Sl No. 1	Development of a cervical cancer progress prediction tool for human papillomavirus-positive Koreans: A support vector machine-based approach
Author	1- Jimin Kahng , 2-Eung-Hee Kim, 2-Hong-Gee Kim and 3-Wonbae Lee
Year	2015
Methodology	Records were retrospectively analysed from women who were positive for HPV on initial testing (before any treatment). Information concerning age, Papanicolaou (PAP) smear result and presence of 15 high-risk HPV genotypes was used in a support vector machine (SVM) model, to identify the patient features that maximally contributed to progression to high-risk cervical lesions. HPV testing -The HPV tests were performed using the HPVDNAChip kit BioMedLab, Seoul, Republic of Korea as previously described. Using a sterilized vaginal speculum, cervical specimens were collected during colposcopic examination by inserting a cytobrush attached to an HPVDNAChip Sampler BioMedLab into the endocervical and exocervical areas. After removal, the cytobrush was placed into transport medium 2 ml phosphate buffered saline and immediately refrigerated at 4-degree Celsius for 72 h before analysis.
Model Used	Support vector machine (SVM) model
Data sets used	

Table 1. Papanicolaou (PAP) smear and biopsy results in subjects positive for human papillomavirus.

PAP smear	n (%)
No result	66 (9.03)
Normal/inflammation	183 (25.03)
Atypical squamous cells of undetermined significance	183 (25.03)
Low-grade squamous intraepithelial lesion	174 (23.80)
High-risk lesion[a]	125 (17.10)
Total	731 (100.00)
Biopsy	
No result	144 (19.70)
Normal	50 (6.84)
Koilocytosis and cervical intraepithelial neoplasia I	226 (30.92)
High-risk lesion[b]	311 (42.54)
Total	731 (100.00)

[a]High-grade squamous intraepithelial lesion (HSIL), atypical squamous cells that cannot exclude HSIL, carcinoma in situ, squamous cell carcinoma, adenosquamous carcinoma, atypical glandular cells-favour neoplasia, adenocarcinoma or neuroendocrine carcinoma.
[b]Cervical intraepithelial neoplasia (CIN) 2, CIN3, carcinoma in situ or invasive carcinoma.

Table 2. The seven most frequently occurring genotypes in subjects ($n = 731$) positive for human papillomavirus (HPV), classified according to the development of high-risk lesions.

Genotype	Subjects, n	Group I $n = 35$	Group II $n = 386$	Group III $n = 310$
HPV16	207 (28.32)	17 (48.57)[a]	70 (18.13)	137 (44.19)[a]
HPV18	51 (6.98)	1 (2.86)	28 (7.25)	23 (7.42)
HPV35	40 (5.49)	—	20 (5.18)	20 (6.45)
HPV51	65 (8.89)	2 (5.71)	49 (12.69)[a]	16 (5.16)
HPV52	103 (14.09)	7 (20.00)[a]	64 (16.58)	39 (12.58)
HPV56	57 (7.80)	1 (2.86)	44 (11.40)[a]	13 (4.19)
HPV58	90 (12.31)	5 (14.29)	42 (10.88)	48 (15.48)

Data presented as n(%) of subjects.
Group I, subjects with benign lesions at initial HPV test who subsequently developed high-risk lesions; group II, subjects with regressive or persistent HPV infection who did not develop high-risk lesions; group III, subjects diagnosed with high-risk lesions at initial HPV test.
[a]$P < 0.05$; Fisher's exact test.

Table 3. Specificity, sensitivity and negative and positive predictive values for the development of high-risk lesions for the seven most frequently occurring genotypes in subjects positive for human papillomavirus (HPV).

