

Breast Cancer Detection

Soni Ahirrao¹, Vinanti Bhoer², Shivani Patil³ and Himani Jawale⁴

¹Student, Department of Information Technology, FCRIT, Vashi, Navi Mumbai, INDIA

²Student, Department of Information Technology, FCRIT, Vashi, Navi Mumbai, INDIA

³Student, Department of Information Technology, FCRIT, Vashi, Navi Mumbai, INDIA

⁴Professor, Department of Information Technology, FCRIT, Vashi, Navi Mumbai, INDIA

¹Corresponding Author: ahirraosoni@gmail.com

ABSTRACT

Breast Cancer is highly predominant in women in today's world. It starts in the breast during the initial stages and spreads to other areas of the body after some period of time. Breast cancer is the second-largest disease leading to the death of women. The disease is curable if detected early enough. Breast Cancer Application monitors the abnormal growth of breast cells during the early stages. They are often diagnosed during the advanced stages of breast cancer. It is the second most diagnosed cancer in women, affecting one in every eight women. Our project comprises two modules, first consists of an application with user login and self-test examine section where and the second section consists of identifying benign and malignant cells. The second section will be used by doctors' side for the detection of abnormalities of breasts as early as possible by providing the user screening data set. It contains Machine Learning techniques for the classification of malignant and benign tumors. There are more treatment options and a better chance of survival. If breast cancer is detected during the early stages then there is a 93 percent of higher survival rate in the first five years.

Keywords— Breast Cancer, Cancer Detection Application, Early Stage Detection, Java

are not much accurate, an alternative we are using Machine Learning techniques for the classification of benign and malignant tumours. The prior diagnosis of Mammary carcinoma can enhance the prediction and survival rate, so that patients can be informed to take clinical treatment at the right time. Classification of benign tumours can help the patients avoid undertaking needless treatments. The early Breast Cancer Detection and advancements in treatments have resulted in an increased survivorship. Breast Cancer is an abnormal, uncontrolled cell growth arising in the breast tissue and is the second leading cause of death in women. To minimize this, we are presenting our Breast Cancer Detection project. There are two main modules of our project, first is the application which provides maximum information related to breast cancer, guides with the self test examination and also provides a chat box section to the user so that he/ she can contact the doctors. Lastly comes the prediction of the result which is done at the hospital end to provide higher accuracy of the result in classification of malignant and benign cells.

I. INTRODUCTION

Nowadays, the early BC detection and advancements in treatments have resulted in an increased survivorship. Breast Cancer is an abnormal, uncontrolled cell growth arising in the breast tissue. Breast Cancer is the second leading cause of death in women. It primarily affects women but about 1 percent of all cases affect men. One third of women with Breast Cancer die from mammary cancerous cell growth. A breast is made up of three main parts: lobules, ducts, and connective tissue. The lobules are the glands that produce milk. The early Breast Cancer Detection and advancements in treatments have resulted in an increased survivorship.

The connective tissue (which consists of fibrous and fatty tissue) surrounds and holds everything together. When mammary carcinoma spreads to other parts of the body, it is said to have metastasized. As existing systems

II. LITERATURE SURVEY

The authors in [1] have addressed the problem of enhancement and denoising of mammography images. The algorithm which they have used is based on dyadic wavelet. Its adaptability of this method to different image analysis the same algorithm is used for both micro calcifications and mass detection. The processed images quality was improving and was considered by radiologists as significant for early detection of breast cancer. In [2] the authors focus on classifying the images automatically according to their maturity. It provides a repeatable and quantitative measure for studies. It works as an important measure to classify the images. Authors of [3] presents an initial investigation for mammary carcinoma detection using a special mode of bi static radar system known as forward scattering radar (FSR). [4] the objective of this paper was to find the performance of all the classification algorithms. All the algorithms which are used for analyzing mammary carcinoma data through analyzing the

