International Journal of Data Informatics and Intelligent Computing (IJDIIC)
Vol. 2, No. 3, September 2023, pp. 11~20
DOI: 10.59461/ijdiic.v2i3.72

Breast Cancer Forecasting Using Machine Learning Algorithms

A.Thilaka[1], E.Sundaravalli[1]
[1]Department of Computer Science, New Prince Shri Bhavani Arts and Science College, Chennai, Tamilnadu, India.

Article Info

Article history:

Received July 07, 2023
Revised August 05, 2023
Accepted August 18, 2023

Keywords:

Breast cancer classification
Breast cancer prediction
Benign
Malignant
Support Vector Machine
Random Forest

ABSTRACT

One of the most prevalent and significant causes of malignancies in women is breast cancer. It is presently a widespread health problem, and it has recently become more frequent. The greatest method for managing breast cancer symptoms is early identification. The only kind of cancer that primarily affects women globally is breast cancer, which has the potential to be a major cause of mortality. Early detection of breast cancer is crucial in order to properly treat it and save many lives. This paper covers the results and analyses of several machine-learning algorithms for identifying breast cancer. Several machine learning models used the information once it was analyzed. In this paper, the random forests and SVC algorithms were applied, and the performance of these algorithms was compared. The dataset was taken from the UCI repository. Analyze and compare the classifiers' performance in terms of accuracy, precision, and f1-Score. For implementing the ML algorithms, the dataset was split among the training and testing phases. The notebook application Jupyter was used to implement these models. When compared to the other two models, it was successfully proven that the SVC model offers the best results. SVC's accuracy of 93% is greater than the method described earlier in that regard. The method used by this model, which will classify cancer into benign and malignant categories, yields the best results.

Corresponding Author:

E. Sundaravalli
Department of Computer Science
New Prince Shri Bhavani Arts and Science College
Chennai
Tamilnadu, India.

Email: sudarp2627@gmail.com

1. INTRODUCTION

It is the most well-known form of cancer across the world, along with lung and bronchus cancer, cancer of the breast and prostate, malignancy of the colon, and malignancy of the pancreas. The cancer of the breast may be an ordinary cause of fatalities and is the most common form of malignant tumour affecting women generally[1]. The study [2] of genetics, obesity, hormonal substances, treatment with radiation, and even reproductive health conditions are among the many of the many possible causes of breast cancer. Breast carcinoma is one of the most prevalent medical conditions among individuals, with an estimated 2.5 million new occurrences being identified every year, according to the International Health Organization's report on the year 2023. In [3] ninety-five of the nations in the entire world, carcinoma of the breast is the throughout the process or second factor in cancer-related fatalities for women. The rates of recurrence for breast tumours vary significantly between and across nation-states, nevertheless. In nations with low to middle incomes, both cervical and breast cancer fatalities constitute more than eighty per cent of all cases.

The proper diagnosis and classification of carcinoma of the breast into both malignant and benign classifications is now an established field of investigation[4]. Machine learning (ML) techniques [5] have been frequently employed to recognize cancers in women and extrapolate additional ideas from information trends in the course of the twentieth century. Machine learning is commonly employed to classify breast tumours[6].

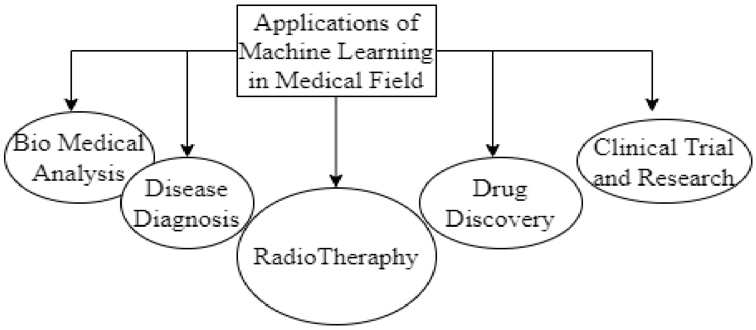

Figure 1. Displays the most common applications of machine learning in the medical industry.

Screening radiography can be performed on individuals who have low-risk groups of carcinoma symptoms in order to identify unconventional cancer of the breast tissue. Since it involves the assessment of possible cancerous regions [7,8], this approach is inadequate unless there are several images. As reported by an investigation [9], almost fifty per cent of carcinomas of the breast have been overlooked in examinations of women, particularly in the thickness of tissue in their breasts.