Genotype	Specificity	Sensitivity	Negative predictive value	Positive predictive value
HPV16	333/386 (86.27)	154/345 (44.64)	333/524 (63.55)	154/207 (74.40)
HPV18	359/386 (93.01)	24/345 (6.96)	359/680 (52.79)	24/51 (47.06)
HPV35	366/386 (94.82)	20/345 (5.80)	366/691 (52.97)	20/40 (50.00)
HPV51	339/386 (87.82)	18/345 (5.22)	339/666 (50.90)	18/65 (27.69)
HPV52	329/386 (85.23)	46/345 (13.33)	329/628 (52.39)	46/103 (44.66)
HPV56	343/386 (88.86)	14/345 (4.06)	343/674 (50.89)	14/57 (24.56)
HPV58	349/386 (90.41)	53/345 (15.36)	349/641 (54.45)	53/90 (58.89)

Data presented as n subjects affected/total n subjects (%).

Table 4. Support vector machine model construction using age, Papanicolaou smear result (PAP) and 15 high-risk genotypes in subjects positive for human papillomavirus (HPV): at each selection step a further feature is added and the predictive accuracy of the steps up to that point is calculated.

FSS/AF/accuracy(%)	Confusion matrix			FSS/AF/accuracy(%)	Confusion matrix		
1/PAP/71.13		ω_1	ω_2	10/HPV33/71.54		ω_1	ω_2
	ω_1	358	28		ω_1	311	75
	ω_2	183	162		ω_2	133	212
2/HPV16/75.10		ω_1	ω_2	11/HPV51/71.81		ω_1	ω_2
	ω_1	335	51		ω_1	308	78
	ω_2	131	214		ω_2	128	217
3/HPV52/74.55		ω_1	ω_2	12/HPV66/71.81		ω_1	ω_2
	ω_1	319	67		ω_1	310	76
	ω_2	119	226		ω_2	130	215
4/HPV35/74.41		ω_1	ω_2	13/HPV45/72.36		ω_1	ω_2
	ω_1	316	70		ω_1	314	72
	ω_2	117	228		ω_2	130	215
5/HPV58/74.14		ω_1	ω_2	14/HPV18/72.22		ω_1	ω_2
	ω_1	321	65		ω_1	312	74
	ω_2	124	221		ω_2	129	216
6/HPV31/72.64		ω_1	ω_2	15/HPV69/72.09		ω_1	ω_2
	ω_1	316	70		ω_1	315	71
	ω_2	130	215		ω_2	133	212
7/HPV56/72.22		ω_1	ω_2	16/HPV68/72.09		ω_1	ω_2
	ω_1	312	74		ω_1	316	70
	ω_2	129	216		ω_2	134	211
8/HPV39/72.22		ω_1	ω_2	17/Age/65.80		ω_1	ω_2
	ω_1	313	73		ω_1	282	104
	ω_2	130	215		ω_2	146	199
9/HPV59/72.09		ω_1	ω_2				
	ω_1	314	72				
	ω_2	132	213				

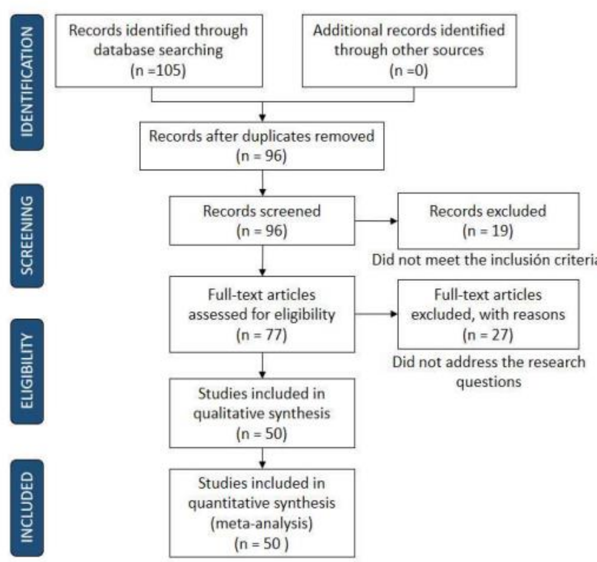
FSS, forward selection step; AF, added feature.

