the cells that line the cervix, which is the lower portion of the uterus (womb). The cervix connects the top section of the uterus (where a foetus grows) to the vaginal canal (birth canal). When

our body cells grow out of control cancer develops. The cervix is divided into two sections and is covered by two types of cells. The endocervix is the entrance to the uterus from the cervix. It has glandular cells covering it. During a speculum exam, the doctor can see the exocervix (or ectocervix), which is the exterior section of the cervix. Squamous cells cover the surface. The transformation zone is the area of the cervix where these two cell types meet. As you get older and if you have children, the exact location of the transformation zone shifts. The transformation zone is where the majority of cervical malignancies begin.

### Types of cervical cancer

The type of cervical cancer you've got influences your prognosis and treatment options. The following are the most common kinds of cervical cancer:

**Squamous cell carcinoma:** Cervical cancer of this sort starts within the skinny, flat cells (squamous cells) that line the outer section of the cervix, which protrudes into the vaginal canal. Squamous cell carcinomas account for the bulk of cervical malignancies.

**Adenocarcinoma:** The column-shaped glandular cells that line the cervical canal are where this type of cervical cancer originates. Cervical cancer can involve both sorts of cells sometimes. Cancer within the cervix's other cells occurs very rarely.

#### Symptoms:

- Vaginal bleeding during or after a sexual encounter, during periods, or after menopause.
- Vaginal discharge that is watery, bloody, and has a terrible odour.
- Pelvic pain or discomfort during sexual activity.

## II. PROPOSED SYSTEM

According to the World Health Organization, cervical cancer is the fourth most common cancer in women, accounting for 7.9% of all malignancies in women.

The absence of appropriate early therapy has been one of the primary reasons of this. As a result, early identification of cervical cancer is critical. The asymptomatic form of cervical cancer poses a significant barrier in early detection. Machine learning has proven to be beneficial in a variety of medical applications, and it can be utilised as a classifier in the early detection of malignant cells in the uterus's cervix region. Because of the complicated interdependence of different factors, diagnosing cervical cancer at an early stage is a difficult undertaking. The development of medical diagnostic decision support systems that can assist medical practitioners in the diagnostic process is crucial.

The major goal of this study is to compare the performance of machine learning algorithms that are used to predict diabetes. Finally, the best model for predicting cervical cancer is developed.

## III. RELATED WORK

Shimizu et al., [4] has expressed views on approach of AI in Oncology. Artificial Intelligence (AI) works similar to the human intelligence many computational algorithms are applied for recognition, prediction, classification and many other. Machine Learning (ML) is evaluated from AI and Deep learning is a part of ML. Deep Learning is applied in medicine for the purpose of diagnosis of diseases, classification of cells based on images obtained from medical tests done to predict the type of disease. AI has drawn its own specialization in the field of Medicine which made changes in workflow of clinical approaches towards the diseases. Accuracy, specificity, sensitivity and efficiency of medical results have been increased. In upcoming decades it is expected to have computational knowledge even in Medical Sciences for more accurate results with in less time span.

Hinton et al., [5] has a belief that Deep learning has changed the health-care in the upcoming decades. Deep learning has deep networks with input, output and many intermediate layers. Deep learning deals with complicated models in the field it is being used. It extracts features for the further workflow and has successfully exceeded the human experts in outcomes with less flaws. Neural Networks add an alternative “hidden layers” which deals the problems in more efficient way. The number of hidden layers are limited in order to avoid the complexity in the developed model. As new algorithms are discovered every time the potential of applications in various fields will increase rapidly.

Geeitha et al., [6] has proposed Machine Learning models in analysing gene expression combined with different Data Mining techniques. Also the importance of feature selection and classification techniques which classify benign and malignant cervical cells. According to the survey, it is said that Machine Learning can deal with large volumes of data and can easily provide accessibility to the detection of cancer. Comparative analysis of Machine Learning models such as hold out method, Backward Elimination Hilbert-Schmidt Independence Criterion (BAHSIC), Singular Value Decomposition Entropy gene selection (SVD Entropy), SMOTE technique, J48, Random tree forest, Boosted Decision Tree, Sequential Elimination approach and classifiers such as SVM, Naïve Bayes, Logistic Regression, KNN, ANN, MLP, WEKA segmentation classifier etc., are implemented. SVM provides higher prediction accuracy when combined with KNN and SMOTE which is developed for imbalanced data. Even though the prediction rate is higher the feature extraction and selection is a bit challenging thus researchers have decided in future to carry out the work in Deep Learning which can deal with number of large datasets with dimensionality reduction

Rayavarapu et al., [7] used two popular Machine Learning techniques Voting and Deep Neural Network (DNN) classifiers to predict growth of cancerous cells. By applying supervised algorithms prediction is done. The preprocessed data is given to the classifiers which divide the data into training and testing data. The result is given to the voting classifier based on the majority voting the class label is acknowledged. DNN is applied whereas when compared voting classifier is highly accurate. The feature extraction can be improved in extinction as DNN takes more time complexity

Devi et al., [8] expressed views on prediction of cells by applying Artificial Neural Networks (ANN) from the screening method results. ANN uses different architectures to obtain high accuracy rate in prediction and classification of cancerous cells. In future combination of Gene Algorithm with Feed Forward Neural Networks is used to predict cervical cancer cells from the pap and liquid-based cytology test images.

