##### A Project report on

**Breast Cancer Detection**

###### A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

**Bachelor of Technology**

**in**

**Computer Science and Engineering (AI&ML)**

Submitted by

E. MAYUKHA

(20H51A6679)

N. SAI NISHANTH

(20H51A6642)

B. DHEERAJ

(20H51A6604)

Under the esteemed guidance of

Ms. A. Deepika

(Assistant professor)



**Department of Computer Science and Engineering (AI&ML)**

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

(An Autonomous Institution under UGC & JNTUH, Approved by AICTE, Permanently Affiliated to JNTUH, Accredited by NBA.)

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

#### 2020- 2024

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (AI&ML)**



#### CERTIFICATE

This is to certify that the Major Project Phase-1 report entitled **"Breast Cancer Detection"** being submitted by E. Mayukha (20H51A6679), N. Sai Nishanth (20H51A6642), B. Dheeraj (20H51A6604) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering(AI&ML)** is a record of bonafide work carried out his/her under my guidance and supervision.

###### The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

**Ms. A. Deepika Dr. P. Sruthi**

**Assistant professor Associate Professor and HOD**

**Dept. of CSE(AI&ML) Dept. of CSE(AI&ML)**

#### ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Ms. A. Deepika , Assistant professor**, Department of Computer Science and Engineering(AI&ML) for her valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. P. Sruthi,** Head of the Department of Computer Science and Engineering(AI&ML), CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academic, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the Teaching & Non- teaching staff of Department of Computer Science and Engineering (AI&ML) for their co-operation

We express our sincere thanks to **Mr. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

E. Mayukha 20H51A6679

B. Dheeraj 20H51A6604

N. Sai Nishanth 20H51A6642

**TABLE OF CONTENTS**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chapters** | | **Description** |  |  | **Page No.** |
|  | | **List of figures** |  |  |  |
|  |  | **Abstract** |  |  |  |
| **1** |  | **Introduction** |  |  |  |
|  | 1.1  1.2  1.3  1.4 | Literature Review  Problem Statement  Research Objectives  Project Scope and Limitations |  |  |  |
| **2** |  | **Background Work** |  |  |  |
|  | 2.1 | GONN for Breast Cancer Classification | 2.1.1  2.1.2  2.1.3 | Introduction  Merits, Demerits and Challenges  Implementation |  |
|  | 2.2 | Computer Based Breast Cancer Prediction | 2.2.1  2.2.2  2.2.3 | Introduction  Merits, Demerits and Challenges  Implementation |  |
| **3** |  | **System Design** |  |  |  |
|  | 3.1 | Proposed System Introduction |  |  |  |
|  | 3.2 | Proposed System Architecture |  |  |  |
|  | 3.3 | Merits of Proposed method |  |  |  |
|  | 3.4 | Implementation |  |  |  |
|  | 3.5 | Source Code |  |  |  |
| **4** |  | **Results and Discussions** |  |  |  |
|  | 4.1  4.2  4.3 | Data collection and  Performance metrics  Comparison of Existing  Solutions  Results and Screenshots |  |  |  |
| **5** |  | **Conclusions** |  |  |  |
| **6** |  | **References** |  |  |  |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title of the Figure** | **Page No.** |
| 1 | GONN for Breast Cancer Classification |  |
| 2 | Implementation of GONN |  |
| 3 | Computer Based Breast Cancer Prediction |  |
| 4 | Implementation of Computer based method (a) |  |
| a | Implementation of Computer based method (b) |  |
| 6 | Proposed System Architecture |  |
| 7 | Implementation of Proposed system |  |
| 8 | Performance metrics and accuracy |  |
| 9 | Accuracy of Logistic regression |  |
| 10 | Error rate of Logistic regression |  |
| 11 | Accuracy of Naïve Bayes |  |
| 12 | Error rate of Naïve Bayes |  |
| 13 | Home page of website (a) |  |
| 14 | Home page of website (b) |  |
| 15 | Home page of website (c) |  |
| 16 | Home page of website (d) |  |
| 17 | Predict page of website (a) |  |
| 18 | Predict page of website (b) |  |
| 19 | Predict page of website (c) |  |
| 20 | Predict page of website (d) |  |

# **ABSTRACT**

Breast cancer is one of the leading causes for the death of women. In women, breast cancer is treated as the most significant issue. Early diagnosis of this helps to prevent the cancer. If breast cancer is detected in early stage, then survival rate is very high. Machine Learning methods are effective ways to classify data. Especially in the medical field, where those methods are widely used in diagnosis and analysis for decision making.

Breast cancer is one of the most common cancers among women worldwide, representing most new cancer cases and cancer related deaths according to global statistics, making it a significant public health problem today.

Further accurate classification of benign tumors can prevent patients undergoing unnecessary treatments. Thus, the correct diagnosis of BC and classification of patients into malignant or benign groups is the subject of much research.