Accuracy Percentage	The maximum number of correct cancer predictions (228/345) is seen when four features (PAP, HPV16, HPV52 and HPV35) are used, with an accuracy of 74.41%.
Advantages	Applicable where number of features is larger than the number of samples, such as in image recognition tasks. Less affected by the presence of outliers. Works well with both linear and non-linear problems. Fast prediction.
Limitations	Computationally expensive- requires significant amt of time and memory to train, especially for large datasets.
Sl No. 2	Machine Learning Analysis for Cervical Cancer Prediction, a Systematic Review of the Literature

Author	1-Sandy Gutierrez-Espinoza, 2-Michael Cabanillas-Carbonell.
Year	2021
Methodology	The proposed study is a systematic review of the scientific literature focused on the use of machine learning for the prediction and diagnosis of cervical cancer. The research questions posed include (RQ1) identifying models and metrics that can improve the accuracy of diagnosing cervical cancer, (RQ2) exploring machine learning methodologies to improve the prediction of cervical cancer results, and (RQ3) identifying countries that have conducted the most cervical cancer research with machine learning. The search strategy involved searching for articles in major databases, including ProQuest, IEEE Xplore, PubMed, ScienceDirect, Springer, IopScience, and Scopus. The search keywords used were "cervical cancer AND machine learning AND prediction model." Inclusion and exclusion criteria were applied, including inclusion criteria for articles related to cervical cancer prediction, implemented machine learning models in cervical cancer, and medical articles related to cervical cancer, and exclusion criteria for unrelated articles on the development of machine learning for cervical cancer, articles not applied to cervical cancer prediction, and articles published before 2014.

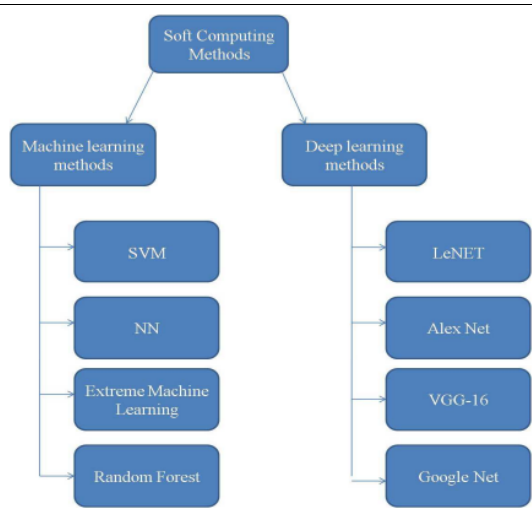
TABLE I. INCLUSION AND EXCLUSION CRITERIA

CRITERIA		
Inclusion	I01	Articles related to cervical cancer prediction
	I02	Implemented articles with machine learning model in cervical cancer
	I03	Medical articles related to cervical cancer.
Exclusion	E01	Unrelated articles on the development of machine learning for cervical cancer.
	E02	Articles not applied to cervical cancer prediction.
	E03	Articles that have been published before 2014

