1. Introduction

• Introduction

• Design thinking

• Project summary and Purpose

• Scope

• Objective

1.**1 INTRODUCTION:-**

* Breast cancer is when a group of cells begin to divide and grow uncontrollably in the breast. Uncontrolled cell growth is dangerous because it can form clumps of cells that spread to other parts of the body. Clumps of cancerous cells can harm the digestive system, lungs, or blood vessels, and can also create chemical imbalances throughout the body.
* The Image processing method filters the data using different algorithms. Three main approaches are used for Image processing.

1. One is image filtering i.e. in image filtering resize the image and remove unwanted part of mammogram.
2. Second is Segmentation filtering, where we try to Show the mass tissue of the breast from the mammogram.
3. Third is detection, in which system detect the white spot from the mammogram image based on the given Threshold.

* The Transform learning method which is part of deep learning which help to classify the white spot from the image is of cancer or not. We have tried different type of model to get accuracy like InceptionV3, VGG16, and CNN etc.
* The system also ask different type of questions and which are selected after research from website, research paper, doctor and articles related to breast cancer.

1. If the cancer spot is detected it will give some possible reason behind the detected cancer.
2. If the cancer is not detected it will give some possible reason for the future risk of the cancer.

**1.2 Scope:-**

Several types of researches have been done to detect breast cancer early so that the treatment can be started to increase the chance of survival. A mammography procedure for early detection and diagnosis of breast cancer is commonly advised. There are so many techniques that are used for malignancy prediction. Using Artificial Intelligence-powered machine learning is widely recommended. Researches are conducted in machine learning to detect cancerous tumors in the human body. Mainly used algorithms which give high accuracy are SVM, naïve bays, decision trees, KNN. Deep learning, the machine learning sub-branch, can also be used to classify breast cancer. Deep learning is a method that is mostly used to clear, rectify, and detect machine learning errors or disadvantages. Convolutional Neural Networks and Inception are the perfect deep learning method for overcoming the drawbacks of machine learning in malignancy detection; however, other strategies such as a recurrent neural network and a deep belief network are being used to overcome the shortcomings of machine learning. As a consequence, using deep learning rather than machine learning yields better results. This review paper's primary motivation is to make budding researchers aware that breast cancer is a serious issue among women and we need to be swift in using different technologies to detect and to improve accuracy as efficiently as possible, it is very important to save our mothers, sisters, loved ones and our society from this dangerous predator.

**1.3 Project summary and Purpose :-**

1.3.1 Problem:

Currently, breast cancer is the second leading cause of death for women worldwide. Although more women today are being diagnosed with breast cancer than in recent years. Mammograms are the best breast cancer screening tests we have at this time. But mammograms have their limits. For example, they aren’t 100% accurate in showing if a woman has breast cancer. Solution for this problem we thought to improve the Detection system.

1.3.2 Observations:

Observation 1:

Recent studies show:

* 1 in 8 women will be diagnosed with breast cancer
* 3/4 of women with breast cancer will survive more than 10 years
* 2/3 of women with breast cancer will survive more than 20 years

Observation 2:

Mammograms are low-dose x-rays of the breast. Regular mammograms can help find breast cancer at an early stage, when treatment is most likely to be successful. To evaluate an area of your breast that has a lump or other signs that might indicate cancer but sometimes due to dense breast tissue it become difficult to find cancer spot out of a mammogram images for which this system enhance the image and try to get proper result.

**1.4 Overview of the project**

The aim of this study was to evaluate the diagnostic accuracy of multipurpose image analysis software based on deep learning with InceptionV3 for the detection of breast cancer in an independent, mammography data set. A neural network is trained to detect the spot at based on Threshold. The system also do image enhancement and image segmentation of the mammogram images and detect the spot of the cancer. And the list of questions are asked to know the risk of the cancer occur in future and to know the reason behind the cancer. Waiting for diagnosing a breast cancer for a long time may increase the possibility of the cancer spreading. Therefore a computerized breast cancer diagnosis has been developed to reduce the time taken to diagnose the breast cancer and reduce the death rate.

**1.4.1 Problem definition:-**

Detect cancer spot from mammogram with use of different types of Deep learning techniques and image processing.