mammography’s images based on their characteristics. Different attribute values of cancer affected mammogram images are considered for analysis in this work. The eating habits of patients, their age, their professional life, their personal lifestyle, all these different styles of attributes are taken into consideration for classification. [5] authors in this paper incorporated a model of 10-year post-treatment period for risk of developing CBC and/or dying of the primary cancer or CBC. For each patient profile, we used 100,000 micro simulation trials to estimate quality-adjusted life expectancy for the clinical strategies CPM and no CPM. Women who are suffering from first stage breast cancer and also with the first degree relative with bilateral breast can have QALY benefit from CPM. CPM is similar to that reported for BRCA1/2 mutation carriers. For most subgroups of women, CPM has a minimal to no effect on quality-adjusted life expectancy, irrespective of family history of breast cancer. [6] the authors in this paper help them to mimic the histopathologist while designing the system and increase the accuracy and reliability of the system. The authors in this paper covers all the information related to breast biopsy. It also tells about histological slide description, image processing techniques for automated histopathology analysis and breast cancer. [7] the authors in this paper analyzed the privacy policy of mHealth apps for breast cancer patients, developing a scale to check if GDPR is complied. Despite privacy being a key factor in the adoption of the use of mHealth apps, the low level of compliance with the GDPR of the analyzed applications was quite surprising. [8] authors in this paper embedded, automated health-care monitoring service. Proposed service is based on a multi-agent system, and can be used for mobile computing. In proposed a modified (i.e., two-way handshaking) iSCSI protocol applied for higher transmission rate. The proposed system compared with resource constrained devices, and conventional systems.

III. PROPOSED SYSTEM

The features included in the breast cancer mobile app development:

- Registration/ sign-in
- List of medical health information
- Self-Test Examination
- Real-time Chat Window

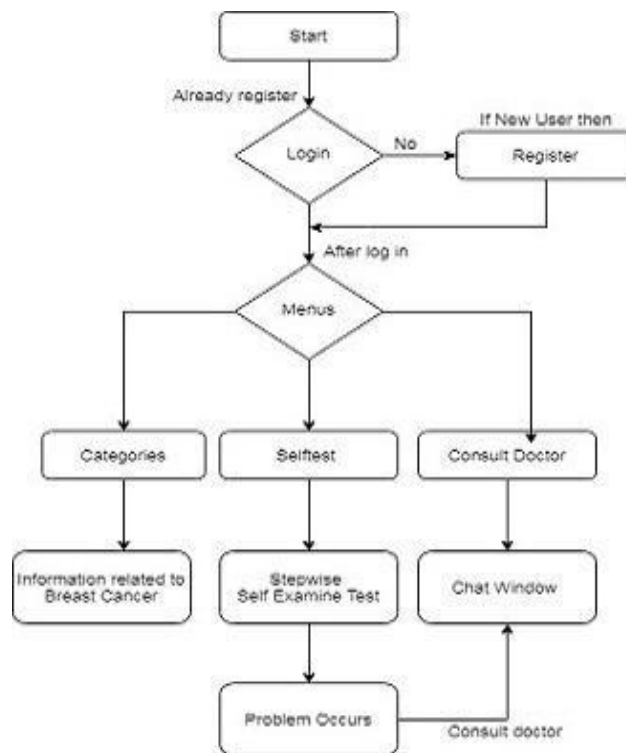


Figure 1: Flow Chart of Module 1

The flowchart in Fig. 1, shows the flow of our application the user must have an account to login into the application. If he doesn't have an account he will have to register first by providing all the credentials. After Login the user will be redirected to the menu page where he will get options like categories, self test module, consult to doctor. For application java and fire base is used for back-end.

The flowchart in Fig. 2, shows the steps included in the model. First the dataset can be collected from various sources such as a file, database, sensor and many other free sources such as kaggle, UCI Machine Learning Repository. Next step is data pre-processing. It is an important step for building machine learning models. It is the process of cleaning the raw data into clean data, so that can be used to train the model. There are various features based on which the data is divided into training and testing dataset. After classification the dataset is divided into benign and malignant cells. Model evaluation aims to estimate the generalization accuracy of machine learning models on future data.