Kurnianingsih et al., [9] has expressed views that segmentation of cells and classification of cells play a major role in prediction of cancerous cells. In the proposed work Mask R-CNN is used for segmentation. The classification phase uses VGG-like Net. For the prior information ResNet10 is used in addition with Mask RCNN. Future work mainly focuses on the implementation of deeper network models to increase the performance.

Athinarayanan et al., [10] used an automated detection system to find the cells from the input images produced at the time of testing the patient. Texture features are premeditated in the decision making system along with SVM. The experimental results are compared with other classifiers KNN and ANN which resulted SVM is better classifier. This proposed system helps the physician to make decision

faster for further treatment of the patient as the classification of cells is done accurately.

Ghoneim et al., [11] has drawn a conclusion from previous cases that computer-aided systems are more accurate and Deep learning based systems are more efficient. Cervical Cancer prediction here is dealt with Convolutional Neural Networks (CNN) model followed with Extreme learning machine (ELM) based classifier. CNN extract features from raw input and that data is fed in to ELM-classifier, here two CNN models are used “VGG-16 Net and CaffeNet” for the efficient extraction of features and then output is fed into ELM-classifier which results the detection of cervical cancer cells.

Taha et al., [12] has focused mainly on the classification step as in the prediction of cancer classification of cells play a crucial role. Rather designing classification algorithm from scratch CNN combined with SVM is used which saves time consumption and is more accurate. CNN here is used for feature extraction as it studies the pap images layer by layer and SVM is used for the decision making purpose which helps to classify normal cells from the cancerous cells

Ungarpalli et al., [13] performed two stage of classification first is it classifies number of cells present in the input images and the second is the class of the cell. CNN is applied for the categorization of cells. Proposed work can predict two classes from the seven classes where as in the future it is expected to predict all the seven classes also provide the cell segmentation for the all classes in the given data accurately.

## IV. Methodology

### Algorithms for Classification

Classification is a model (functions) for describing and distinguishing classes of data or concepts with the

goal of predicting the class of an object whose label class is unknown. Classification is an aspect of data mining, which is a term used to describe the process of discovering knowledge in databases. Data mining is also a technique for extracting and finding meaningful information and relevant knowledge from massive databases using statistical approaches, mathematics, artificial intelligence, and machine learning.

### Classifier based on Decision Trees

The decision tree is a classification approach in which the splitting criteria are used to classify the data. The decision tree is a tree-like flow chart that sorts instances based on attribute (feature) values to classify them. Every node in a decision tree represents an attribute in a classification instance. Each leaf node stores the class label, and each branch represents a test result. The feature value is used to classify the instance. There are a variety of approaches for determining which feature divides the training data the best, including information gain, gain ratio, Gini index, and so on. Starting with the training set and recursively discovering a split is the most frequent strategy to build decision trees utilising top down greedy method partitioning. a feature that seeks to maximise a local requirement The classification rule for the data set is generated via a decision tree. ID3, C4.5, and CART are the three core algorithms that are extensively employed.

#### 4.1.1 ID3 algorithm

Iterative Dichotomer 3 (ID3) is a type of iterative dichotomer. Quinlan Ross first proposed the decision tree method in 1986. The core idea is to use a top-down greedy technique to create a decision tree. ID3 selects the best feature based on the information gathered. To determine the information gain, we must first determine the entropy.

$$\text{Entropy (s): } - \sum [P(I) \log_2 P(I)]$$

$$\text{Euclidean} = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}$$

Where  $P(I)$  is the proportion of  $S$  that belongs to Class  $I$ ,  $S$  denotes all records (instances),  $C$  denotes Class, and is over  $C$ , i.e. the sum of all classifiers.  $\text{Entropy}(S) - ((|S_v|/|S|) \text{Entropy}(S_v))$  Information Gain( $S, A$ ) =  $\text{Entropy}(S) - ((|S_v|/|S|) \text{Entropy}(S_v))$  Where  $A$  is the feature for which gain will be calculated,  $V$  is the feature's entire range of possibilities, and  $S_v$  is the number of elements for each  $V$ .

#### 4.1.2 C4.5 algorithm

Quinlan created the C4.5 decision tree algorithm. It's a variation on the ID3 algorithm. The statistic Classifier might be referred to as the C4.5. Gain ratio is used in this technique to choose features and build the decision tree. It is capable of dealing with both continuous and discrete features. Because of its speedy categorization and excellent precision, the C4.5 algorithm is widely employed. The information gain was "Normalized" using the gain ratio as follows:

$$\text{Information Gain}(S, A) / \text{Entropy}(S, A) = \text{Gain Ratio}(A, S)$$

#### 4.1.3 CART algorithm

Breiman introduced the Classification Regression Tree (CRT). CART uses binary splitting, which implies that each node has precisely two outgoing edges, and the Gini index is used to split them.

$$P_2 = 1 - \text{Gini Index}(I)$$

4.2 The KNN classifier is a K Nearest Neighbours (KNN) classifier.

K Nearest Neighbors (KNN) is a simple algorithm that maintains some available examples and categories (e.g., distance functions). A case is classified by a majority vote of its neighbours, with the case being allocated to the most common class among its KNN, as determined by a distance function. The equation gives the Euclidean distance between two points  $x$  and  $y$ . Inspection of the data set yields the value of  $k$  (the positive integer). Cross-validation is another

method for retrospectively determining a good  $k$  value by validating the  $k$  with an independent data set. We've used the values ( $k=1, 3$ ), and the outcome is good at  $k=5$ . This means that when the  $k$  value increases, the outcome will get more precise. Most of the time, the best  $k$  value is between 3 and 10.