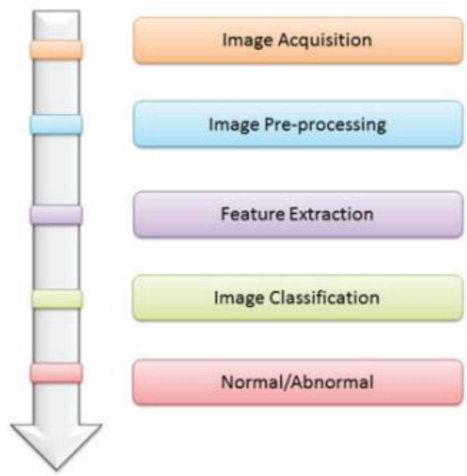


Model Used	Neural Network, Tree C5,

	SVM, Decision Trees, Random Forest, KNN, Naive Bayes, Convolutional Neural Network, Decision Trees, XGBoost and AdaBoost, Random Forest, KNN, Support Vector Machine, Multi-Layer Perceptron.
Data sets used	A series of searches of articles published in the main databases ProQuest, IEEE Xplore, PubMed, ScienceDirect, Springer, IopScience and Scopus were carried out. A total of 105 scientific articles were collected and 50 were systematized.
Accuracy Percentage	The prediction of cervical cancer that has been found to be most accurate in the analysis is that of the Convolutional Neural Networks - 99.5% and C5 Tree classifiers - 97% have performed reasonably well, as well as being extremely accurate through reliable results with the highest accuracy in identifying women presenting with clinical signs of cervical cancer.
Advantages	CNN- Parameter Sharing -reduces the number of parameters and allowing for faster training and better generalization. Hierarchical Learning: CNNs learn features hierarchically, meaning that lower-level features such as edges and corners are learned first, followed by more complex features such as object parts and textures. Efficient Memory Usage. C5 Tree Classifier- Accuracy: uses statistical methods to determine the best split for each node. handle large datasets, handle both categorical and continuous variables and can also handle missing data, C5.0 generates a decision tree that is easy to understand and interpret Decision Trees Easy to understand and interpret, can handle both categorical and numerical data, Can handle missing data, Can handle nonlinear relationships by using a combination of splits and decision rules. Fast and efficient. KNN- KNN is a simple algorithm that can be easily implemented and understood by beginners. No assumptions about data distributions. No training phase. Can be used for both classification and regression.
Limitations	CNN- CNNs require large amounts of labeled data for training, Computationally Expensive, overfitting-meaning that they may perform well on the training data but poorly on new, unseen data. C5 Tree Classifier- Overfitting, Biased, Sensitivity to outliers. Decision Trees Overfitting, an be unstable as small changes in the data can result in a different tree structure. May be biased towards features with many levels or categories. Decision trees may not be robust as they can easily be influenced by outliers and noise in the data and hence need for careful tuning.

	KNN- Computationally expensive when the dataset is large. Requires a proper choice of K, if K is too small, the algorithm may be too sensitive to noise, while if K is too large, it may miss important patterns in the data, Can be biased towards the majority class.
Sl No. 3	**An extensive study of cervical cancer detection methods using machine and deep learning approaches.**
Author	1- Divya Francis, 2- R Chitra, 3- Kasthuri Rengan.P, 4- Sri Vignesh.G
Year	2022
Methodology	 **Figure 2 Classification approaches for cervical cancer detection**
Model Used	Support Vector Machine (SVM), Neural Networks (NN), Random Forest (RF) algorithm and Extreme Machine Learning (EML) algorithm. Further, the deep learning models are categorized into LeNET, Alex net, VGG-16 and Google Net models.
Data sets used	The open access dataset was utilised by numerous researchers for their studies on cervical cancer detection. The cervical images in paid dataset can be obtained by paid manner and they are limited to open access. Mostly used dataset are MobileODT [16], Kaggle [17], SEVIA [19], and COCO2017 [18].

Accuracy Percentage	Colposcopy Ensemble Network, by incorporating the system with several deep learning models, Visual Geometry Group (VGG)-19, accuracy achieved was 92%. Neural Networks (NN) model- accuracy was 98.1%. The developed CNN cervical with 95% of classification accuracy, in first category dataset containing the cervical images belonging to low resolution category and the developed model obtained 96% of classification accuracy on second dataset category containing the cervical images belonging to high resolution category. SVM and a Gaussian kernel gave accuracy of 88.7%. SVM and a polynomial kernel gave accuracy of 85.1%. Radial kernel's incorporation of SVM gave accuracy of 90.6% .
Advantages	Neural Network(NN) more accurate than linear models. Neural networks can adapt to new data, automatically extract relevant features from the input data. Ensemble Network with Visual Geometry Group (VGG)- Improved accuracy, Scalability, robustness. SVM and a Gaussian kernel- Non-linear decision boundaries making them effective for a wide range of classification problems. highly flexible and can model complex relationships between features, relatively robust to noise, Fewer assumptions SVM and a polynomial kernel- Non-linear decision boundaries, flexible, robust , fewer assumptions. Radial kernel's incorporation of SVM- Non-linear decision boundaries, flexible, robust , fewer assumptions.
Limitations	Limitation- Neural Network- challenging to interpret and debug, long time to train, Overfitting. Ensemble Network with Visual Geometry Group (VGG)- More complex than single models, require more resources, such as computational power and memory, than single models, less interpretable than single models. Ensemble learning method failed to detect the low resolution cervical images as the main limitation of this work SVM and a Gaussian kernel- Overfitting, black box nature-a black box that maps the input to the output without providing info about relationships between features.