**2.** System Requirement Study

• Tools and Technology

• User Characteristics

• Software and Hardware Requirements

• Constraints

• Assumptions and dependency

• Brief History of work done

**2.1 About Tools and Technology**

FRONT END: Html, CSS, java script, bootstrap.

BACKEND: DL Algorithms, Flask, Transform learning methods, Image processing method.

DATABASE: php

PREDICTOR Libraries: Python, Matplotlib, Numpy, Pandas, Seaborne, Opencv, Sklearn, Tensorflow.

TOOLS: Jupyter notebook, VS Code, Anaconda, Chrome, Default Image reader, Github.

**2.2 User Characteristics**

* User must have basic knowledge of Computers.
* User should understand the use of all modules.
* User can easily interact with the proposed system.
* User must have knowledge of mammogram and mammogram images.
* User must be aware of the system.

**2.3 Hardware and software Requirement**

HARDWARE REQUIREMENTS: The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. A hardware requirements list is often accompanied by a hardware compatibility list (HCL), especially in case of operating systems. An HCL lists tested, compatibility and sometimes incompatible hardware devices for a particular operating system or application. The following subsections discuss the various aspects of hardware requirements.

Microsoft Windows XP Professional SP3/Vista SP1/Windows 7 Professional:

Processor: 800MHz Intel Pentium III or equivalent, i3

Memory: 512 MB Disk

space: 750 MB of free disk space

Ubuntu 9.10:

Processor: 800MHz Intel Pentium III or equivalent

Memory: 512 MB

space: 650 MB of free disk space

**2.4 SOFTWARE REQUIREMENTS**

Software Requirements deal with defining software resource requirements and pre-requisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre- requisites are generally not included in the software installation package and need to be installed separately before the software is installed.

OPERATING SYSTEM: Windows 7/ XP/8 And above

**2.5 Constraints**

Regulatory Policies: - For the accuracy of the model we need a proper classified dataset which was not available publically and the radiologist didn’t provide private dataset.

Reliability Requirements: - System depends only on the mammogram images.

**2.6 Assumptions and Dependencies:**

Assumptions:

• Database of patients are assumed to be secure and reliable.

• User is the person having enough knowledge for the traversing operation.

• We will provide a user friendly interface so that any user can easily navigate through the system.

• The server used for data storing is always secured.

Dependencies:

• The system is dependent upon the patient’s valid images.

• All the users of the system will be assigned a specific role. According to these roles each and every user will be allowed to access predefined set of feature

**2.6 Brief History of work done**

Breast cancer is a major concern among women. It is the second-most common and leading cause of cancer deaths among women. According to published statistics, breast cancer has become a major health problem in both developed and developing countries over the past 50 years, and its incidence has increased in recent years. At present, there are no effective ways to prevent breast cancer, because its cause remains unknown. However, efficient diagnosis of breast cancer in its early stages can give a woman a better chance of full recovery. Therefore, early detection of breast cancer can play an important role in reducing the associated morbidity and mortality rates.

Computer-aided detection or diagnosis (CAD) systems, which use computer technologies to detect abnormalities in mammograms such as calcifications, masses, and architectural distortion, and the use of these results by radiologists for diagnosis [4], can play a key role in the early detection of breast cancer and help to reduce the death rate among women with breast cancer. Thus, in the past several years, CAD systems and related techniques have attracted the attention of both research scientists and radiologists. For research scientists, there are several interesting research topics in cancer detection and diagnosis systems, such as high-efficiency, high-accuracy lesion detection algorithms, including the detection of masses, detection of architectural distortion, and the detection of bilateral asymmetry. Radiologists, on the other hand, are attracted by the effectiveness of clinical applications of CAD systems.