Graphical user interface, text, application

Description automatically generated

# **CHAPTER 1**

**INTRODUCTION**

**CHAPTER 1**

**INTRODUCTION**

After lung cancer, BC (Breast Cancer) is the second popular cause of death in both developed and undeveloped worlds. BC is characterized by the mutation of genes, constant pain, changes in the size, color(redness), skin texture of breasts. The most frequent classification is binary (benign cancer/malign cancer). Today, Machine Learning (ML) techniques are being broadly used in the breast cancer classification problem. They provide high classification accuracy and effective diagnostic capabilities.

In our project, we present the Naive Bayes (NB) classifier for breast cancer classification. We propose a comparison between the already existing Logistic Regression implementation of the classification and the NB classifier and evaluate their accuracy. Results show that NB gives the highest accuracy of 93.51% with lowest error rate then Logistic Regression classifier which gives 92.19% accuracy.

Text

Description automatically generated with medium confidence

* 1. **Literature Review**

Arpit B. and Aruna T. proposed a genetically optimized neural network for breast cancer classification. To evaluate their work, they used WBCD and compared the classification accuracy, confusion matrix. This method presents a good accuracy classification. However, it can be improved by using a larger dataset than WBCD.

Ashraf O. I. and Siti M. S. [6] proposed a computer-based method to automatically classify breast cancer disease. The method applied multilayer perceptron (MLP) neural network based on enhanced non-dominated sorting genetic algorithm to optimize the accuracy. Compared to other methods, this work improves classification accuracy. However, MLP can get stuck in local minima.

Haifeng W. designed an SVM-based ensemble learning model for breast cancer diagnosis. The proposed ensemble model includes two types of SVM structures, i.e., a C-SVM and a -SVM, and six types of kernel functions. The proposed model increases diagnosis accuracy compared to other works based on single SVM. However, it is a computationally expensive method, and the training time is high.

Na L. proposed an intelligent classification model for breast cancer diagnosis based on a hybrid feature selection approach using support vector machine learning algorithm. This process can improve the classification accuracy and reduce the computational cost. The proposed work shows a good performance and decreases the calculation complexity.

* 1. **Problem Statement**

The aim of our project is to predict breast cancer, which is the second leading cause of death among women worldwide, and with early detection and prevention can dramatically reduce the risk of death, using several machine-learning algorithms by choosing the most effective of them. Few of them are Random Forest, Logistic Regression, Naïve Bayes, Support Vector Machines, and K-Nearest Neighbors.

The main objective is to evaluate the efficiency and effectiveness of the algorithm used for classification of data by predicting the outputs with high accuracy. We will be using Naïve Bayes Classifier to help understand the severity of breast cancer. Labels will be given as per the severity and as soon as we enter the details required to get the prediction, we will get an output stating whether the patient’s cancer level is Malignant or Benign.

Diagram

Description automatically generated

* 1. **Research objective**

The goal is to increase the proportion of breast cancers identified at an early stage, allowing for more effective treatment to be used and reducing the risks of death from breast cancer. Since early detection of cancer is key to effective treatment of breast cancer, we use various machine learning algorithms to predict if a tumour is benign or malignant, based on the features provided by the data.

This analysis aims to observe which features are most helpful in predicting malignant or benign cancer and to see general trends that may aid us in model selection and hyper parameter selection. To achieve this, we have used machine learning classification methods to fit a function that can predict the discrete class of new input.

**1.4**  **Project Scope and Limitations**

**Scope:**

We would like to work on reducing the error rate for the Naïve Bayes Algorithm and increase its accuracy. We would also like to visualize the output in better and beautiful ways for complete understanding of the data. We would also like to visualize the output in better and beautiful ways for complete understanding of the data.

**Limitations:**

The precision will decrease when the size of the dataset is small. If our test data set has a categorical variable of a category that wasn’t present in the training data set, the Naive Bayes model will assign it zero probability and won’t be able to make any predictions in this regard.