	SVM and a polynomial kernel- Overfitting, black box nature. Radial kernel's incorporation of SVM- Overfitting, black box nature, computationally intensive.
Sl No. 4	**Automated Cervical Cancer Image Analysis using Deep Learning Techniques from Pap-Smear Images: A Literature Review**
Author	1- Harmanpreet Kaur, 2- Dr. Reecha Sharma, 3- Dr. Lakhwinder Kaur
Year	2021
Methodology	 Illustratating the various steps of cervical cancer detection First, digital images are acquired from pap-smear slides, and then pre-processing techniques are applied to remove redundant information and improve accuracy. Data augmentation techniques can be used to increase the number of images in datasets and prevent over-fitting. Next, deep transfer learning models, which are pre-trained on natural Image-net dataset, are used to extract discriminant features from raw images in the dataset. Several pre-trained models are available, including LeNet-5 (1998), AlexNet (2012), VGG-16 (2014), Inception-v1 (2014), Inception-v3 (2015), ResNet-50 (2015), Xception (2016), Inception-v4 (2016), Inception-ResNets (2016), and ResNeXt-50 (2017). Finally, classification techniques are used to evaluate the stages of cancer or pre-cancerous lesions.
Model Used	InceptionV3 pre-trained, VGG19 pre trained, SqueezeNet pre-trained, and ResNet50 CNN pre-trained model, SVM, Artificial Neural Network based on Multi-Layer Perceptron (ANN-MPL).

<table>
<tr><td></td><td>Enhanced nucleus cell segmentation algorithm by using Fuzzy c-means (FCM) clustering and Back Propagation Neural Network (BPNN).