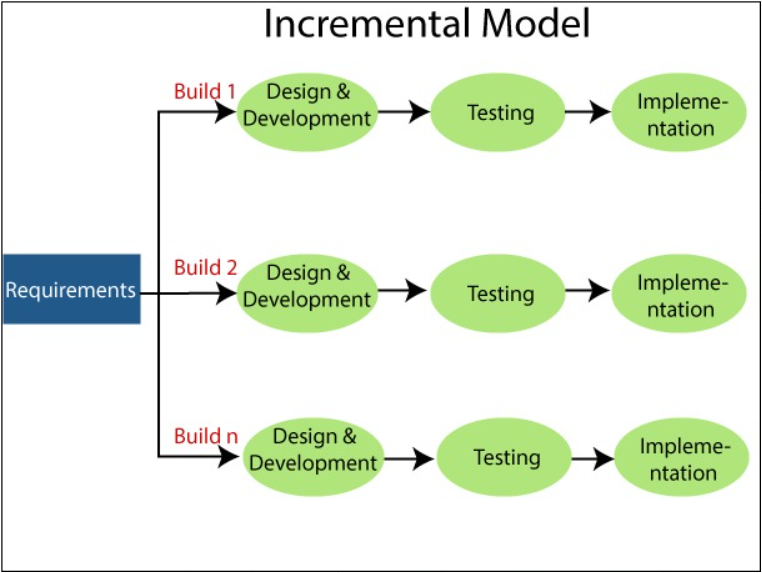
3. Project Planning

• Project Development Approach

• Prototype

* 1. **Project Development Approach:-**

The model that is referred for the development of the project is “INCREMENTAL”.



1. Requirement analysis: In the first phase of the incremental model, the product analysis expertise identifies the requirements. And the system functional requirements are understood by the requirement analysis team. To develop the software under the incremental model, this phase performs a crucial role.

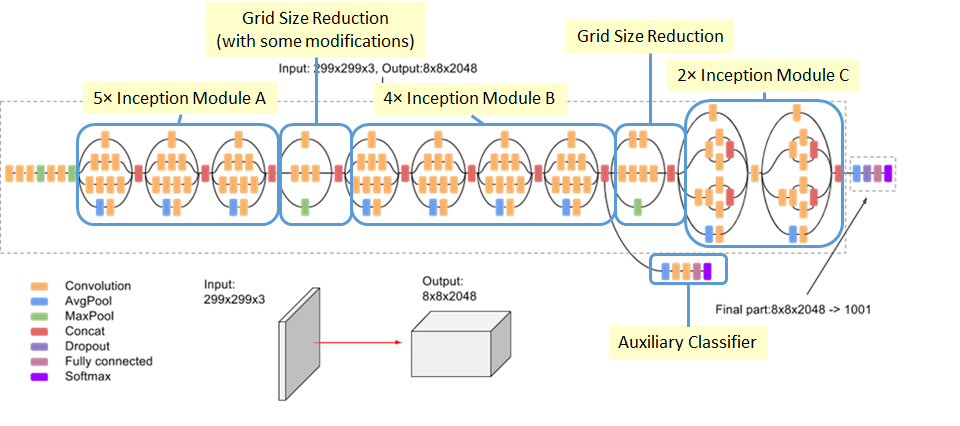
2. Design & Development: In this phase of the Incremental model of SDLC, the design of the system functionality and the development method are finished with success. When software develops new practicality, the incremental model uses style and development phase.

3. Testing: In the incremental model, the testing phase checks the performance of each existing function as well as additional functionality. In the testing phase, the various methods are used to test the behavior of each task.

4. Implementation: Implementation phase enables the coding phase of the development system. It involves the final coding that design in the designing and development phase and tests the functionality in the testing phase. After completion of this phase, the number of the product working is enhanced and upgraded up to the final system product.

* 1. **Prototype:-**

To analyze the Choice of the user, we have to use some transform learning algorithms which is the part of deep learning. We selected the inceptionV3 algorithm to design a model.



*Inception-v3 is a pre-trained convolutional neural network model that is 48 layers deep. The network has an image input size of 299-by-299. The model extracts general features from input images in the first part and classifies them based on those features in the second part.*

4. System weakness and requirements

• Problem and weakness of current system

• Functional And Non Functional Requirements

**4.1 Problem and Weaknesses of Current System**

Low quality of mammogram images and images with dense mass tissue it will make a system difficult to detect the cancer spot because the system is based on fixed threshold and one problem is also that with fixed threshold there will be a problem with flexibility of a system.