**CHAPTER 2**

**BACKGROUND**

**WORK**

**CHAPTER 2**

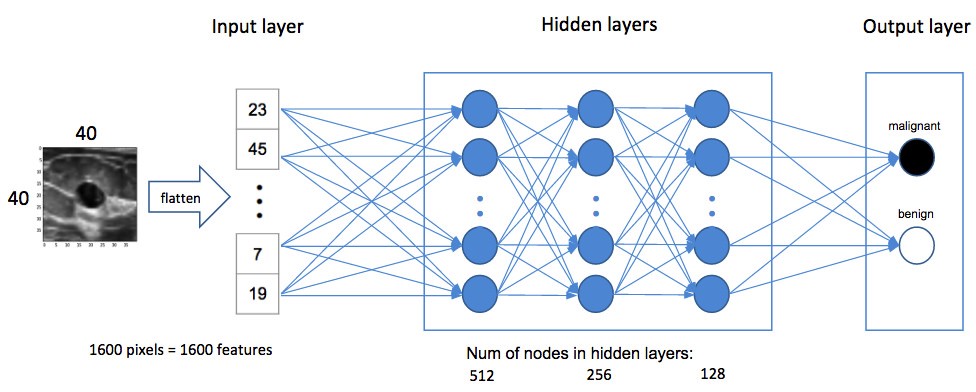
**BACKGROUND WORK**

**2.1 GONN for Breast Cancer Classification**

**2.1.1 Introduction**:

A Genetically Optimized Neural Network (GONN) model is proposed which simultaneously evolved the structure and weight of neural network for classifying the Wisconsin breast cancer dataset (WBCD) as benign or malignant. It is observed that the proposed algorithm yielded an accuracy of 98.24% for 50–50, 99.63% for 60–40 and 100% for 70–30 training–testing partition respectively and classification accuracy of 100% for 10-fold cross validation scheme.

The results show that the approach works well with the breast cancer database and can be a good alternative to the well-known machine learning methods. Measures such as sensitivity, specificity, ROC curves, Area under the ROC curves (AUC) and Mann–Whitney two tailed test are used to validate the performance.

A wide range of methods have been proposed to forecast medical diagnosis of breast cancer with WBCD also with recently proposed algorithms applied on the WBCD database.

**2.1.2 Merits, Demerits and Challenges:**

**Merits:**

* The GONN model can be a very helpful tool to assist the physicians to diagnose the patient or it can be used as a second opinion for their final diagnosis.
* This GONN model shows the competent or better results when compared with the state-of-the art methods applied to the WBCD database.
* An average accuracy of 97.73%, 99.11% and 99.21% is achieved for 50–50, 60–40 and 70–30 training–testing partitions respectively and 99.26% for 10-fold cross validation scheme.
* The results of the confusion matrix show that with increased number of training samples, the number of false positive and false negative rates decreases.

**Demerits:**

* Only crossover and mutation operators are improved.
* One other demerit is that the GONN model is applied on WBCD dataset available on UCI repository which is very small. Thus, generalizability of the proposed model must be further tested with variety of actual data from different hospitals.

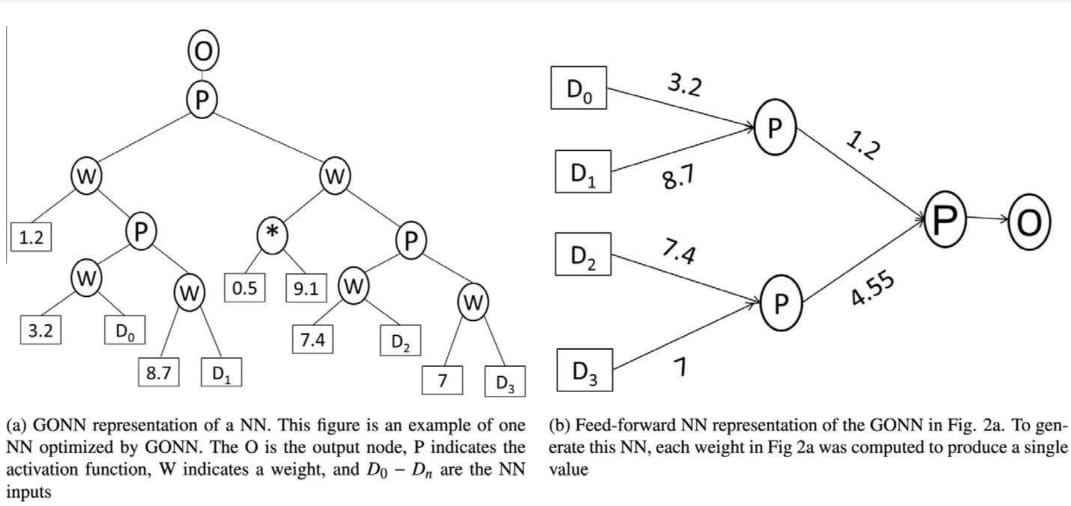
**Challenges:**

* The proposed model is applied on numeric data only; it would be interesting to see its behavior when it is applied on different types of data available in medical field such as images, signals.
* However, for practical implementation, future work is required to evaluate the usefulness of the proposed method with a greater number of patients**.**

**2.1.3 Implementation:**

To evolve the ANN architecture using GP, for solving the classification of WBCD dataset, we must form a GONN architecture in such a way to treat it like an ANN structure. A final GONN architecture represents an ANN with appropriate mapping of GP parameters to ANN learning parameters. The detail description of the mapping of GONN architecture to its equivalent Feed Forward Neural Network representation is described next.

We input the WBCD training data and GP parameters and output the GONN architecture. While No. of fitness evaluations < Maximum number of fitness evaluations, perform fitness evaluation and calculate the fitness value of all individuals. Then we apply modified crossover operation on all the remaining parent pairs. For all crossover pairs, repeat till we get offspring better than parents. Take the parent pair and generate two offspring from them by giving more chances to function nodes (90%) to swap.