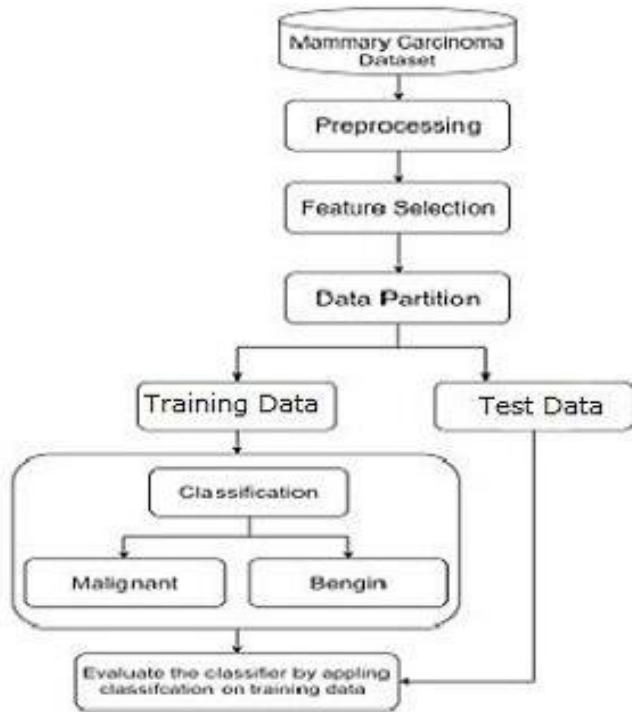


Figure 2: Flow Chart of Module 2

IV. SYSTEM DESIGN



Figure 3: Register Page

For the user interface, there is a register page in our application where the user can register himself and then login. The username and password which the user had given during the registrations is used for logging-in into the system. Fig 3, Fig 4 shows the snapshots of an application created.

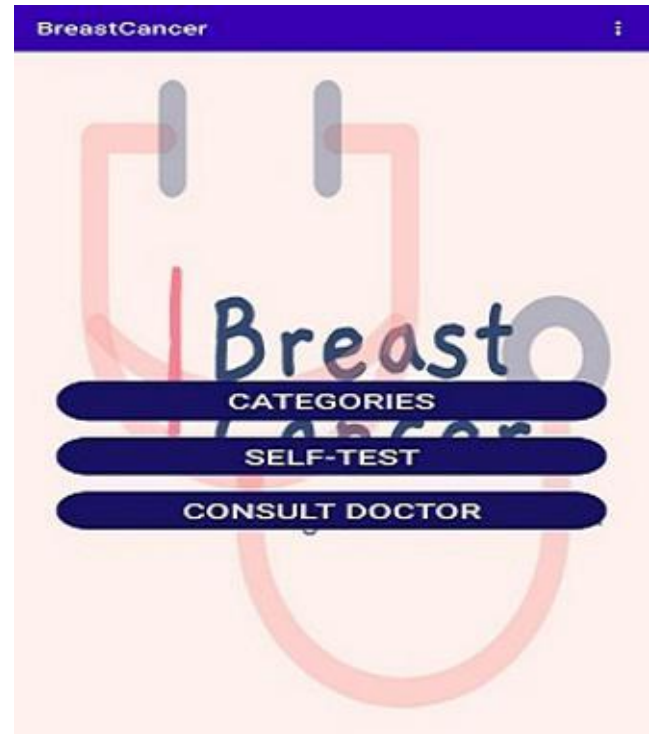


Figure 4: Main Page

After logging-in into the system the user comes to the main page of the application which consists of three sections that are categories, self-test and consult doctor. The categories consist of various menus which provides all the information related to breast cancer. It helps the user to know more about breast cancer and also the steps to take self-test examinations at their own pace. The third section consult doctor provides the user to personally connect with the doctor and explain the problems faced while self-test examination if any.

V. EXPERIMENTAL RESULTS

We have provided categories in the Breast Cancer Application which provides all information related to Breast Cancer, all risk factors, yoga exercises, diagnosis procedure, etc. The self examine section provides step wise procedure with images to carry out the test smoothly and efficiently. The client can carry out tests easily at their own pace. If any sort of uneven breasts surface or pain is identified during self-test then clients can connect with the doctor through application and consult with them.

Carrying out self-test helps us to identify problems earlier at these initial stages. Entire application is easy to understand and compelling. Dataset is divided into two subsets Training Dataset and Testing Dataset. We use a training data set to train our model so it can accurately predict our outcome. After classification Fig 5 shows the visualization count of benign and malignant cells.