**4.2 Functional System Requirements**

1) For User (doctor):-

Download functionality

Search option for categories

Adding video or upload

Removing video

Manage query

**4.3 Non-Functional System Requirements**

1. Performance Requirements: -

The database shall be able to accommodate a minimum of 10,000 records of patient.

The software shall support use of multiple users at a time.

1. Safety Requirements:-

The database may get crashed at any certain time due to virus or operating system failure. Therefore, it is required to take the database backup.

1. Security Requirements:-

Some of the factors that are identified to protect the software from accidental or malicious access, use, modification, destruction, or disclosure are described below.

Keep specific log or history datasets, Assign certain functions to different modules ,Restrict communications between some areas of the program , Check data integrity for critical variables, Later version of the software will incorporate encryption techniques in the user/license authentication process.

1. Communication needs to be restricted when the application is validating the user or license. (i.e., using https)

5. Function of System

• Current Feature

• Data Dictionary

• Entity Relationship Diagram

• Class Diagram

• Use case Diagram

**5.1 Current Feature**

1. Filter the image and enhance the quality
2. Predict the cancer
3. Aware you about a future risk of a cancer

**5.2 Data dictionary**

* A data dictionary contains metadata i.e., data about the database. The data dictionary is very important as it contains information such as what is in the database, who is allowed to access it, where is the database physically stored etc. The users of the database normally don't interact with the data dictionary, it is only handled by the database administrators.

* The data dictionary in general contains information about the following –

• Names of all the database tables and their schemas.

• Details about all the tables in the database, such as their owners, their security constraints, when they were created etc.

• Physical information about the tables such as where they are stored and how.

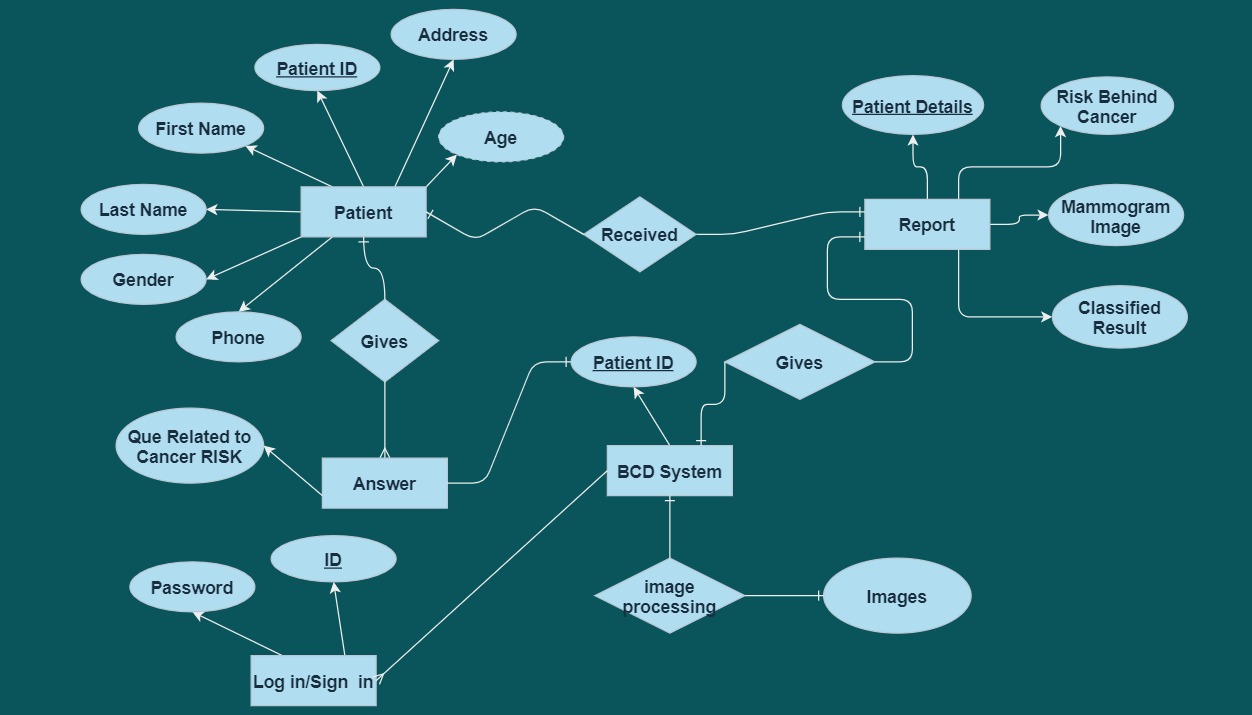
• Table constraints such as primary key attributes, foreign key information etc.