Place the top offspring of that pair into a table sorted according to fitness values. Select the top Pc% offspring from the sorted table and transfer them to the next generation. Take the parents of remaining Pm% individuals in the table and apply modified mutation. Repeat till we get offspring better than parents. Take the parent and generate child from them by giving more chances to terminal nodes (90%) . Return best individual in terms of fitness value which is considered as Genetically Optimized Neural Network (GONN) architecture.

**2.2 Computer Based Breast Cancer Prediction**:

**2.2.1 Introduction**:

Diagnosis of breast cancer disease depends on human experience. It is time consuming and has an element of human error in the results. This method presents an intelligent multi-objective classifier to Diagnose breast cancer diseases using multilayer perceptron (MLP) neural network with Differential Evolution technique. The Differential Evolution (DE) algorithm is used to solve multi-objective optimization problems by tuning MLP neural network parameters.

In addition, it utilizes the advantages of multi-objective differential evolution to optimize the number of hidden nodes in the hidden layer of the MLP neural network and to reduce network error rate. In this method, MLP technique is used as a prediction model for the breast cancer tumors , as a set of experiments have been conducted on a variant parameter of the MLP model to enhance the prediction accuracy of the generated model.

In addition to the importance of detecting the tumor in the early stages, the necessity to find sub-optimal models is required. This model presents the results of the MLP method that would be baselines for researchers to focus more on improving such techniques for the future of breast cancer diseases.

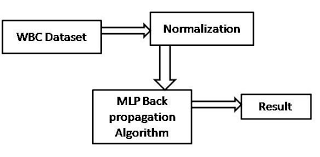
Diagram, schematic

Description automatically generated

**2.2.2 Merits, Demerits and Challenges:**

**Merits:**

* For regression and mapping, the MLP algorithm is a generally very good algorithm.
* It tends to be utilized to map an N-dimensional input signal to an M-dimensional output signal, this mapping can likewise be non-linear.

****

**Demerits:**

* The principal restriction of the MLP algorithm is that the number of hidden neurons must be set by the user.
* Setting this value too low may result in the MLP model underfitting while at the same time setting this value too high may result in the MLP model overfitting.
* Another restriction of the MLP algorithm is that due to how it is trained, it can't ensure that the minima it stops at amid training are the global minima.
* The MLP algorithm can stall out in local minima.

**2.2.3 Implementation:**

This model consists of the following steps:

* Data Selection and Preparation: The dataset is selected from WDBC. This dataset includes 569 instances and 32 attributes.
* Preprocessing Data: In this step the data set is loaded, then a supervised filter which is attribute-based is applied on the data set which removes unnecessary attributes and therefore 1 attribute was selected.
* Classification Iteration: In this step, we classify and select cross-validation 10 and select the MLP algorithm.
* Ensemble feature selection: In this step, we define and select attributes and choose a search method and attribute evaluators.
* Hence, the number of features will be reduced.

Diagram

Description automatically generated

**CHAPTER 3**

**SYSTEM DESIGN**

**CHAPTER 3**

**SYSTEM DESIGN**

**3.1 Proposed System Introduction**

Our proposed model uses Naïve Bayes Classifier for predicting the patient’s cancer severity. This is the best algorithm for prediction of breast cancer since it has a better accuracy as compared with other algorithms. The Naive Bayes Classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naïve Bayes can often outperform more sophisticated classification methods.

The Naïve Bayes technique depends on the famous Bayesian approach following a clear and fast path. A Naïve Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong independence assumptions. A Naïve Bayes classifier assumes that the presence (or absence) of a particular feature of a class is un-related to the presence (or absence) of any other feature, given the class variable.

We will be using the scikit-learn load\_breast\_cancer dataset for our project involving binary classification. The dataset contains more than 500 instances of subject records containing 30 unique features. We have compared the accuracies of the algorithms useful for predicting breast cancer and found that Naïve Bayes Classifier gives the best results.

The system suggested that machine learning techniques can be acted as a clinical assistant for the diagnosis of breast cancer and will be very helpful for new doctors or physicians in case of misdiagnosis. Our breast cancer prediction model could help the early diagnosis of the disease with necessary care plans.