A computational technique that has been increasingly gaining popularity for identifying numerous medical conditions is a form of machine learning [10]. It includes the process of extracting knowledge from information and discovering hidden coincidences. Whereas other investigations employed mammographic images or individual biopsy information to construct their forecasts, certain studies [11] simply utilized demographic risk indicators (lifestyle and scientific data) for forecasting breast cancer. Others demonstrated how genetic information could be employed to predict the development of breast cancer. The establishment of an approach that takes into account all recognized risk factors is an essential challenge in predicting the likelihood of cancer of the breast [12].

2. LITERATURE REVIEW

In accordance with the research conducted, the single most widespread carcinoma among Indian women is breast carcinoma, exhibiting an age-adjusted number of cases of 25.7 in every 100,000 women with an overall mortality rate of 12.8 every 110,000 women. For the purpose of evaluating occurrence and death rates, an investigation encompassing several of the latest national cancer surveillance systems contrasted information reporting. Specific geographical regions distinguished notably due to their age-adjusted prevalence levels of carcinoma of the breast, with Delhi, India, holding the greatest incidence rate at 40 in 110,000 women, followed by Chennai (37.9), Bengaluru (34.4), then Trivandrum Region (33.7). Many Population-Based Cancer Registration Systems (PBCRs), such as Bengaluru (annual proportion modification: 2.84%), Barshi (1.87%), Chennai (2.00%), Delhi (2.44%), Mumbai (1.44%), and Bhopal (1.42%), possessed a significant increase in the age-adjusted rate during 1983 and 2015. The number of fatalities to occurrence proportion numerous reaching substantial in countryside records at 68 and considerably lower in those in cities at 8. Furthermore, especially for Indian women, growing younger has grown into an important risk indicator of breast cancer. According to estimates, there could be a substantial rise in the total number of scenarios of breast cancer diagnosed throughout the nation by the end of the decade, with an estimated maximum of around 1,798,900 cases. An enhanced medical atmosphere could originate through improved awareness of health concerns in combination with the accessibility of early detection of breast cancer detection programmes as well as effective treatment programmes across the nation [13].

Conducted at Bengaluru's ICMR National Centre for Disease Informatics and Research, the study utilized data from the National Cancer Registry Programme (NCRP) spanning 2020 to 2022. Its objective was to assess India's cancer burden in 2025. According to the findings, it is anticipated that one out of every nine Indians will likely develop cancer during their lifespan, reaching this point around the age of 74.

Carcinoma of the breast, which had the greatest overall prevalence of every one of the common malignancies [14] among women, was promptly followed by malignancies of the cervical region, ovary, as well as corpus uteri. Lung, oral tradition, and tongue malignancies remained widespread among men as well as demonstrated significant correlations between cigarette smoking as well as additional tobacco-related behaviours.

A research assistant educator in the Laboratory of Cancer Radiation Oncology at the All India University of Medical Sciences in Patna, Punjab, divulged that demonstrating data recommended that there would be approximately two million cases of breast cancer occur by the year 2030, regardless of the reality that the paper failed to provide specific figures for the number of cases of the breasts, the cervical cavity, ovary, and uterine tumours in India. In the Indian subcontinent, there are at present, 183,000 scenarios of breast cancer have been reported.

From the survey of WHO, Breast cancer is a condition characterized by the uncontrolled growth of abnormal cells in the breast, forming tumours that can become life-threatening if not treated. These cancerous cells typically originate in the breast's milk ducts or lobules, and the earliest form, known as in situ, is generally not lethal. However, as cancer progresses, it can invade nearby breast tissue, leading to the development of lumps or thickening [15,16].

Invasive breast cancers have the potential to spread to nearby lymph nodes or other organs, a process called metastasis, which can be fatal. The treatment approach [17] for breast cancer depends on various factors, including the individual's health, the type of cancer, and its extent. Common treatments include surgery, radiation therapy, and medications.

Here are some key points about breast cancer:

- In 2020, breast cancer claimed the lives of 685,000 people worldwide.
- Approximately half of all breast cancer cases occur in women without specific risk factors; other
- than being female and ageing.
- Breast cancer can affect women in every country across the globe.
- While breast cancer is more common in women, around 0.5% to 1% of cases occur in men.