K nearest-neighbor, naïve Bayes, logistic regression, random forest, and support vector machines
are used for the classification.</td></tr>
<tr><td>Data sets used</td><td>
<table>
<tr><th>AUTHOR'S NAME</th><th>NAME OF SOURCES (N/S) AND NO OF SAMPLES (SIZE) IN THE DATASET</th><th>TECHNIQUES USED</th><th>QUANTITATIVE RESULTS</th></tr>
<tr><td>Khamparia et al. (2020) [8]</td><td>N/S: Herlev University (Pap smear images) dataset
Size: 917 Pap smear images; the size of 256 × 256 pixels</td><td>InceptionV3, VGG19, SqueezeNet and ResNet50 for feature extraction
K nearest neighbor, naïve Bayes, logistic regression, random forest and support vector machines</td><td>Accuracy = 97.89%</td></tr>
<tr><td>A. Ghoneim et al. 2020 [9]</td><td>N/S: Herlev University (Pap smear images) dataset
Size: 917 Pap smear images; the size of 256 × 256 pixels</td><td>VGG-like Net, Shallow, CaffeNet, Extreme Learning Machine</td><td>Accuracy = 91.2%</td></tr>
<tr><td>Kudva et al. (2019)[11]</td><td>N/S: Private (Kasturba Medical College, Manipal, India) dataset</td><td>Alexnet and VGG-16 neural network</td><td>Accuracy = 91.46%</td></tr>
<tr><td>HaomingLinet al.(2019) [12]</td><td>N/S: Herlev University Hospital; Size: 917 images (256 × 256 × 3 pixels)</td><td>AlexNet, GoogleNet, ResNet and DenseNet- CNN models</td><td>for 2-class -
Accuracy = 94.5%;
for 7-class
Accuracy = 64.5%</td></tr>
<tr><td>Kurnianingsih et al. (2019) [13]</td><td>N/S: Herlev University Hospital; Size: 917 images (256 × 256 × 3 pixels)</td><td>R-CNN and VGGNet</td><td>Precision = 92%; ,
Recall= 91%, and ZSI value of 90%,</td></tr>
<tr><td>W. William et al. (2019) [14]</td><td>N/S: Private and Herlev University Hospital; Size: 917 images (256 × 256 × 3 pixels)</td><td>fuzzy c-means algorithm</td><td>Accuracy = 95.20%</td></tr>
<tr><td>Long Nguyen et al. (2019) [15]</td><td>N/S: Hela dataset, and Hep-2 datasets;
Size: 2D Hela datasets: 917 images (45 × 43 to 768 × 284); Hep-2 dataset: images (33 × 38to396 × 295)</td><td>Inception-v3, ResNet152, Inception- ResNet-v2.</td><td>Accuracy= 93:04%</td></tr>
<tr><td>Sompawong et al. (2019) [17]</td><td>N/S: Thammasat University (TU)Hospital.; Size: 178 liquid-based histological images (3,048×2048 pixels)</td><td>Mask Regional Convolutional Neural Network (Mask R-CNN)</td><td>Accuracy = 89.8%;
Sensitivity = 72.5%,
Specificity = 94.3%.</td></tr>
<tr><td>Wu et al. (2018) [20]</td><td>N/S: Xinjiang Medical University; Size: 3012 images</td><td>convolutional neural network (CNN)</td><td>Accuracy = 93.33%</td></tr>
<tr><td>J. Hyeon et al. (2017) [21]</td><td>N/S: Bethesda system (BS)
Size: 71,344 microscopic images</td><td>VGGNet-16 CNN model and Support vector machine, random forest, logistic regression, AdaBoost</td><td>F-score value of 78%.</td></tr>
<tr><td>Makris etal. (2017) [22]</td><td>N/S: "Attikon" Hospital, Medical School;
Size: 416 liquid based histologically images</td><td>ANN-MPL</td><td>Accuracy = 95.91%,
Specificity = 93..44%,
Sensitivity = 99.42%</td></tr>
<tr><td>G.Forslid et al.(2017) [23]</td><td>N/S: CerviSCAN private and Herlev University datasets
Size: CerviSCAN dataset (9809 normal and 2234 abnormal images); Herlev dataset: 917 images, size of 100×100 pixels</td><td>VGG and ResNet CNN architecture</td><td>For VGGNet CNN: Accuracy = 82% %, For ResNet CNN: Accuracy = 83% %,</td></tr>
<tr><td>B.Sharma et al. (2016) [26]</td><td>N/S: Herlev University (Pap smear images) dataset; Size: 917 images (256×256 pixels)</td><td>Shape based features extracted using Back Propagation Neural Network.</td><td>Precision = 86%;
Recall = 90%; ZSI = 85%.</td></tr>
<tr><td>Y. Song et al. (2015) [27]</td><td>N/S: Private dataset (Hospital of Shenzhen);
Size: 21 cervical cytology images (1024×1360 pixels)</td><td>MSCN and graph partitioning</td><td>Accuracy= 90% for MSCN and 85% for graph partitioning-based method</td></tr>
</table>

</td></tr>
<tr><td>Accuracy Percentage</td><td>Most of the existing deep learning techniques acquire accuracy of 97.89% which is still low, on open source Pap smear (Herlev dataset) images. The reviewed accuracy can be improved with the help of meta-heuristics techniques using a hybrid of pre-trained CNN models.</td></tr>
<tr><td>Advantages</td><td>CNN-
Parameter Sharing -reduces the number of parameters and allowing for faster training and better generalization.
Hierarchical Learning: CNNs learn features hierarchically, meaning that lower-level features such as edges and corners are learned first, followed by more complex features such as object parts and textures. Efficient Memory Usage.</td></tr>
</table>