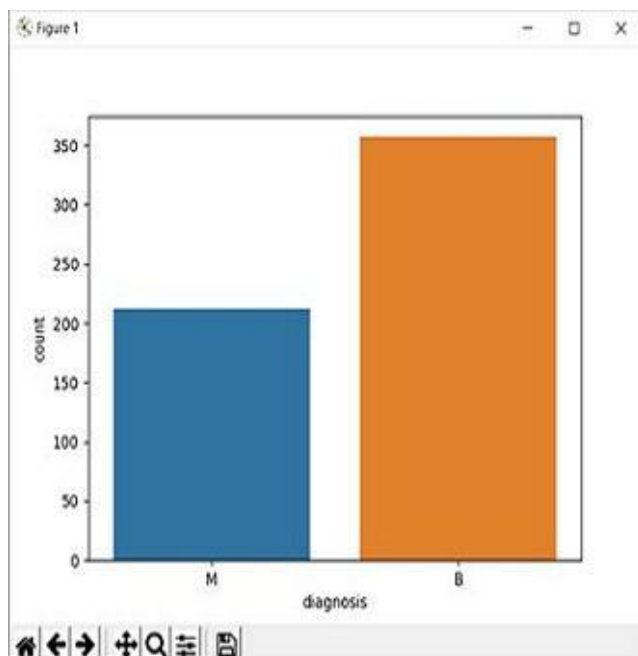


Figure 5: Visualization Count of Benign and Malignant Cells

VI. ADVANTAGES OF SYSTEM

This system provides handy information to users related to Breast Cancer to create awareness. The Self-test examine section helps users to get acknowledgement about abnormal growth of breasts. If any kind of issues faced by the user during Self-test he can concern with the doctor using a chat window. The stepwise procedure along with the images have been provided in the application to carry out self-test anytime, anywhere at your own pace. Module 2 of our project can be used by doctors to identify malignant cells and benign cells.

VII. CONCLUSION

An Early detection breast cancer application as our module one we have provided users with handy information related to breast cancer to create awareness and self examine themselves. In the second module to analyze the dataset we are using various data mining algorithms and also machine learning methods. An important challenge in the data mining and machine

learning area is to build accurate and computationally efficient on the Mammary Carcinoma (original) data set. We also have tried to compare effectiveness and efficiency of various algorithms in terms of accuracy, precision, recall, f1-score, support, macro average and weighted avg to find the best classification accuracy. According to our system, The Random Forest Classifier Accuracy has proved its accuracy around 96.5 percent.

REFERENCES

- [1] Arianna Mencattini, Marcello Salmeri, Roberto Lojaco, Manuela Frigerio, & Federica Caselli. (2008). *Mammographic images enhancement and denoising for mammary carcinoma detection using dyadic wavelet processing*. Available at: <https://ieeexplore.ieee.org/document/4454239>.
- [2] Tobias Christiian Cahoon, Melanie A., Suttorf James, & E. Bezdek. (2000). *Breast cancer detection using image processing techniques*. Available at: <https://ieeexplore.ieee.org/document/839171>.
- [3] Sumalatha G & Archana S. (2017). A study on early prevention and detection of mammary carcinoma using data mining techniques. *International Journal of Innovative Research in Computer and Communication Engineering*, 5(6), 11045-11050. DOI: 10.15680/IJIRCCE.2017.
- [4] Padmapriya B & Velmurugan T. (2016). Classification algorithm based analysis of mammary carcinoma data. *International Journal of Data Mining Techniques and Applications*, 5(1), 43-49.
- [5] A. K. Arrington, L. Goldstein, L. Kruper, C. Vito, J. Yim, & S.L. Chen. (2014). Life expectancy after curative-intent treatment of breast cancer: Impact on long-term follow-up care. *Am. Surg.*, 80(6), 604-609.
- [6] R. Sangeetha & K. S. Murthy. (2017). A novel approach for detection of breast cancer at an early stage using digital image processing techniques. *International Conference on Inventive Systems and Control (ICISC), Coimbatore*, pp. 1-4. DOI: 10.1109/ICISC.2017.8068625.
- [7] J. Benjumea, E. Dorrnzoro, J. Roper, O. Rivera-Romero, & A. Carrasco. (2019). Privacy in mobile health applications for breast cancer patients. In: *IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain*, pp. 634-639. DOI: 10.1109/CBMS.2019.00131.
- [8] A. Md. & H. J. Kim. (2008). A health-care service for mobile computing. In: *Advanced Software Engineering and Its Applications, Hainan Island*, pp. 190-193. DOI: 10.1109/ASEA.2008.28.