By the conclusion of the year in question, 7.9 million women were newly confirmed to have carcinoma of the breast around the world, thus becoming one of the most prevalent illnesses in the world as a whole. In the year 2020, there were approximately 2.4 million novel cases of mammary tumours discovered in women, making it the cancer with the greatest incidence in the entire world. [10] Women are susceptible to cancer of the breast at any develops after their teenage years. Nevertheless, the number of hazards tends to increase with age. Deaths from breast cancer were very low and stable from the 1931s to the 1971s when the primary treatment was surgery (radical mastectomy). Significant improvements in survival rates started to emerge in the 1990s, coinciding with the establishment of breast cancer early detection programs in various countries. These programs were closely linked to comprehensive treatment approaches that included effective medical therapies.

Machine learning is a modelling technique used to extract knowledge from data and uncover hidden relationships, and it has gained significant traction in the field of healthcare, particularly in predicting various diseases [18]. In recent years, there have been studies focused on predicting breast cancer using different approaches. Some of these studies have exclusively relied on demographic risk factors such as lifestyle and laboratory data to make predictions, while others have used mammographic patterns or patient biopsy data [19]. Additionally, genetic information has been applied in certain cases to predict breast cancer risk.

The establishment of a comprehensive framework that incorporates into consideration all recognized factors that increase risk is one of the main problems in cancer prevention and prediction [20]. Current forecasting techniques typically involve demography factors for risk or the examination of mammographic images, possibly disregarding additional significant variables. The administration of frequently received examinations and disruptive treatments like MRI, which stands for magnetic resonance imaging and ultrasonography, could come from these types of models, regardless of whether treatments have been successful at identifying people at increased risk, weighing down individuals both economically and emotionally. Effective breast cancer risk prediction requires the inclusion of various factors, including demographic data, laboratory results, and mammographic risk factors [21]. Therefore, models that incorporate multiple risk factors in their analysis can offer more precise assessments of breast cancer risk. This current study aims to predict breast cancer using diverse machine-learning approaches while taking into account a wide range of factors in the modelling process.

3. METHODOLOGY

This section encompasses all approaches and resources used, as well as the description of the dataset, block diagrams, flowcharts, and evaluation grids.

3.1. Dataset

The investigation was conducted using the Wisconsin Breast Cancer Diagnostic (WBCD) data source. The dataset was taken from a well-known UCI repository. The repository for machine learning has a simplified size of 569 samples and contains 32 features). From this data source, we only took 500 samples and five major features to prepare the model. The atomic features of one needle aspirations (FNAs) taken from patients' exhibited breasts make up the example dataset. A tiny needle is placed into a body Kuid or tissue that appears aberrant to receive an instance for the purpose of determining the cause or forecast illness of tumour tissue. There are no errors in Figure 2, which shows the total number of malignant and benign data from the dataset.

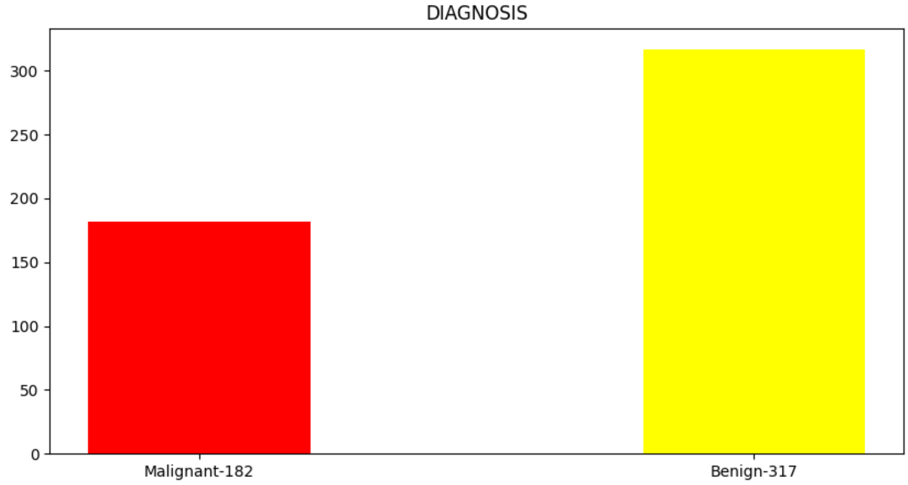

Figure 2. Total number of cancerous and normal cells

In this study, we employ five basic characteristics to forecast the development of breast cancer. symmetry_mean, area_mean, perimeter_mean, Radius_mean, and concave points_mean are the most significant parameters. Finding the median for each of these features makes it possible to forecast the likelihood of the presence or absence of malignant tumours in an illustration with simplicity. The five features we selected for the present project are displayed in Figure 3.