The Naive Bayes Classifier technique is based

on the so-called Bayesian theorem and is par-

ticularly suited when the dimensionality of the

inputs is high. Despite its simplicity, Naive

Bayes can often outperform more sophisticated

classification methods

The Naive Bayes Classifier technique is based

on the so-called Bayesian theorem and is par-

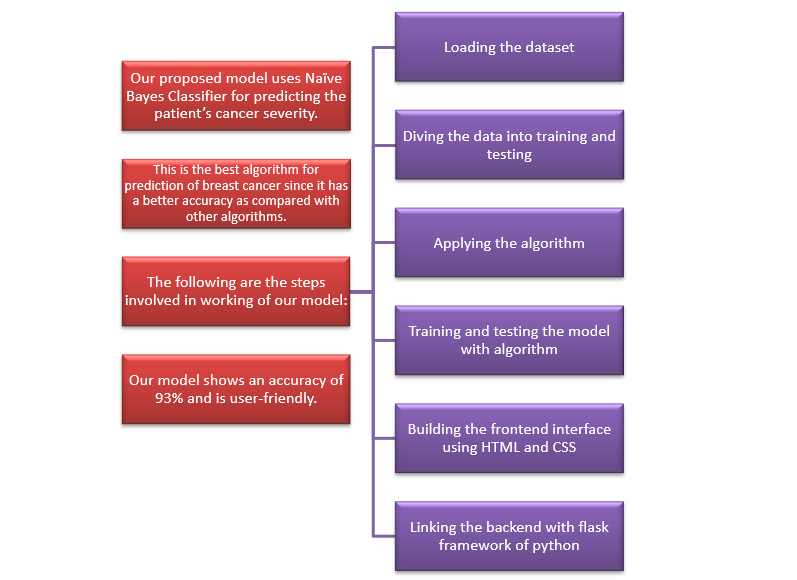
ticularly suited when the dimensionality of the

inputs is high. Despite its simplicity, Naive

Bayes can often outperform more sophisticated

classification methods

**3.2 Proposed system Architecture**

****

**3.3 Advantages of Proposed method**

Voice Input

WhatsApp

Our proposed model has the following features:

* Can be used as a clinical analysist in early stages of the breast cancer.
* Has the highest accuracy and less error rate as compared with other models.
* Helps with the early diagnosis of the disease with necessary care plans.
* Can act as a clinical assistant for new doctors or physicians in case of misdiagnosis.

**3.4 Implementation of the Proposed method**

We start off by importing the load\_breast\_cancer dataset from the scikit-learn library and loading the same dataset. We split the dataset into training data and testing data in the ratio of 80:20 with the help of the train\_test\_split attribute.

Then we import the Naïve Bayes classifier and create and instance of the class.

We fit the model with the x\_train and y\_train values using the .fit method and apply the .predict method on the x\_test value. Using the same information, we get the accuracy of our model. We can also calculate the precision and recall of our model.

For the interface of the website for our predictive model, we made use of HTML, CSS, and Bootstrap to build the front-end of the website. We have linked our Machine Learning code written in python language to the website, via the Flask framework of python. Once we enter the input values in the appropriate fields, we will get the prediction result of whether the patient has a Benign (non-severe) or Malignant (severe) cancer.

Diagram

Description automatically generated

**3.5 Source Code:**

import numpy as np

import pandas as pd

import sklearn.datasets

from sklearn.model\_selection import train\_test\_split

from sklearn.naive\_bayes import GaussianNB

from sklearn.metrics import accuracy\_score

from flask import Flask ,render\_template,request

breast\_cancer\_dataset = sklearn.datasets.load\_breast\_cancer()

data\_frame = pd.DataFrame(breast\_cancer\_dataset.data, columns = breast\_cancer\_dataset.feature\_names)

data\_frame['label'].value\_counts()

data\_frame.groupby('label').mean()

X = data\_frame.drop(columns='label', axis=1)

Y = data\_frame['label']

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size=0.2, random\_state=2)

model = GaussianNB()

model.fit(X\_train, Y\_train)

X\_train\_prediction = model.predict(X\_train)

training\_data\_accuracy = accuracy\_score(Y\_train, X\_train\_prediction)

X\_test\_prediction = model.predict(X\_test)

test\_data\_accuracy = accuracy\_score(Y\_test, X\_test\_prediction)

app=Flask(\_\_name\_\_)

@app.route("/")

def r1():

    return render\_template("index.html")

@app.route("/predict")

def r2():

    return render\_template("predict.html")

@app.route("/age")

def r3 ():

    return render\_template("age.html")

@app.route("/ivsd")

def r4():

    return render\_template("ivsd.html")

@app.route("/sevVsNonsev")

def r5():

    return render\_template("sevVsNonsev.html")

@app.route("/pairplot")

def r5():

    return render\_template("pairplot.html")

@app.route("/index")

def r6():

    return render\_template("index.html")

@app.route("/team")

def r7():

    return render\_template("team.html")

@app.route("/diagnosis")

def r8():

    return render\_template("diagnosis.html")

@app.route("/features")

def r9():

    return render\_template("features.html")