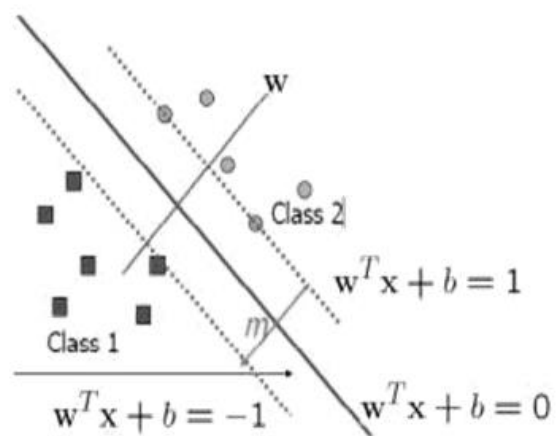
#### 4.3 Classification using the naive Bayes method.

The Bayesian Classification approach is both a supervised learning and a numerical classification method. It is based on an essential probabilistic model and allows us to ethically constrain uncertainty about the model by establishing probabilities of the outcomes. As a result, it can solve the process' diagnostic and predictive problems. The Bayesian classification is named after Thomas Bayes, 1702-1761, who discovered the Bayes Theorem. It provides practical learning methods, as well as prior information and observed data.

$$P(C_i | S) = P(S | C_i) P(C_i) / P(S)$$

The Bayesian Classification provides a useful framework for analysing and comprehending various learning methods. As a result, it estimates precise probabilities for hypotheses and is resistant to noise in the categorization input data.

#### 4.4 Support Vector Machine classifier



Two classes of SVM training

### Pre-Processing of Data

The extraction of valuable information and results is primarily dependent on the data's quality, while medical data is influenced by factors such as missing values, noisy data, inconsistencies, and outliers. As a result, before beginning the machine learning process, the data must be processed, and data pre-processing is an important step in increasing data efficiency. Data pre-processing includes data preparation and dataset transformation, which improves the effectiveness of knowledge discovery. The following steps were used to pre-process data in this project:

#### 1. Missing data

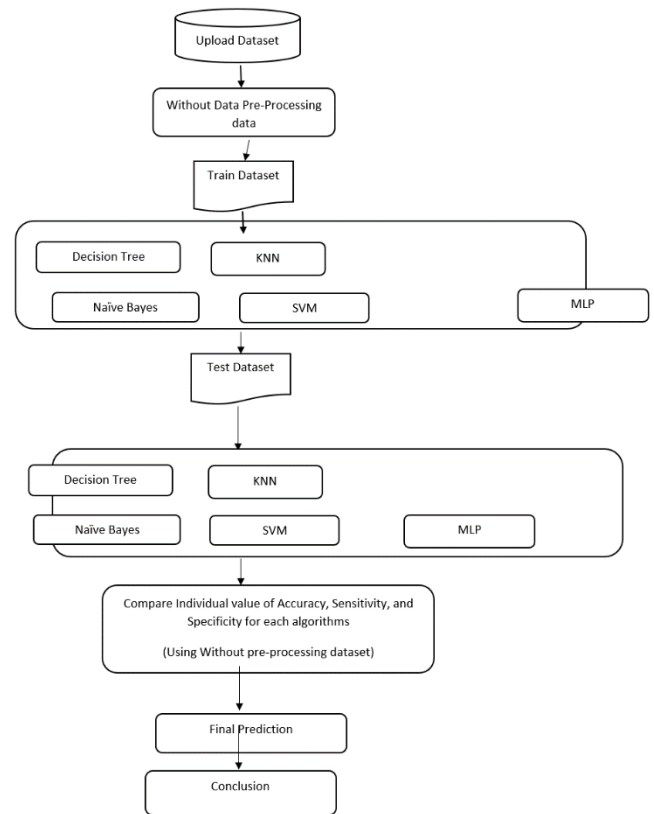
Missing data refers to data values that are not stored in the dataset for a variable or attribute. Missing data are a significant issue in the data analysis process because data loss is common, particularly in medical data. Missing data is critical because most analytical methods cannot be applied to an incomplete dataset because the quality of the machine learning model suffers as a result. As a result, missing values must be taken into account when calculating missing values with reasonable values [38]. Some algorithms, such as scikit-learn methods, make the assumption that all values are present and have a meaningful value. The cervical cancer dataset contains many blank values, which are the result of some patients declining to answer certain questions due to concerns about their personal privacy. Missing data in the cervical cancer dataset can be handled in two ways: ignore (remove) missing values or impute (fill) missing values.

#### 2. Ignore missing values

Ignoring some features helps to make data consistent because they have a high percentage of missing values. This method is useful because some features in the dataset have missing values, such as "Time since first diagnosis" and "Time since last diagnosis," where the missing values account for more than 80% of all data in these two attributes. These two attributes were

excluded due to the difficulty in filling in such a large proportion of missing values with meaningful values and the lack of attribute dependencies that can be used to derive values for the missing data.