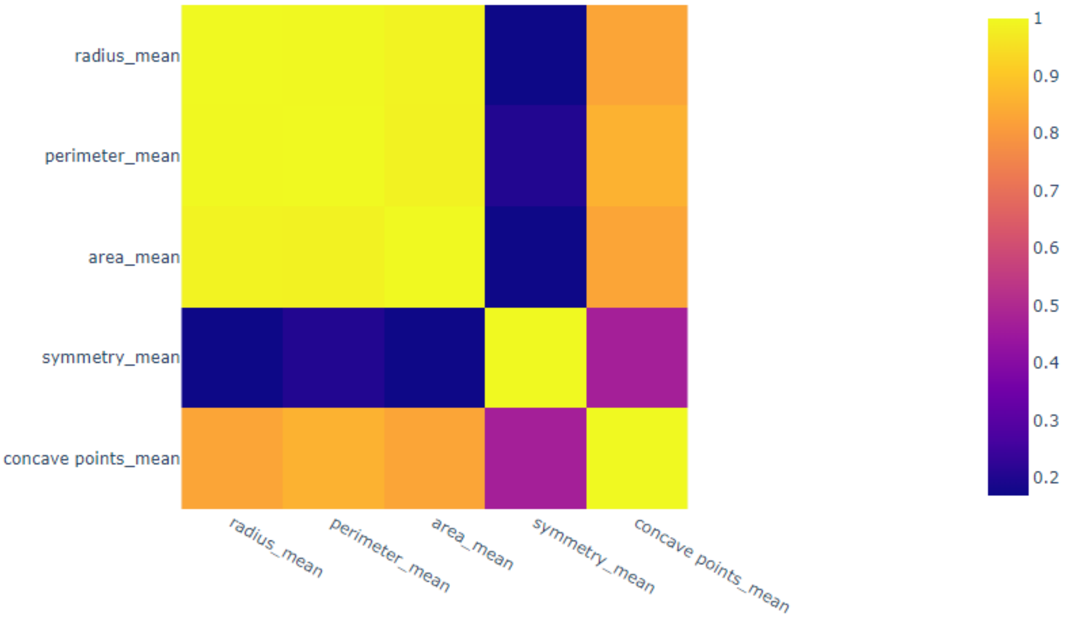

Figure 3. Correlation diagram for the proposed system

3.2. Classifier using random forests

As a non-parametric technique, Random Forest (RF) employs a classification methodology. With each dataset, RF rapidly carries out categorization and employs numerous decision trees. Within each tree, an assorted set of input variables is present, and subsequently, all these trees are amalgamated to enhance inference drawn from the variables.

3.3. SVC Classifier

In the realm of machine learning, support vector machines operate as supervised models. These machines establish a plane with a multi-dimensional classification of objects. This multi-dimension plane, essentially a line within a plane, serves to differentiate between the two classes. An SVC training approach generates a classification model employing a number of experiment specimens that were previously categorized into one of the two distinct groups. This model becomes the deterministic multimodal linear classification system as soon as it assigns fresh instances to a single of the two groups. While SVCs are inherently non-probabilistic, techniques like Platt scaling can be employed to apply SVC within a probabilistic classification context. These novel examples are then projected into the same feature space, and based on which side of the divide they land on, their classification is anticipated.

4. PROPOSED METHODOLOGY

Employing a set of data taken from the state of Wisconsin information sets, this suggested approach evaluates the outcomes of a number of machine learning techniques, including SVC (Support Vector Classifier) and random forest modelling (RF). The information set has been divided into testing and training sets in order to use different machine learning methods, which enabled an in-depth assessment of every one of the algorithms. The approach that generates the most favourable outcomes will be employed as an accurate prediction model when predicting breast cancer in the future. The dataset in question consists of 5 multimodal practical problem characteristics, which can be obtained from the UCI Machine Learning Data Portal. Only 498 points of information from an aggregate of 571 circumstances were used for developing the model that we used. The suggested system can be observed in the image below.