• Information about the database views those are visible.

**5.2.1 Patient details:-**

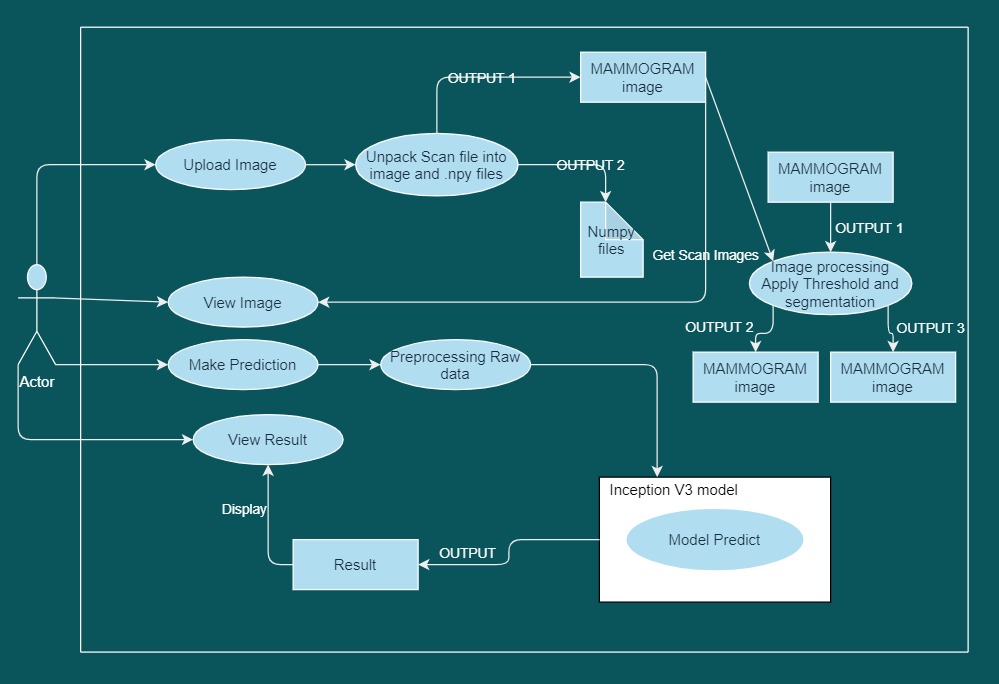
|  |  |  |  |
| --- | --- | --- | --- |
| Column\_name | Data type | Size | Constraint |
| Patient\_id | Int | 10 | Primary Key |
| Patient\_name | Varchar2 | 50 | Not null |
| Patient\_age | Int | 4 | Not null |
| Email\_id | Varchar2 | 20 | Not null |
| Patient\_pno. | Int | 13 | Not null |
| Patient\_address | Varchar2 | 100 | Not null |
| Prediction | Varchar2 | 3 | Not null |