@app.route("/submit", methods=['GET','POST'])

def submit():

    a=float(request.form.get('mean\_radius'))

    b=float(request.form.get('mean\_perimeter'))

    c=float(request.form.get('mean\_area'))

    d=float(request.form.get('mean\_compactness'))

    e=float(request.form.get('mean\_concavity'))

    f=float(request.form.get('mean\_concave\_points'))

    g=float(request.form.get('radius\_error'))

    h=float(request.form.get('perimeter\_error'))

    i=float(request.form.get('area\_error'))

    j=float(request.form.get('worst\_radius'))

    k=float(request.form.get('worst\_perimeter'))

    l=float(request.form.get('worst\_area'))

    m=float(request.form.get('worst\_compactness'))

    n=float(request.form.get('worst\_concavity'))

    o=float(request.form.get('worst\_concave\_points'))

    input\_data = (a,b,c,d,e,f,g,h,i,j,k,l,m,n,o)

    input\_data\_as\_numpy\_array = np.asarray(input\_data)

    input\_data\_reshaped = input\_data\_as\_numpy\_array.reshape(1,-1)

    prediction = model.predict(input\_data\_reshaped)

    print(prediction)

    if (prediction[0] == 0):

        return render\_template("predict.html",p='Report : The Breast Cancer is Severe!')

    else:

        return render\_template("predict.html",p='Report : The Breast Cancer is Not Severe!')

if \_\_name\_\_=='\_\_main\_\_':

    app.config['TEMPLATES\_AUTO\_RELOAD'] = True

    app.run( debug =True)

**CHAPTER 4**

**RESULTS AND**

**DISCUSSIONS**

**CHAPTER 4**

**RESULTS AND DISCUSSIONS**

**4.1 Data Collection and Performance Metrics**

**Technologies Used:**

* Programming language used for ML: Python
* Tool used for backend: Flask Framework
* Tools used for frontend: HTML, CSS, Bootstrap

**Performance Metrics:**

Table

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

**4.2. Comparison of Existing Solution**

Accuracy of the model using Logistic Regression:

Graphical user interface, text, application, email

Description automatically generated

Error rate of the model using Logistic Regression:

Graphical user interface, text, application, email

Description automatically generated

**---------------------------------------------------------------------------------------**

Accuracy of the model using Naïve Bayes:

Graphical user interface, text, application, email

Description automatically generated

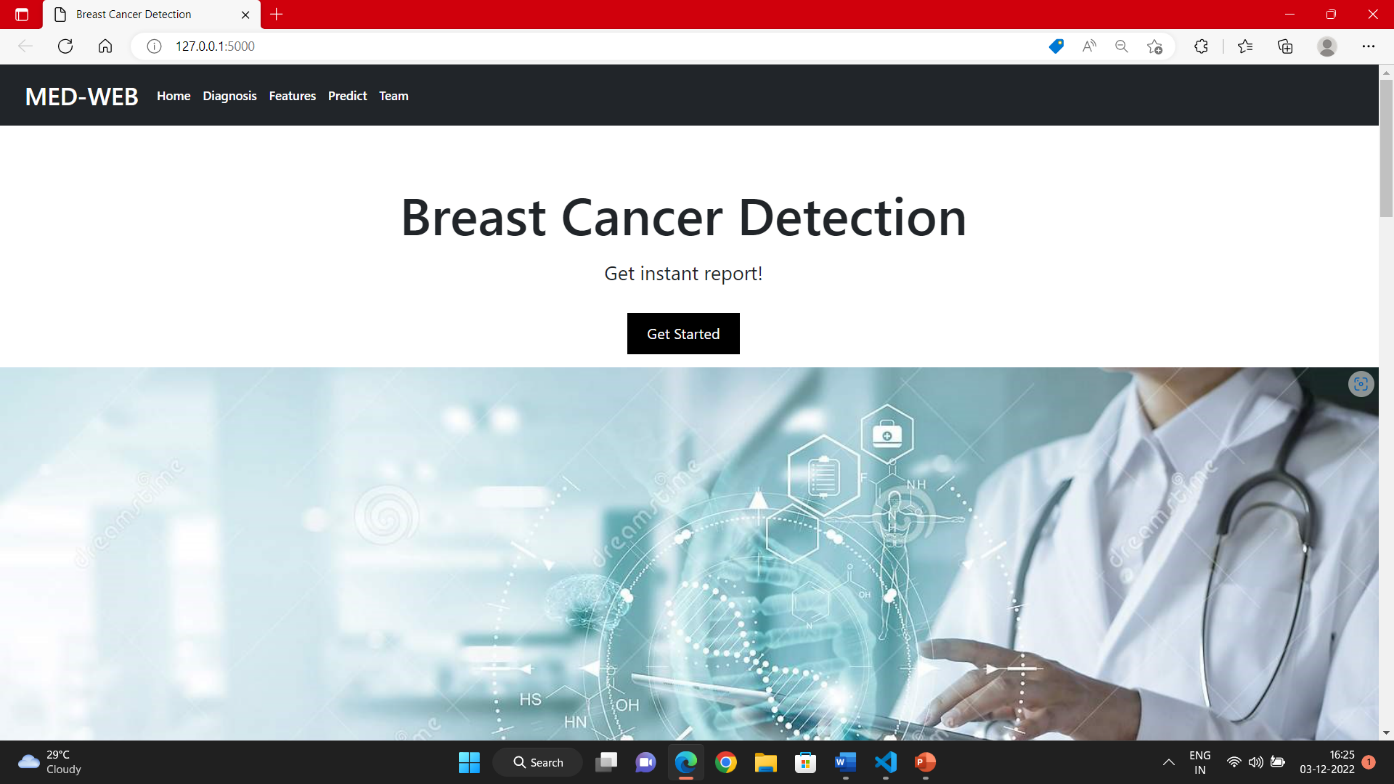
Error rate of the model using Naïve Bayes:

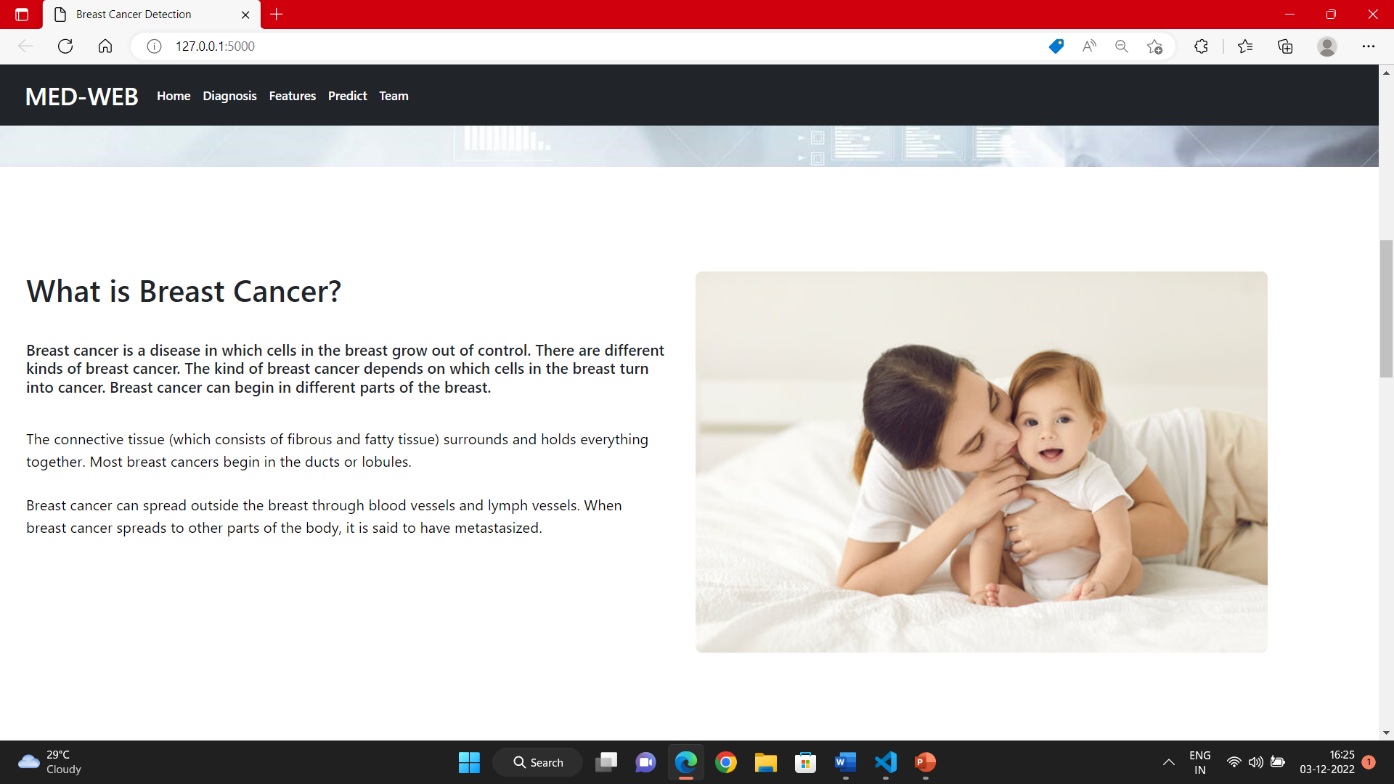
Graphical user interface, text, application, chat or text message, email