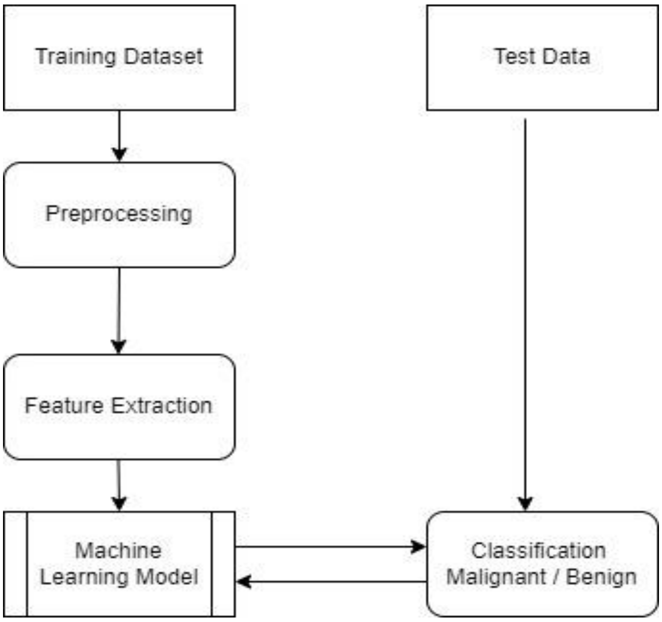

Figure 4. Histogram of the proposed system

Here, we are providing a simplified outline of the key steps occupied in the Random Forest algorithm:

Data Preparation:

Input dataset

Split into training and testing datasets

Random Sampling (Bootstrap) selection:

We can Randomly select subsets of data for each and every tree

Growing Decision Trees:

For each subset of data and features, we can create the following steps:

Build a decision tree (binary tree)
Split all the nodes based on the best feature
Repeat until a stopping criterion is met

Ensemble Creation:

Next, we Combine the outcome from individual trees
Based on the outcome of classification, we are making use of majority voting

Prediction:

Use the ensemble to make predictions on new data

Evaluation:

Assess the performance of the Random Forest on the testing dataset

Repeat:

Repeat steps 2-7 for a specified number of times (n_estimators)

Final Prediction:

Before getting the output, merge the predictions from all trees for the final result

Output:

Finally, we find the final prediction or classification results

A table of information labelled a matrix of confusion is frequently employed to prove how effectively a model of classification performs when used with an array of test information whose true values have been established. In this section, we represent ourselves as the matrix of uncertainty corresponding to the randomly generated forest approach for the information we have within the graph.

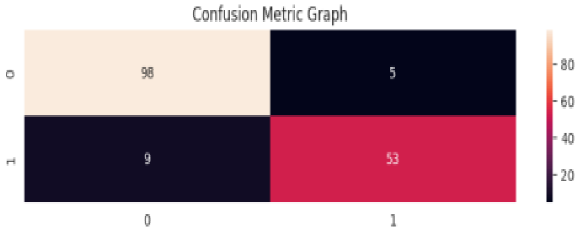

Figure 5. Confusion matrix for random forest

For the Random Forest approach, the suggested system correctly predicted 98(TP) benign and 53(TN) malignant. It mislabeled benign 5(FP) as malignant and malignant 9(FN) as benign. Out of 165 datasets, this method predicted 151 records.

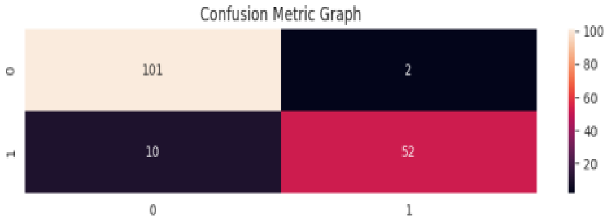

Figure6. Confusion matrix for SVC

SVC classified 153 records as TP and TN; it misclassified 12 (FP and FN) records. This model accurately classified 93% of the records. The accuracy of SVC is higher since the misclassification is lower than in the Random Forest model.

