BVRIT HYDERABAD College of Engineering for Women

Department of Information Technology

BREAST CANCER PROGNOSIS USING MACHINE LEARNING

Under the guidance of
 Guide Name: Mr. B.Srinivasulu
 Designation: Assistant Professor

Team- 09
 R. Parimala (19WH1A1268)
 P. Aditi Kiran(19WH1A1277)
 Ch. Lakshmi Durga(19WH1A1295)
 Raveena Yadlapalli (19WH1A1296)

Agenda

- Introduction
- Existing System
- Problem statement
- Literature Survey
- Proposed system
- Tools and Technology
- Feasibility study
- Societal impact
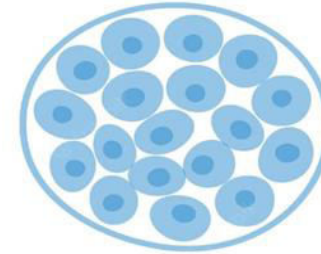- Project timeline
- References

Introduction

- Basically, there are two types in breast cancer. They are malignant and benign.

   Types of breast cancer:
   > **Malignant**
   >**Benign**

- The main objective of this project to build the model for predicting cancer using various machine learning techniques and predicting the chances of reoccurrence of cancer.

Benign tumor
- Non-cancerous
- Capsulated
- Non-invasive
- Slow growing
- Do not metastasize (spread) to other parts of the body
- Cells are normal

Maligant tumor
- Cancerous
- Non-capsulated
- Fast growing
- Metastasize (spread) to other parts of the body
- Cells have large, dark nuclei, may have abnormal shape

Problem Statement

- Computer Science & Engineering is used in Bioinformatics and Biomedical to diagnose and prognoses disease Cancer. This can be further directed to a field called Machine Learning where various techniques are available to predict the cancer on the basis of collected standard data sets.
- We need to apply some classifiers of Machine Learning Techniques to signify the cancer in a human.

Existing system

There are 2 existing systems that are already being used and they are

- Breast Cancer Risk Assessment Tool (BCRAT)

- Breast and Ovarian Analysis of Disease Incidence (BOADICEA) model

Literature Survey

S.No	Author	Title	Published year	Observation
1	Krishna Mridha	Early Prediction of Breast Cancer by using Artificial Neural Network and Machine Learning Techniques	2021	The paper focuses on numerous models that are applied to the dataset taken. In terms of accuracy, cross-validation, sensitivity, and specificity gained, each of the algorithm was calculated and compared.Random Forest gives the best accuracy is 98.83%
2	Prateek P.Sengar, Mihir J.Gaikwad, Ashlesha S.Nagdive	Comparative study of Machine Learning Algorithms for Breast Cancer Prediction.	2020	2 machine learning algorithms is proposed to compare namely Logistic Regression and Decision Tree algorithm on the Wisconsin (Diagnostic) Data Set and use the algorithm with the best accuracy for predicting Breast Cancer.
3	Mamatha Sai Yarabaria, Lakshmi Kavya Ravi, A. Sivasangari	Breast Cancer Prediction via Machine Learning.	2019	predicting whether a person is suffering with breast cancer or not is done with the help of the trained data.

Proposed System

- The objective of this project is to predict and diagnose breast cancer, using machine learning algorithms, and find out the most effective result based on the performance of each classifier/algorithm in terms of confusion matrix, precision, accuracy, and sensitivity. And we are also predicting the recurrence of breast cancer by applying SVM algorithm.

Tools and Technology

Environment	Specification
Hardware	System Hard disk Ram - 4GB Processor - intel core i5 7th generation Hard drive - 64 GB
Software	Operating System - Windows 10 Jupyter Notebook Google Colab Google Chrome Ms-word Python

Department of Information Technology

Feasibility Study

- The system utilizes Machine Learning Algorithms like Logistic Regression,Decision Tree,Naive Bayes,K Nearest Neighbors,Support Vector Machine,Random Forest.
- Tools required are Google Colab,Jupyter Notebook.

Societal Impact

- Breast cancer is considered to be one of the significant causes of death in women. So, early detection of breast cancer plays an essential role to save woman's life.
- Machine Learning models are getting better than pathologists at accurately predicting the development of cancer.
- We are intending how to parametrize our classification techniques to achieve high accuracy. We are looking into many datasets and how further Machine Learning algorithms can be used to characterize Breast Cancer. We want to reduce the error rates with maximum accuracy.

Department of Information Technology

Project Timeline

Dates(To-From)	Duration	Tasks
06.09.2022-22.10.2022	4 weeks	Domain Selection References Specifications Requirements
23.10.2022-12.11.2022	3 weeks	Literature Survey
13.11.2022-18.12.2022	5 weeks	Data Pre-processing & Algorithms
19.12.2022-10.01.2023	3 weeks	Full Implementation

Department of Information Technology

References

- Usman Naseem, Junaid Rashid, "An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers" 12 May 2022

- L Yang, B Fu, Y Li, Y Liu, W Huang, S Feng et al., "Prediction model of the response to neoadjuvant chemotherapy in cancers by a Naive Bayes algorithm", *Computer methods and programs in biomedicine*, vol. 192, pp. 105458, 2020.

- B. Akbugday, "Classification of Breast Cancer Data Using Machine Learning Algorithms," 2019 Medical Technologies Congress (TIPTEKNO), Izmir, Turkey, 2019, pp. 1-4.

THANK YOU