Department of Information Technology
NBA Accredited

A.P. Shah Institute of Technology

G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2021-2022

A Project Report on

Breast Cancer Detection

Submitted in partial fulfillment of the degree of
Bachelor of Engineering(Sem-6)
in
INFORMATION TECHNOLOGY
By
Simran Choudhary(19104023)
Janhavi Kulkarni(19104062)
Loveritu Itnare(19104028)

Under the Guidance of
Prof.Kiran Deshpande
          &
Prof.Jayshree Jha

1.Project Conception and Initiation

- Breast cancer detection is done with the help of mammograms, which are basically X-rays of the breasts.

- It's a tool which can help detect and diagnose breast cancer.

- The result of a mammogram are images that can show any calcifications or deposits of calcium in the breasts.

- Cancer is a disease that occurs when there are changes or mutations that take place in genes that help in cell growth.

1.1 Objectives

- Help doctors analyze the huge datasets of cancer data and find patterns with the patient's data and that cancer data available.

- With this analysis we can predict whether the patient might have breast cancer or not. Machine learning algorithms will help with this analysis of the datasets.

- These techniques will be used to predict the outcome.

- The Machine Learning algorithms will analyze the data from the datasets to predict whether the patient has breast cancer or not.

1.2 Literature Review

Author	Image Database	Feature	Remark
Caberrera et al. [2]	DDSM database	Texture feature	Highly dependent on a reference gray level
Pavelkral et al. [4]	DDSM and MIAS database	LBP features	Accuracy achieved – 84%
Anujkumar et al. [5]	---	Gradient	Dependent on threshold parameter selection and size of the averaging filter
Oliveira et al. [5]	DDSM database	Texture characteristics	Accuracy achieved – 85%
S. Naresh et al. [7]	MIAS database	LBP features	Results of CLBP are better than LBP
Oliver et al. [8]	---	Eigen faces approach	Result is $A_Z = 0.92$.
Peirrira et al. [9]	MIAS database	LBP features	ROC performance > 0.8
Oliver et al. [10]	DDSM database	LBP features	ROC performance = 0.909 +/-0.043
Spandana et al. [11]	Mammographic images of 14 patients were acquired from KIMS, Hyderabad	Texture, wavelet and geometric features	--

1.3 Scope

- If a patient has breast cancer, ML can be used to predict the case based on the huge amount of data on breast cancer.

- This system helps the patients as it reduces the amount of money they need to spend just for the diagnosis

- The model is efficient as there is no requirement of manual data entry work or any work by doctors

1.4 Problem Definition

A continuous evolution related to cancer research has been performed. Scientists used various methods, like early stage screening, so that they could find different types of cancer before it could do any damage. With this research, they were able to develop new strategies to help predict early cancer treatment outcome . With the arrival of new technology in the medical field, huge amount of data related to cancer has been collected and is available for medical research. But, physicians find the accurate prediction of the cancer outcome as the most interesting yet challenging part.

2.1 Proposed System

- Traditionally, the diagnosis of breast cancer and the classification of the cancer as malignant or benign was done by various medical procedures like:
- Breast exam – The doctor would check the breasts and lymph nodes in the armpits to check if there are any lumps or abnormalities.
- Mammogram – These are like X-ray of the breast.
- Breast ultrasound - Ultrasound uses sound waves to produce images to determine whether a new breast lump is a solid mass or a fluid-filled cyst
- Removing a sample of breast cells for testing (biopsy) – This is probably the only definite way of checking if a patient has breast cancerRI of the breasts - An MRI machine uses a magnet and radio waves to create pictureseethe interiors of the breast tissues.
- Blood tests, CT scans and PET scans are also done to check for breast cancer.

2.2 Features and Functionality

- The project provides a deep glance for accurate and quick result if the user is having cancer or not .

- The project is user friendly and quick and also helps in the detection purpose.

- Having a ethical dataset helps it give accurate results up to 95% and avoids false positive results which is a must.

2.3 Technology stack

- Python: language
- NumPy: library for numerical calculations
- Pandas: library for data manipulation and analysis
- SkLearn: library which features various classification, regression and clustering algorithms
- Flask: microframework for building web applications using Python.

3. Project Design/Algorithm

1.DECISION TREES

This algorithm is used to predict the value of an output or target variable based on many input variables. They are a collection of divide and conquer problem solving strategies. It takes the shape of a tree like structure. It starts with root nodes and this splits into sub-nodes or child nodes.