#### Architecture diagram



#### Impute missing values:

One of the most common methods in the field of missing data is imputation, which fills missing values with appropriate values. Many features in the cervical cancer dataset have missing values of less than 20%. These missing values were recorded in the dataset as "?" and imputed in one of two ways Imputation rested on mean values this is the most universally used insinuation tactics. This method works by calculating the mean value of non-missing data in a specific column and then replacing missing values with that value in that column.

#### 4. Data transformation:

During the data transformation step, the data is converted or consolidated so that the processing results are more efficient and the existing patterns are



easier to understand. The data is then ready for processing and the application of machine learning algorithms. Normalization is a data transformation strategy that refers to the process of scaling feature values to be within a small specified range or common range, such as [0,1] or [-1,1]. Normalization methods include Min-Max, decimal, and Z-score normalisation. This dataset was normalised using the min-max method.

**5.Outliers:**

Finding outliers in a dataset is a difficult and time-consuming process, especially for high-dimensional datasets. Outliers are observations or subsets of observations that appear to be inconsistent with the rest of the dataset, whereas outlier detection is the search for objects in the dataset that are not subject to the laws that apply to the majority of the remaining data. Outliers in medical data are most commonly caused by medical device malfunctions, human errors, patient-specific behaviours, natural changes in the patient, medication intake, food or alcohol, stress, and others. In some cases, the outlier value provides useful information because it may indicate a rare disease; thus, outlier values are typically treated by retaining, removing, or modifying them. Outlier values were retained in this paper because they explain people's social situations and differences.

**II. EXPERIMENTATION**

In this work, we used a cervical cancer dataset from UCI that contains the demographics, behaviours, and medical histories of 858 patients at Hospital University de Caracas in Venezuela. The dataset has 32 attributes and four target variables. The type and description of the attributes. These characteristics are cervical cancer. The Hinselmann, Schiller, Cytology, and Biopsy results, which are the most common cervical cancer diagnosis methods, are the target variables. To categorise the data using the techniques stated above, we use those four variables as classification labels. The time from initial diagnosis and time since last diagnostic missing data are excessive, so they are eliminated to lessen the impact of missing values. Some pieces of data that had too many missing values were also eliminated to verify the accuracy of the results. Then there are 668 records in the dataset we used to examine.

**III. RESULTS AND DISCUSSIONS**

Six classifier approaches for classifying the cervical cancer dataset are proposed in this paper. The classifier is trained and tested, with the results being kept track of. The following figures exhibit a graphical depiction of the Confusion matrix and its ROC curve for five classification techniques.

Table 1 shows the description and kind of certain properties

S.no	Attribute name	Type	S.no	Attribute name	Type
1	Age	int	19	STDs: pelvic inflammatory disease	bool
2	No. of sexual partners	int	20	STDs: genital herpes	bool
3	First sexual intercourse(age)	int	21	STDs: molluscum conttagiosum	bool
4	No. of pregnancies	int	22	STDs: AIDS	bool
5	Smokes	bool	23	STDs: HIV	bool
6	Smokes(years)	bool	24	STDs: hepatitis B	bool
7	Smokes(packs/year)	bool	25	STDs: HPV	bool

8	Hormonal contraceptives	bool	26	STDs: no.of diagnosis	int
9	Hormonal contraceptives years	int	27	STDs: time since first diagnosis	int
10	IUD	bool	28	STDs: time since last diagnosis	int
11	IUD(years)	int	29	Dx: cancer	bool
12	STDs	bool	30	Dx: CIN	bool
13	STDs (number)	int	31	Dx: HPV	bool
14	STDs:condylomatosis	bool	32	Dx	bool
15	STDs: cervical condylomatosis	bool	33	Hinselmann: target variable	Bool
16	STDs: vaginal condylomatosis	bool	34	Schiller: target variable	Bool
17	STDs: vulvo-perineri condylomatosis	bool	35	Cytology: target variable	bool
18	STDs: syphilis	bool	36	Biopsy: target variable	bool