	SVM- Applicable where number of features is larger than the number of samples, such as in image recognition tasks. Less affected by the presence of outliers. Works well with both linear and non-linear problems. Fast prediction. Random Forest (RF)- High accuracy, Robustness Linear regression Linear regression is a simple and straightforward, can be applied to a wide range of problems, can produce accurate predictions when the underlying assumptions are met. Easy to implement.
Limitations	CNN- CNNs require large amounts of labeled data for training, Computationally Expensive, overfitting-meaning that they may perform well on the training data but poorly on new, unseen data. SVM- Computationally expensive- requires significant amt of time and memory to train, especially for large datasets. Random Forest (RF) Computationally expensive. Interpretability. Not suitable for small datasets. Linear regression - assumes that the relationship between the dependent and independent variables is linear, can overfit the data if the model is too complex or if there are too few data points. assumes that the independent variables are not highly correlated with each other
Sl No. 5	**Diagnosing Cervical Cancer Using Machine Learning Methods**
Author	1-Derya Yeliz Coşar Soğukkuyu, Oğuz Ata
Year	2022
Methodology	

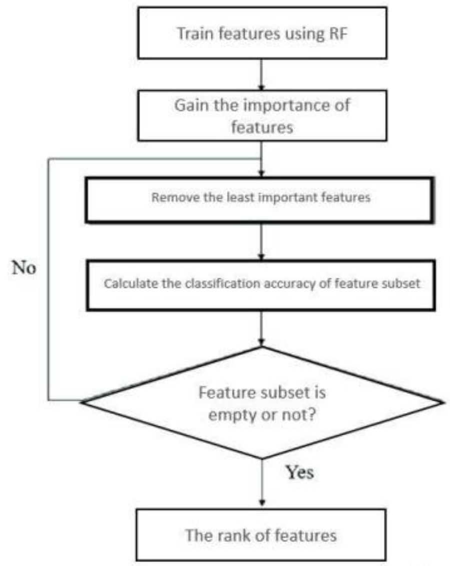


Figure 1:The recursive feature elimination (RFE) method's main procedure.

A. Dataset Description

- A dataset of cervical cancer risk variables collected at 'Hospital Universitario de Caracas' in Caracas, Venezuela has been used.
- The dataset contains 858 patients' demographics, behaviours, and medical records and consists of 36 indicators related to cervical cancer.
- There are numerous missing entries in the data, and unnecessary columns have been dropped from the dataset.
- The data imbalance has been handled by applying the Oversampling Technique - Smote.

B. Decision Tree Machine Learning Algorithms

- Decision Trees are a form of Supervised Machine Learning that separates data based on parameter.
- Decision nodes and leaves are two main parameters that form the tree.

C. Random Forest Machine Learning Algorithms

- Random Forest is a machine learning algorithm that uses the supervised learning method and can be used for both classification and regression issues.
- It is based on ensemble learning, which integrates several classifiers to increase the model's performance.

D. Model Optimization

- Recall score is important since predicting a cancer patient as healthy is extremely risky and can negatively affect the patient's life.
- Imbalance of the target variable has been handled by applying the smote oversampling technique.
- Recursive Feature Elimination (RFE) algorithm has been used for feature selection.