K-NEAREST NEIGHBOUR

This algorithm is one of the simplest Machine learning techniques. It is a lazy learning algorithm used for regression and classification. It classifies the objects using their "k" nearest neighbors. k-NN only considers the neighbors around the object, not the underlying data distribution. If k = 1, it basically assigns the unknown to the class of the nearest neighbor. If k > 1, the classification is decided by majority vote based on the k nearest neighbor prediction result.

Naïve Bayes

Naive Bayes classifier is based on Bayes' theorem and is one of the oldest approaches for classification problems.We used it to classify data.

Design(Flow Of Modules)
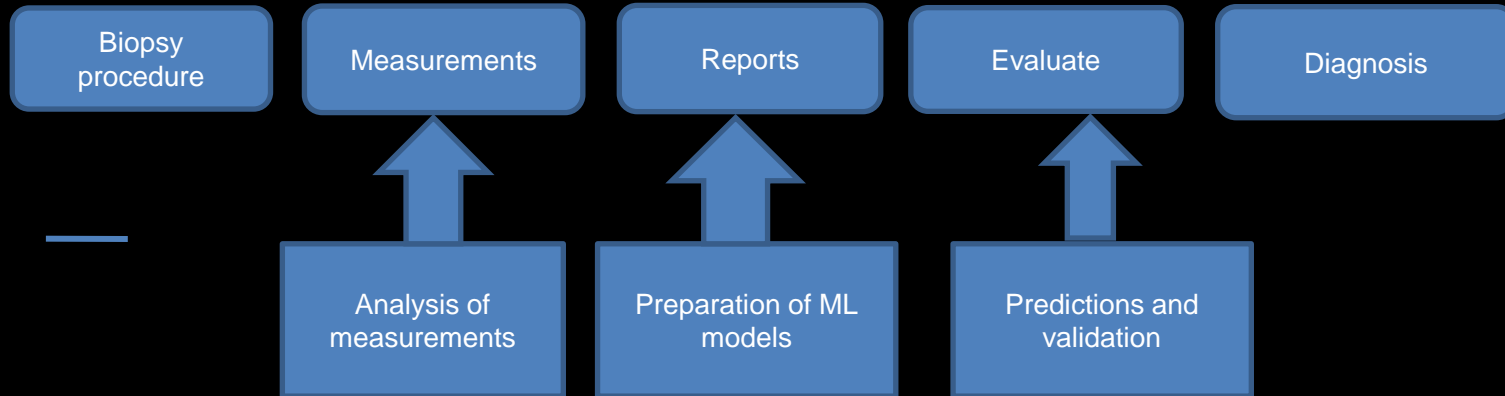
4. Implementation

Breast Cancer Prediction

Predicts whether the cancer is benign or malignant with an accuracy of more than **95%**.

Benign Tumors

Since benign tumors are harmless, doctors may opt to leave the lump alone rather than remove it. If you are experiencing discomfort, irritation or pressure, let your doctor know so that she can plan to remove it for you and improve your comfort.

Malignant tumors

Malignant tumors are cancerous and aggressive because they invade and damage surrounding tissue. When a tumor is suspected to be malignant, the doctor will perform a biopsy to determine the severity or aggressiveness of the tumor.

What is Breast Cancer   Who gets Breast Cancer   What Are the Symptoms of Breast Cancer

What is Breast Cancer

Cancers are typically named after the part of the body from which they originate. Breast cancer originates in the breast tissue. Like other cancers, breast cancer can invade and grow into the tissue surrounding the breast. It can also travel to other parts of the body and form new tumors, a process called metastasis..

Radius Mean	Enter radius_mean
Texture Mean	Enter texture_mean
Perimeter Mean	Enter perimeter_mean
Area Mean	Enter area_mean
Smoothness Mean	Enter smoothness_mean
Compactness Mean	Enter compactness_mean
Concavity Mean	Enter concavity_mean

Concave Points Worst	Enter concave points_worst
Symmetry Worst	Enter symmetry_worst
Fractal Dimension Worst	Enter fractal_dimension_worst