Figure 5.1 Decision tree Confusion matrix

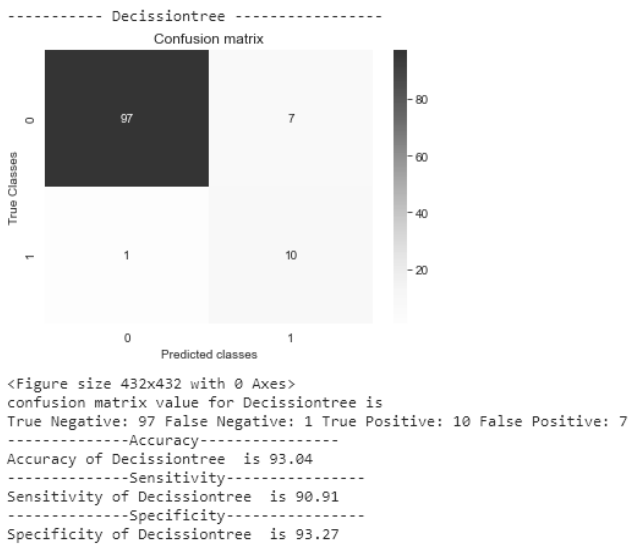


Figure 5.1.1 Decision tree ROC curve

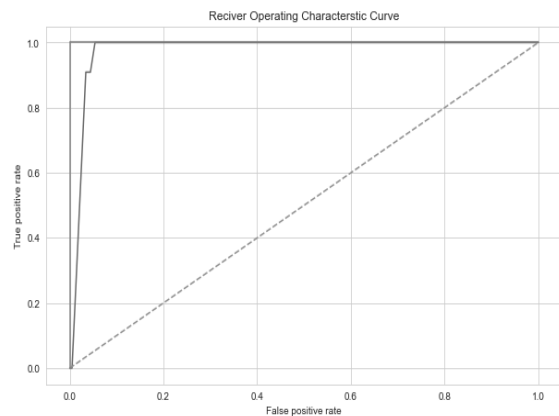


Table 1

Decision Tree	
True Negative	99
False Negative	5
True Positive	9
True Positive	2



Figure 5.2 SVM confusion matrix

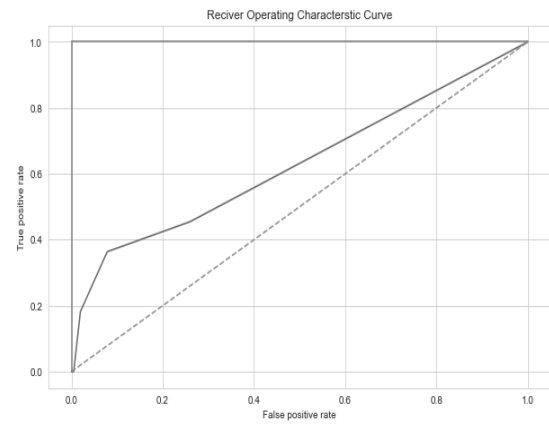
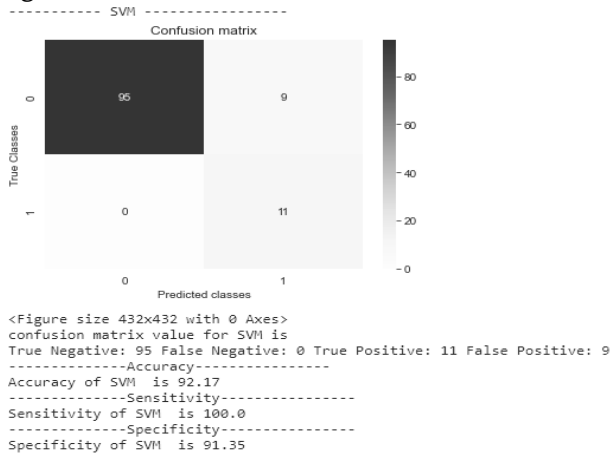


Table 3

Table 2

SVM	
True Negative	98
False Negative	4
True Positive	10
True Positive	3

KNN	
True Negative	1
False Negative	14
True Positive	0
True Positive	0

Figure 5.3.1 KNN confusion matrix

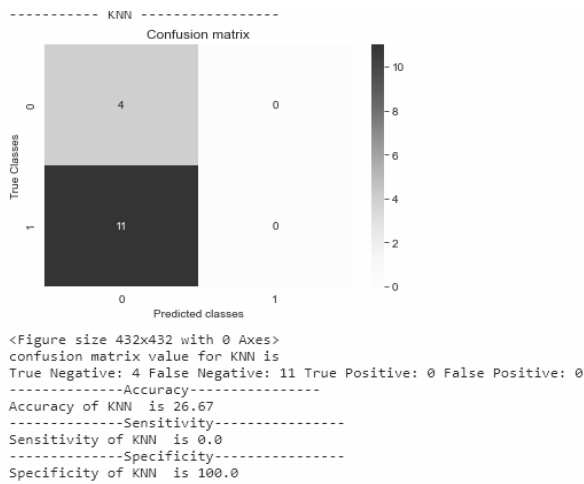


Figure 5.4 NV confusion matrix

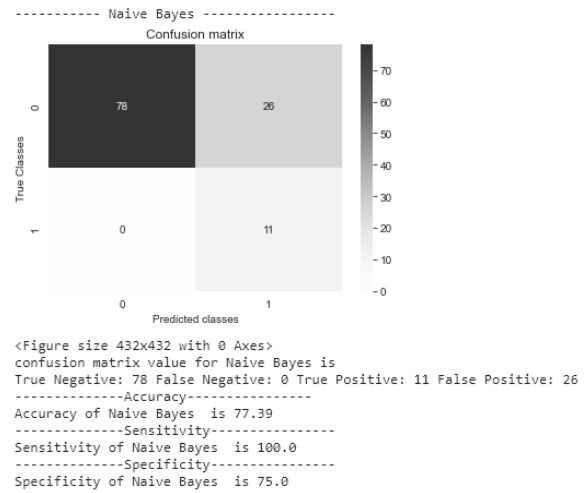
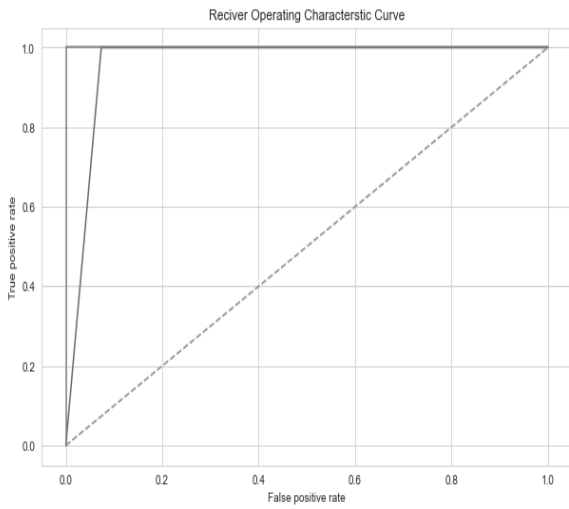