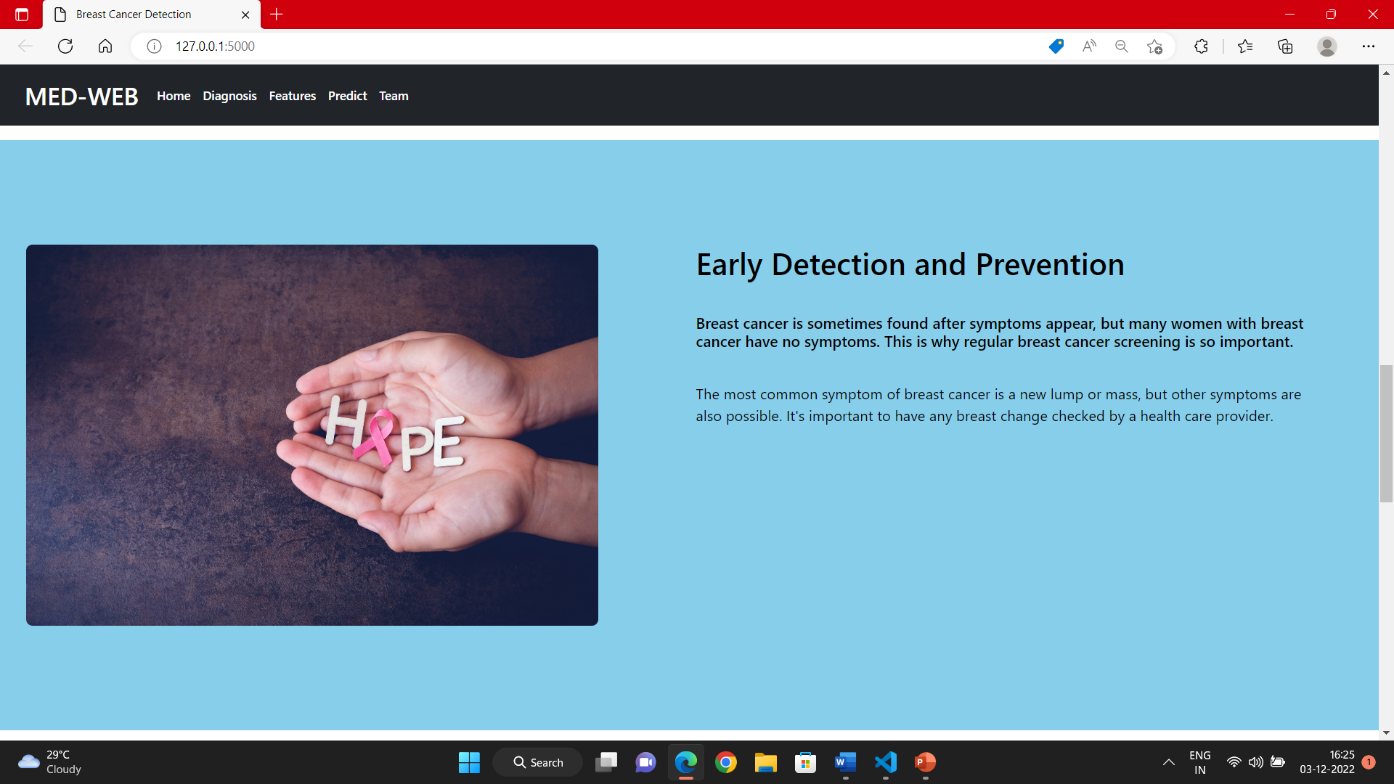
Description automatically generated

**4.3. Results and Screenshots**

The image below shows the main page of our website. We have given some basic information about breast cancer and its preventions.







By clicking on the “Click here” button, we will be redirected to the predict page where you can find out if you are diagnosed with breast cancer:

Graphical user interface, text

Description automatically generated

The following shows the predict page:

Graphical user interface

Description automatically generated

Insertion of appropriate values:

**Graphical user interface

Description automatically generated**

Final output showing the result of the prediction for the given values:

**Graphical user interface, text, table

Description automatically generated**

**Graphical user interface, text, application

Description automatically generated**

**CHAPTER 5**

**CONCLUSIONS**

**CHAPTER 5**

**CONCLUSION**

We would like to work on reducing the error rate for the Naïve Bayes Algorithm and increase its accuracy. We would also like to visualize the output in better and beautiful ways for complete understanding of the data. Breast cancer severity prediction could help the early diagnosis of the disease with necessary care plans.

The system proposed that machine learning techniques can act as a clinical assistant for the diagnosis of breast cancer and will be very helpful for new doctors or physicians in case of incorrect diagnosis. From the study, we can conclude that ML Techniques are able to detect the severity of the disease with high accuracy.

Icon

Description automatically generated with low confidence

**CHAPTER 6 REFERENCES**

**CHAPTER 6**

**REFERENCES**

* <https://www.researchgate.net/publication/308934053_Breast_Cancer_Prediction_using_Naive_Bayes_Classifier>
* American cancer society. Breast cancer facts and figures 2005-06 (http://www.cancer.org)
* <https://www.ijert.org/breast-cancer-detection-using-machine-learning-techniques>
* <https://www.ijert.org/research/breast-cancer-detection-using-machine-learning-techniques-IJERTV10IS070064.pdf>
* Arpita Joshi and Dr. Ashish Mehta “Comparative Analysis of Various Machine Learning Techniques for Diagnosis of Breast Cancer” (2017).
* <https://scholar.google.com/scholarMeghaRathiandVikasPareek2016HybridapproachtopredictbreastcancerusingmachinelearningtechniquesInternationalJournalofComputerScienceEngineering>
* <http://www.breastcancer.org/symptoms/understand_bc/what_is_bc>
* Ashraf O. I. and Siti M. S. [6] computer-based method to automatically classify breast cancer disease.