Submit

5. Result
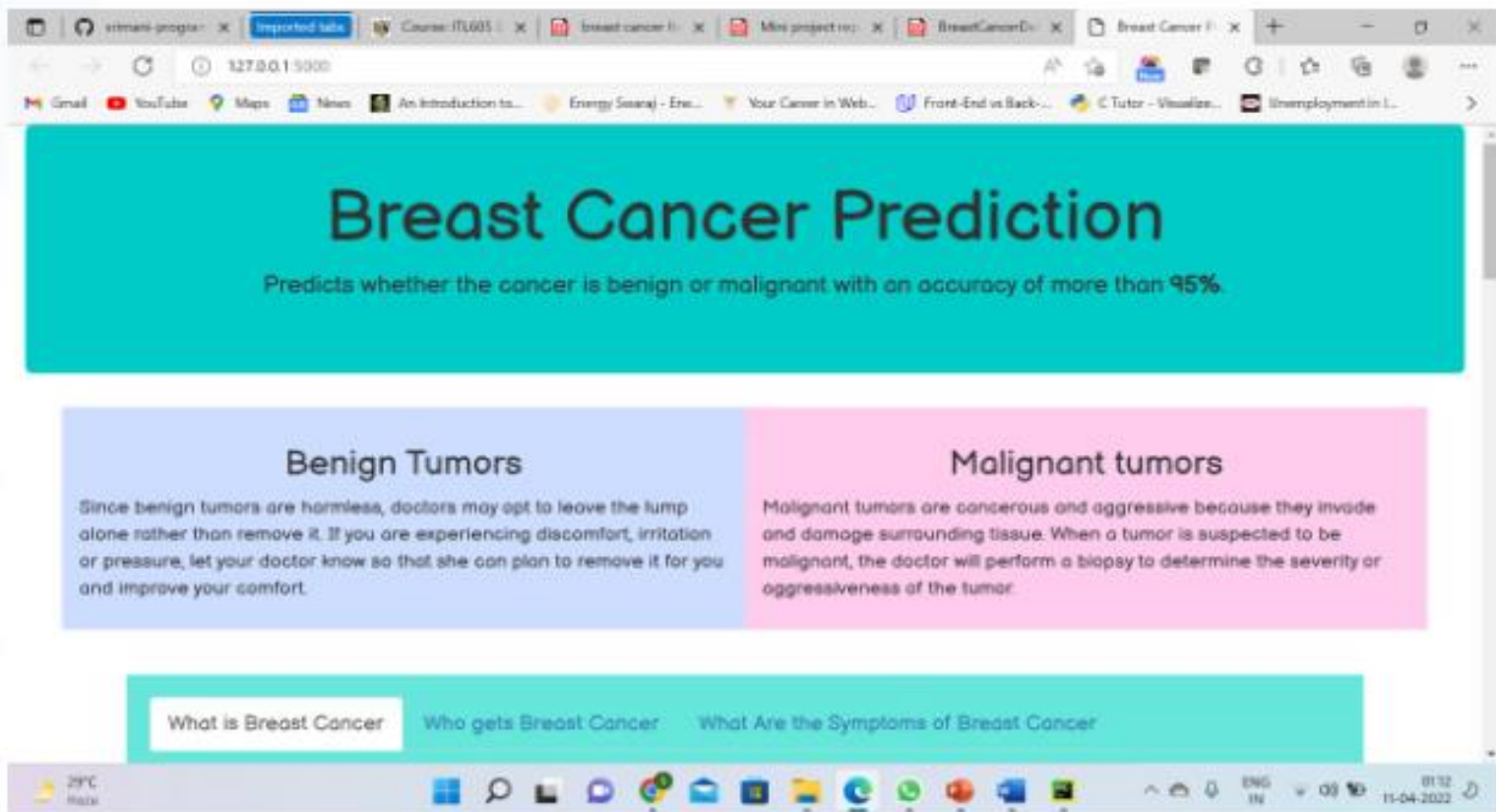
6. Conclusion and Future Scope

The use of machine learning techniques can be applied to detection of breast cancer status but different degrees of accuracy can be obtained. The study has shown the prediction accuracy of three machine learning models to predict the presence or absence of Breast Cancer.

References

Wolberg, Street and Mangasarian, "Wisconsin Diagnostic Breast Cancer Dataset"
http://archive.ics.uci.edu/ml

Mengjie Yu, "Breast Cancer Prediction Using Machine Learning Algorithm", The University of Texas at Austin, 2017.

Wenbin Yue, Zidong Wang, Hongwei Chen, Annette Payne and Xiaohui Liu, "Machine Learning with Applications in Breast Cancer Diagnosis and Prognosis", 2018.

S. Palaniappan and T. Pushparaj, "A Novel Prediction on Breast Cancer from the Basis of Association rules and Neural Network", 2013.

Joseph A. Cruz, David S. Wishart, "Applications of Machine Learning in Cancer Prediction and Prognosis"

Thank You