**5.3 Entity Relationship Diagram:-**

****

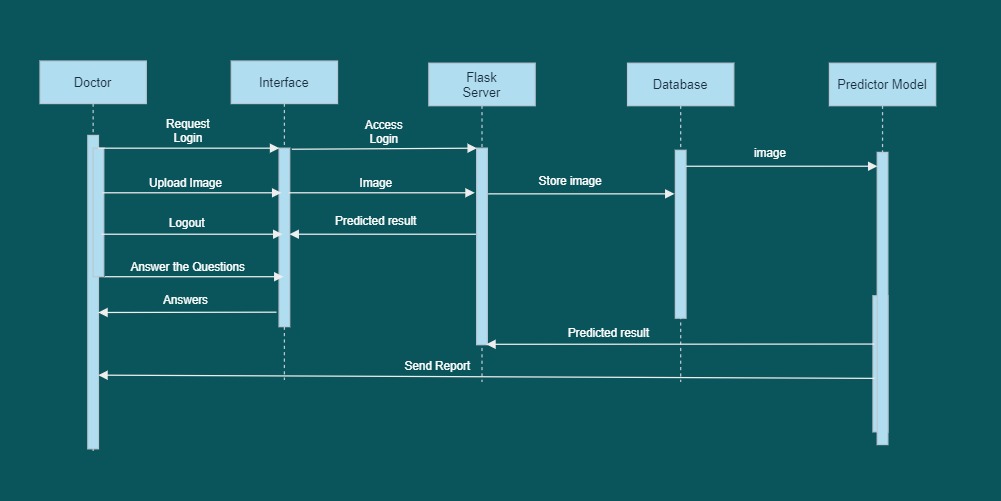
**FIG5.3.ENTITY RELATIONSHIP DIAGRAM**

**5.4 Use Case Diagram :-**



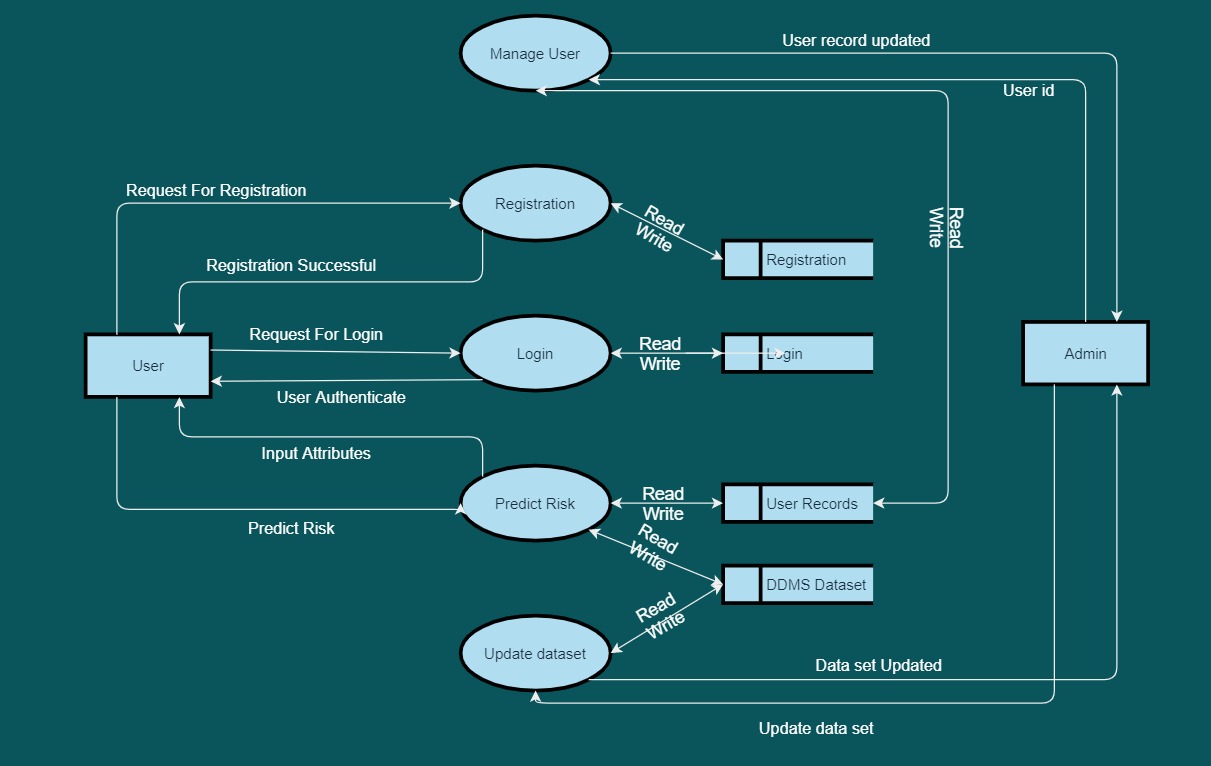
**FIG5.4 USE CASE DIAGRAM**

**5.5 Sequence diagram:-**

****

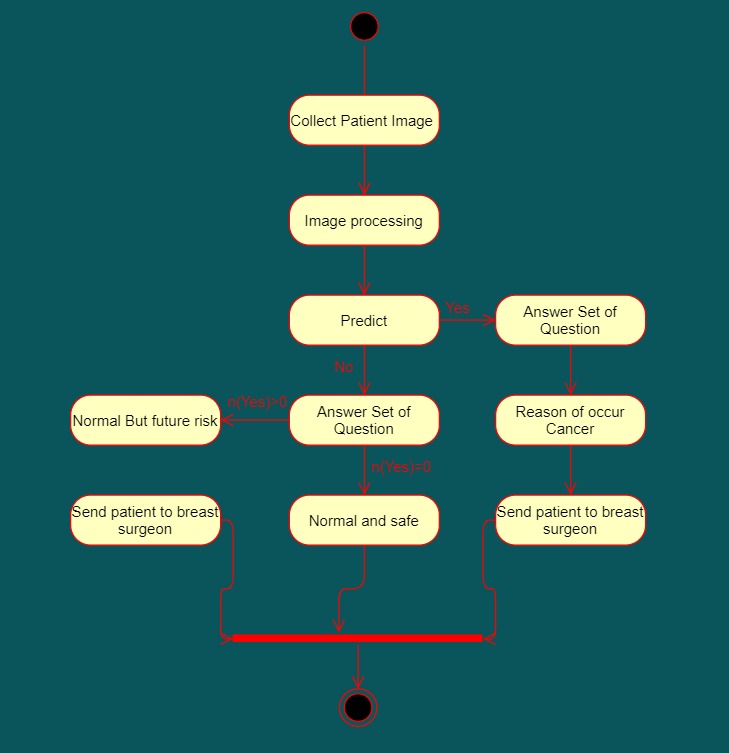
**FIG5.5.SEQUENCE DIAGRAM**

**5.6 DFD Diagram:-**

****

**FIG5.6.DFD DIAGRAM**

**5.7 State Diagram:-**

****

**FIG5.7.STATE DIAGRAM**

6. Estimation

• Estimation Cost of Project