Figure 5.3.2 KNN ROC curve

Figure 5.4.1 NV ROC curve



True positive	6
True positive	2

Figure 5.5.1 MLP ROC curve

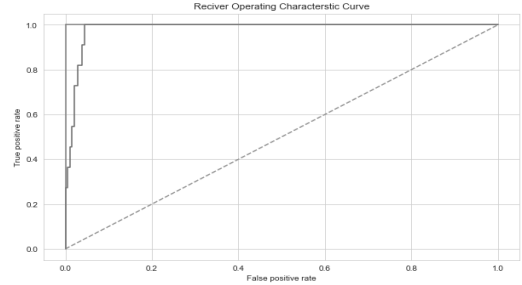


Table 4

Naïve Bayes	
True Negative	9
False Negative	0
True Positive	14
True Positive	92

Figure 5.5 MLP confusion matrix

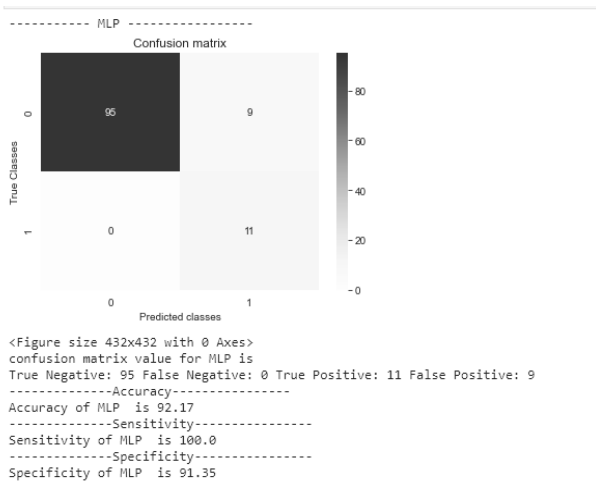


Table 5

MLP	
True negative	99
False negative	8

Here a comparison is made to find which classifier is the best for classification. The resultant graph showed in the following figure5.2 and figure 5.5.1