E. Application and Results

- A decision system based on Artificial Intelligence was created for prediction of cervical cancer.
- Precision, recall, and F-score were used to assess the model's performance.
- Precision = TP / (TP + FP)
- Recall = TP / (TP + FN)

<table>
<tr><td></td><td>

- F1-Score = 2 * Precision * Recall / (Precision + Recall)

</td></tr>
<tr><td>Model Used</td><td>Composition of the C4.5 and Logistic Regression,
Support Vector Machine (SVM),
Convolutional Neural Network (CNN),
Random Forest (RF),
an ensemble-based approach</td></tr>
<tr><td>Data sets used</td><td>

Table 1 Overview of Prediction of Cervical Cancer using machine learning models

<table>
<tr><th>Author & Year</th><th>Method</th><th>Dataset</th><th>Result</th></tr>
<tr><td>Su et al. (2016) [5]</td><td>Composition of the C4.5 and Logistic Regression</td><td>pap-smear images</td><td>95.642%</td></tr>
<tr><td>Ashok et al. (2016) [6]</td><td>Support Vector Machine (SVM)</td><td>pap-smear images</td><td>98.5%</td></tr>
<tr><td>Wang et al. (2019) [7]</td><td>Support Vector Machine (SVM)</td><td>pap-smear images</td><td>96%</td></tr>
<tr><td>Ghoneim et al. (2019) [8]</td><td>Convolutional Neural Network (CNN)</td><td>pap-smear images</td><td>99.7%</td></tr>
<tr><td>Adem et al. (2019) [9]</td><td>stacked autoencoder</td><td>a data set containing 668 samples, 30 attributes and 4 target variables (Schiller, Citology, Biopsy and Hinselmann) from the UCI database</td><td>97.25%</td></tr>
<tr><td>Geetha et al. (2019) [10]</td><td>Random Forest (RF)</td><td>a data set containing 668 samples, 30 attributes and 4 target variables (Schiller, Citology, Biopsy and Hinselmann) from the UCI database</td><td>-</td></tr>
<tr><td>Ijaz et al. (2019)[11]</td><td>Random Forest (RF)</td><td>the repository of UCI collected at Hospital Universitario de Caracas in Caracas, Venezuela 858 instances with 36 features.</td><td>97.02%</td></tr>
<tr><td>Lu et al. [12]</td><td>an ensemble-based approach</td><td>UCI risk factor dataset</td><td>83.16%</td></tr>
</table>

</td></tr>
<tr><td>Accuracy Percentage</td><td>According to model results accuracy performance is the same 94% for both decision tree and random forest.
According to model after applying RFE feature selection technique results, maximum performance achieved with Decision Tree with accuracy 97% even random forest had accuracy score 98%. When all three metrics are considered, the results show that for Recall measurement, the Decision Tree algorithm performs better than other algorithms, with a score of 95% whereas the random forest reached 90%.</td></tr>
<tr><td>Advantages</td><td>Composition of the C4.5 and Logistic Regression
C4.5 can handle complex datasets with many features, while Logistic Regression can handle binary classification problems very well. By combining these two approaches, we can potentially get better predictive performance than using either algorithm alone.

Support Vector Machine (SVM),
Applicable where number of features is larger than the number of samples, such as in image recognition tasks.
Less affected by the presence of outliers. Works well with both linear and non-linear problems.
Fast prediction.

CNN-
Parameter Sharing -reduces the number of parameters and allowing for faster training and better generalization.</td></tr>
</table>

	Hierarchical Learning: CNNs learn features hierarchically, meaning that lower-level features such as edges and corners are learned first, followed by more complex features such as object parts and textures. Efficient Memory Usage. Random Forest (RF)- High accuracy, Robustness
Limitations	Composition of the C4.5 and Logistic Regression Decision trees, like C4.5, are prone to overfitting when the tree becomes too large, Logistic Regression is a linear model and may not be suitable for datasets that have complex nonlinear relationships. When combining, finding the right balance between the two algorithms can be difficult and require some experimentation to get right. Support Vector Machine (SVM) Computationally expensive- requires significant amt of time and memory to train, especially for large datasets. CNN- CNNs require large amounts of labeled data for training, Computationally Expensive, overfitting-meaning that they may perform well on the training data but poorly on new, unseen data. Random Forest (RF) Computationally expensive Interpretability Not suitable for small datasets