5. RESULTS AND DISCUSSION

To compare the performance of these two algorithms, we have to use some parameters to calculate their performance. The Performance metrics are

- Accuracy
- Precision
- F1-Score
- Recall

5.1. Accuracy

The effectiveness of an algorithm for classification, whenever utilized with an array of test outcomes whose accurate values have been established, is frequently displayed in a matrix of confusion, which constitutes a tabular. In this section, individuals represent the data set for the random forest technique in the picture that was provided as a confusion matrix. As a result, it may not be the ideal method for comparing various classifiers, but it can perform a general overview of classification performance. Therefore, accuracy can be calculated with the below-given equation:

Accuracy = (TP + TN) / (TP + FP + FN + TN)

In the above equation:

- True Positives (TP) are situations that were accurately categorized as positive.
- True Negatives (TN) are situations that were appropriately categorized as negative.
- False Positives (FP) are situations that are mislabelled as positive.
- False Negatives (FN) are situations that are wrongly labelled as negative.

5.2. Precision

Precision, often called confidence, measures the proportion of correctly identified true positives out of all instances classified as positives. It indicates the classifier's effectiveness in handling positive observations but doesn't provide substantial insights into its performance regarding negative ones. Precision can be calculated as:

Precision = TP / TP + FP.

5.3. F1-score

The most important metric is the f1-score to compare the prediction power of those algorithms. The weighted harmonic mean of Precision and Recall is the F1-Score. This score consequently considers both false positive and false negative results. F1-Score is equal to two times the product of recall and precision. We can calculate the f1-score by using the below equation,

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

5.4. Recall

The percentage of actual events that are beneficial that a model for prediction effectively identifies into the positive category is referred to as recall, which is additionally known as sensitivity. Here a valuable metric, particularly in fields like healthcare, as it tells us how well the model detects the actual positive cases. In this context, recall is essentially about the accuracy of positive predictions. The equation is,

Recall = TP / (TP+FN)

The below images will show the comparison results of all the metrics for the two algorithms.

Table 1. Classification report of random forest

Features	**Recall**	**Precision**	**F1-score**
Malignant	0.85	0.91	0.88
Benign	0.95	0.92	0.93

Table 2. Classification report of SVC

Features	**Recall**	**Precision**	**F1-score**
Malignant	0.84	0.96	0.90
Benign	0.91	0.91	0.94

Table 3. The comparison result of the proposed algorithms

Algorithms	**Accuracy**
SVC	93%
RF	92%

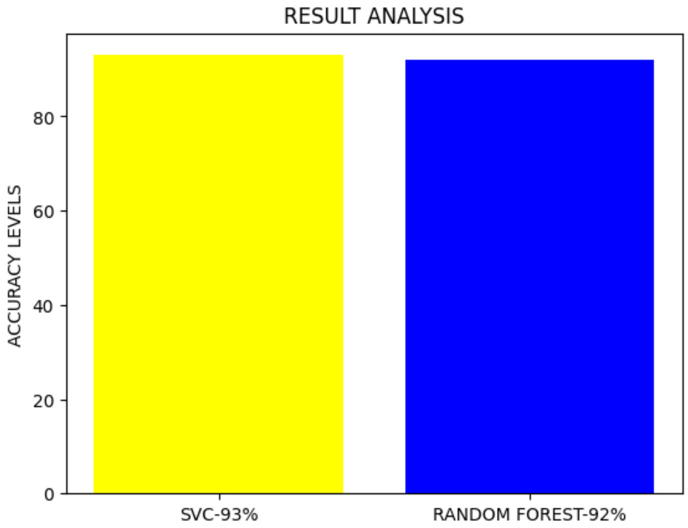

Figure 9. The comparison result of SVC and Random Forest

In the proposed system, the Support Vector Classifier (SVC) achieved an accuracy rate of 93%, strongly signifying the capability to correctly catalogue data points in a specified dataset. This high accuracy suggests that SVC is more effective in distinguishing between altered classes or categories in the data. The Random Forest model achieved an accuracy of 92%, exposing its potential to make clear-cut predictions. Random Forest is an ensemble investigating method that fuses multiple decision trees, making it robust and suitable for a wide range of tasks. Its 92% accuracy signifies its effectiveness in handling complex datasets. These accuracy percentages reflect the performance of the two models in terms of their ability to make correct predictions, with the Support Vector Classifier having a slightly higher accuracy than the Random Forest in this comparison. The table below will portray the accuracy level of the two algorithms.