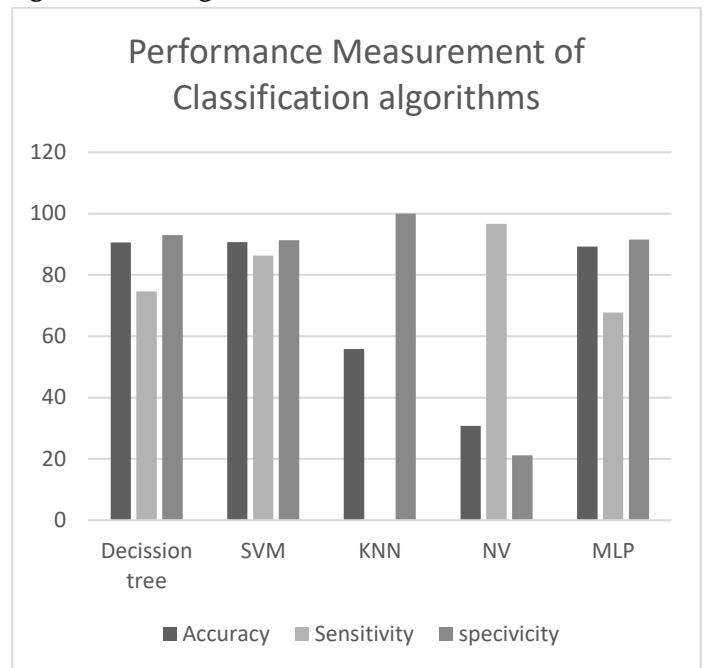


Figure 5.6

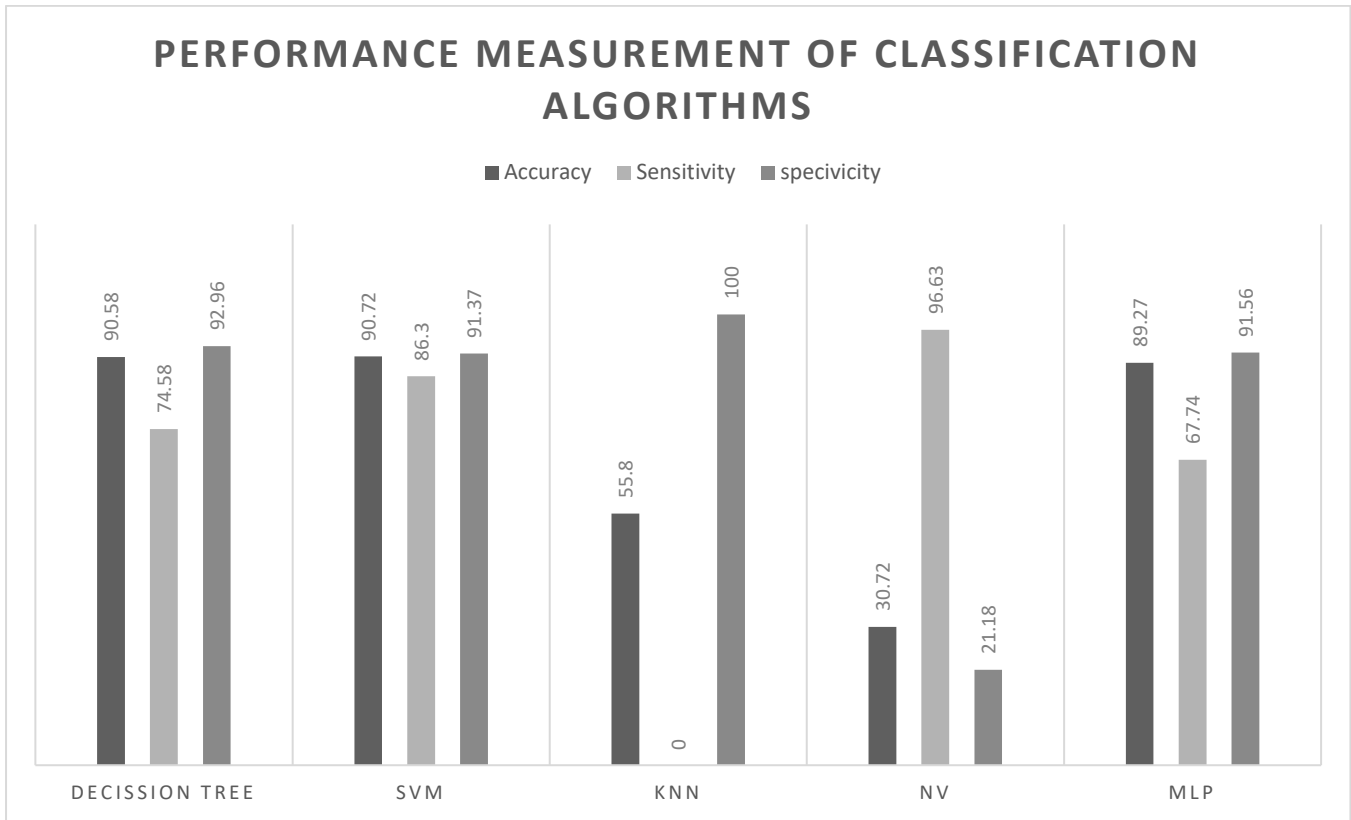


Figure 5.7

Table 6

Sno	Author	Techniques Applied	Dataset	Result
1	Shimizu	CNN, FCN	ILSVRC dataset	CNN
2	Geoffrey Hinton	ANN	-	ANN
3	S. Geeitha, M. Thangamani	SVM, Naive Bayes, Logistic regression, KNN, ANN, MLP	Cervical cancer dataset	ANN 93%
4	Komala Rayavarapu	DNN	The cervical cancer (risk factors) datasets from UCI	DNN 88% accuracy
5	M. Anousouya Devi, S. Ravi, J. Vaishnavi and S. Punitha	ANN, MLP, Backpropagation	-	ANN
6	KURNIANINGSIH	Mask R CNN, VGG	eHerlev dataset (original image data collection)	VGG
7	S. Athinarayanan, M.V. Srinath	SVM, KNN and ANN	Original Image dataset	KNN and ANN

8	Ghoneim, Ahmed, Muhammad, Ghulam; Hossain, M. Shamim	VGG,CNN, Shallow DNN	Original Image dataset	VGG 16
9	Taha	SVM with CNN	Original Image dataset	SVM with CNN
10	Manasa Ungrapalli N S, Myna A N	CNN and DNN with Different classes	-	DNN

#### IV. RESULTS AND DISCUSSIONS

Six classifier approaches for classifying the cervical cancer dataset are proposed in this paper. The classifier is trained and tested, with the results being kept track of. The following figures exhibit a graphical depiction of the Confusion matrix and its ROC curve for five classification techniques.

#### V. CONCLUSION

In the healthcare industry, the healthcare businesses are more crucial than ever for retaining the relationship between various symptoms and past medical history. Understanding the relationship between patient disease symptoms and historical medical records is critical to success and survival in the healthcare industry. In such healthcare methods, machine learning and predictive analytics might provide important information. Its applications are used in practically every field that has a high number of complex data sets and procedures. Its ability to lower the amount of false positives and false negative choices has been demonstrated. On the Cervical cancer risk factor dataset, this project evaluated and compared the classification performance of five different Machine Learning techniques' models Decision tree, Nave Bayes, KNN, SVM, and MLP. We used 858 patients and 36 distinct variables to distinguish who has disease and who does not in this experiment. The classification performances of the five models have been using three statistical measures; Classification accuracy, sensitivity, specificity. This data set has partitioned into training and test by the

ratio 75% and 25%, respectively. Experimental results have shown the effectiveness of models. Decision tree has achieved slightly better performance than MLP and SVM. Meanwhile Naïve Bayes and KNN algorithms went under performance category because the nature of unbalanced dataset.

#### VI. FUTURE WORK

In the future, we will improve the dataset used to detect cervical cancer and increase the efficiency of future prediction models by including several essential attributes that aid in the early detection of cervical cancer. Some data could be gathered and added to the dataset. To find the best classification method, multiple classification algorithms can be used. Because this is a medical problem, and cervical cancer is more of a personal female problem, data authenticity is always an issue. The emphasis should be on obtaining data that is accurate and reliable.

#### VII. REFERENCES

- [1]. All about Cancer, Cancer Society of Finland, Available at: <https://www.allaboutcancer.fi/facts-aboutcancer/detection/#8667b054> (Accessed date 25.07.2020).
- [2]. American Cancer Society, Cancer Facts for Women, Available at: <https://www.cancer.org/healthy/findcancer-early/womens-health/cancer-facts-for-women.html>(Accessed date 25.07.2020).
- [3]. American Cancer Society, What is Cervical Cancer?, Available at:

- <https://www.cancer.org/cancer/cervical-cancer/about/what-is-cervical-cancer.html>  
(Accessed date 25.07.2020).
- [4]. Shimizu, H., & Nakayama, K. I. (2020). Artificial intelligence in oncology. *Cancer Science*, 111(5), 1452–1460. <https://doi.org/10.1111/cas.14377>
- [5]. Devi, M. A., Ravi, S., Vaishnavi, J., & Punitha, S. (2016). Classification of Cervical Cancer Using Artificial Neural Networks. *Procedia Computer Science*, 89, 465–472. <https://doi.org/10.1016/j.procs.2016.06.105>
- [6]. Geeitha, S., & Thangamani, M. (2020). A cognizant study of machine learning in predicting cervical cancer at various levels-a data mining concept. *International Journal on Emerging Technologies*, 11(1), 23–28.
- [7]. Rayavarapu, K., & Krishna, K. K. V. (2018). Prediction of Cervical Cancer using Voting and DNN Classifiers. *Proceedings of the 2018 International Conference on Current Trends towards Converging Technologies, ICCTCT 2018*, 1–5. <https://doi.org/10.1109/ICCTCT.2018.8551176>
- [8]. Hinton, G. (2018). Deep learning—a technology with the potential to transform health care. *JAMA - Journal of the American Medical Association*, 320(11), 1101–1102. <https://doi.org/10.1001/jama.2018.11100>
- [9]. Kurnianingsih, Allehaibi, K. H. S., Nugroho, L. E., Widyawan, Lazuardi, L., Prabuwo, A. S., & Mantoro, T. (2019). Segmentation and classification of cervical cells using deep learning. *IEEE Access*, 7, 116925–116941. <https://doi.org/10.1109/ACCESS.2019.2936017>
- [10]. S., A., & M.V., S. (2016). Classification of Cervical Cancer Cells in Pap Smear Screening Test. *ICTACT Journal on Image and Video Processing*, 06(04), 1234–1238. <https://doi.org/10.21917/ijivp.2016.0179>
- [11]. Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. *Future Generation Computer Systems*, 102, 643–649. <https://doi.org/10.1016/j.future.2019.09.015>
- [12]. N, T. . B. D. . J. (2017). Classification of Cervical Cancer Using Pap-Smear Images: A Convolutional Neural Network Approach. *Department of Electrical and Computer Engineering*, 1(d), 698–706. <https://doi.org/10.1007/978-3-319-60964-5-23>
- [13]. Manasa Ungrapalli, N. S., & Myna, A. N. (2019). Classification of pap smear images for cervical cancer using convolutional neural network. *International Journal of Innovative Technology and Exploring Engineering*, 9(1), 2801–2807. <https://doi.org/10.35940/ijitee.J1226.119119>
- [14]. W. Wu and H. Zhou, "Data-Driven Diagnosis of Cervical Cancer With Support Vector Machine-Based Approaches," in *IEEE Access*, vol. 5, pp. 25189–25195, 2017.
- [15]. Y. E. Kurniawati, A. E. Permanasari and S. Fauziati, "Comparative study on data mining classification methods for cervical cancer prediction using pap smear results," 2016 1st International Conference on Biomedical Engineering (IBIOMED), Yogyakarta, 2016, pp. 1-5
- [16]. R. Vidya1\* and G. M. Nasira2, "Prediction of Cervical Cancer using Hybrid Induction Technique: A Solution for Human Hereditary Disease Patterns", *Indian Journal of Science and Technology*, August 2016.
- [17]. D. Kashyap et al., "Cervical cancer detection and classification using Independent Level sets and multi SVMs," 2016 39th International Conference on Telecommunications and Signal Processing (TSP), Vienna, 2016, pp. 523-528.
- [18]. E. Njoroge, S. R. Alty, M. R. Gani and M. Alkatib, "Classification of Cervical Cancer Cells using FTIR Data," 2006 International Conference of the IEEE Engineering in

Medicine and Biology Society, New York, NY, 2006, pp. 5338-5341.

- [19]. J. Hyeon, H. J. Choi, K. N. Lee and B. D. Lee, "Automating Papanicolaou Test Using Deep Convolutional Activation Feature," 2017 18th IEEE International Conference on Mobile Data Management (MDM), Daejeon, 2017, pp. 382-385.
- [20]. K. Teeyapan, N. Theera-Umpon and S. Auephanwiriyakul, "Application of support vector based methods for cervical cancer cell classification," 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE), George Town, 2015, pp. 514-519.

**Cite this article as :**

Prathamesh D. Mane, Dr. Surabhi Thorat, "Heart Stroke Diagnosis using AI Model", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 6, pp. 241-244, November-December 2021. Available at doi : <https://doi.org/10.32628/CSEIT217652>  
Journal URL : <https://ijsrcseit.com/CSEIT217652>