6. CONCLUSION

Nearly every nation on every continent suffers from cancer of the breast. Every year, the count of breast cancer diagnoses was increasing rapidly irrespective of the risk factors of age and gender. In this work, nearly 500 records were taken with various attributes. The cancer of the breast is capable of being treated when it is identified earlier. The present investigation evaluates techniques using machine learning to recognize cancer of the breast with specific features. The basic principles associated with the mammary are outlined through these distinctive features. To anticipate the disease, predictive machine learning algorithms employing randomly generated forests and SVC were developed. By comparing the accuracy and F1-score of these two models, SVC gives the highest accuracy and F1-score. This SVC model will diagnose breast cancer through the basic features of the breast with maximum accuracy. Breast cancer awareness should be taken to all people through education and social media because breast cancer detection at the early stage is much more important.

7. FUTURE ENHANCEMENT

Nowadays to diagnose the breast cancervariousmachine learning algorithms are used to build the machine learning model. In future we will take the important features of the malignant cells based on the experts (Doctors) suggestion. For that dataset we can apply Ensemble methods to get higher accuracy in Prediction model.

REFERENCES

[1] S. Malvia, S. A. Bagadi, U. S. Dubey, and S. Saxena, “Epidemiology of breast cancer in Indian women,” *Asia. Pac. J. Clin. Oncol.*, vol. 13, no. 4, pp. 289–295, Aug. 2017, doi: 10.1111/ajco.12661.
[2] Kumar Mohan, Yousuf Fadhil Salim AL-Mamari, and Mohammed Ahmed Mohammed AL-Najadi, “Techniques

of Machine Learning for Detecting Heart Failure," *Int. J. Data Informatics Intell. Comput.*, vol. 2, no. 2, pp. 47–54, Jun. 2023, doi: 10.59461/ijdiic.v2i2.62.

[3] S. S. C. Bose, R. Natarajan, G. H L, F. Flammini, and P. V. Praveen Sundar, "Iterative Reflect Perceptual Sammon and Machine Learning-Based Bagging Classification for Efficient Tumor Detection," *Sustainability*, vol. 15, no. 5, p. 4602, Mar. 2023, doi: 10.3390/su15054602.

[4] Z. Guan *et al.*, "Combining Breast Cancer Risk Prediction Models," *Cancers (Basel).*, vol. 15, no. 4, p. 1090, Feb. 2023, doi: 10.3390/cancers15041090.

[5] M. U. Ghani, T. M. Alam, and F. H. Jaskani, "Comparison of Classification Models for Early Prediction of Breast Cancer," in *2019 International Conference on Innovative Computing (ICIC)*, IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/ICIC48496.2019.8966691.

[6] A. K. Shukla and V. Suresh Kumar, "Cloud Computing with Artificial Intelligence Techniques for Effective Disease Detection," *Int. J. Data Informatics Intell. Comput.*, vol. 2, no. 1, pp. 32–41, Mar. 2023, doi: 10.59461/ijdiic.v2i1.45.

[7] S. Subash Chandra Bose, A. Vinoth Kumar, A. Premkumar, M. Deepika, and M. Gokilavani, "Biserial targeted feature projection based radial kernel regressive deep belief neural learning for covid-19 prediction," *Soft Comput.*, vol. 27, no. 3, pp. 1651–1662, Feb. 2023, doi: 10.1007/s00500-022-06943-x.

[8] A. M. McCarthy *et al.*, "The use of the Gail model, body mass index and SNPs to predict breast cancer among women with abnormal (BI-RADS 4) mammograms," *Breast Cancer Res.*, vol. 17, no. 1, p. 1, Dec. 2015, doi: 10.1186/s13058-014-0509-4.

[9] W. L. Bi *et al.*, "Artificial intelligence in cancer imaging: Clinical challenges and applications," *CA. Cancer J. Clin.*, vol. 69, no. 2, pp. 127–157, Mar. 2019, doi: 10.3322/caac.21552.

[10] S. Babichev, L. Yasinska-Damri, and I. Liakh, "A Hybrid Model of Cancer Diseases Diagnosis Based on Gene Expression Data with Joint Use of Data Mining Methods and Machine Learning Techniques," *Appl. Sci.*, vol. 13, no. 10, p. 6022, May 2023, doi: 10.3390/app13106022.

[11] A. F. M. Agarap, "On breast cancer detection," in *Proceedings of the 2nd International Conference on Machine Learning and Soft Computing*, New York, NY, USA: ACM, Feb. 2018, pp. 5–9. doi: 10.1145/3184066.3184080.

[12] W. Yue, Z. Wang, H. Chen, A. Payne, and X. Liu, "Machine Learning with Applications in Breast Cancer Diagnosis and Prognosis," *Designs*, vol. 2, no. 2, p. 13, May 2018, doi: 10.3390/designs2020013.

[13] M. A. Mohammed *et al.*, "RETRACTED ARTICLE: Decision support system for nasopharyngeal carcinoma discrimination from endoscopic images using artificial neural network," *J. Supercomput.*, vol. 76, no. 2, pp. 1086–1104, Feb. 2020, doi: 10.1007/s11227-018-2587-z.

[14] Ashish Kumar Pandey and Prabhdeep Singh, "A Systematic Survey of Classification Algorithms for Cancer Detection," *Int. J. Data Informatics Intell. Comput.*, vol. 1, no. 2, pp. 34–50, Dec. 2022, doi: 10.59461/ijdiic.v1i2.32.

[15] M. A. Naji, S. El Filali, K. Aarika, E. H. Benlahmar, R. A. Abdelouhahid, and O. Debauche, "Machine Learning Algorithms For Breast Cancer Prediction And Diagnosis," *Procedia Comput. Sci.*, vol. 191, pp. 487–492, 2021, doi: 10.1016/j.procs.2021.07.062.

[16] R. Rabiei, "Prediction of Breast Cancer using Machine Learning Approaches," *J. Biomed. Phys. Eng.*, vol. 12, no. 3, Jul. 2022, doi: 10.31661/jbpe.v0i0.2109-1403.

[17] M. Ferdous, J. Debnath, and N. R. Chakraborty, "Machine Learning Algorithms in Healthcare: A Literature Survey," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, Jul. 2020, pp. 1–6. doi: 10.1109/ICCCNT49239.2020.9225642.

[18] S. S. Bose* and D. T. Christopher, "Aggregate Linear Discriminate Analyzed Feature Extraction and Ensemble of Bootstrap with Knn Classifier for Malicious Tumour Detection," *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 3696–3694, Sep. 2019, doi: 10.35940/ijrte.C4802.098319.

[19] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 6308–6325, 2020, doi: 10.1109/JSTARS.2020.3026724.

[20] N. Al-Azzam and I. Shatnawi, "Comparing supervised and semi-supervised Machine Learning Models on Diagnosing Breast Cancer," *Ann. Med. Surg.*, vol. 62, pp. 53–64, Feb. 2021, doi: 10.1016/j.amsu.2020.12.043.

[21] V. Nemade and V. Fegade, "Machine Learning Techniques for Breast Cancer Prediction," *Procedia Comput. Sci.*, vol. 218, pp. 1314–1320, 2023, doi: 10.1016/j.procs.2023.01.110.

BIOGRAPHIES OF AUTHORS

A. Thilaka is a dedicated educator with an impressive academic journey and five years of experience as an Assistant Professor at New Prince Shri Bhavani Arts and Science College. Her commitment to the field of education is underpinned by her extensive qualifications. Dr.A.Thilaka attained her Doctorate from Bharathidasan University, Trichy, in 2021, demonstrating her commitment to advancing knowledge in her field. Prior to her doctoral studies, she completed her M. Phil degree at the same institution in 2015, further honing her research skills. Driven by her passion for computer science, she earned her Master of Computer Science degree from Cauvery College of Arts and Science for Women in 2014 and her Bachelor of Computer Science degree in 2011, laying a strong academic foundation. Her diverse academic background and years of teaching experience make her a valuable asset to the academic community. Her interests in Research are Image Processing, Machine Learning, Big data Analysis, and Deep Learning. For inquiries or collaborations, Dr. A. Thilaka can be reached at email: thilakaart1991@gmail.com.

E. Sundaravalli is an Assistant Professor with 1.5 years of experience at New York Prince Shri Bavani Arts and Science College in Chennai. Her dedication to academia is evident through her recent qualification as an Assistant Professor in UGC-NET in 2021, a prestigious accomplishment that underscores her expertise in her field. Earning her Master of Computer Science degree at SDNB Vaishnav College for Women in 2018, she has a strong academic foundation. This was preceded by her Bachelor of Computer Science degree from Prof. Dhanapalan College of Arts and Science in 2016. Her combination of teaching experience and academic achievements makes her a valuable asset to the academic community. She can be reached at email: sudarp2627@gmail.com.