• Project Timeline Chart

• Milestone and Deliverable

**6.1 Estimation Cost of project:-**

In our project in under a Semidetached system. Consider a software project using semidetached system mode with 10000 line of code. We will obtain estimation for this project as follows:

1. Effort Estimation:

E = a1(KLOC) Exp(a2)

Person-months E = 3.0(10) Exp(1.121)

E = 40 PM(approx)

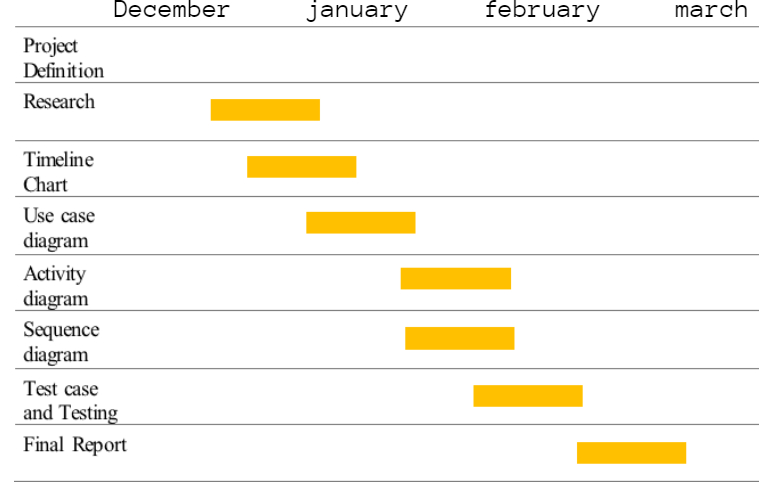
1. Duration Estimation:

D = b1(E) Exp(b2) months

D = 2.5(40) Exp(.35)

D = 9.09

**6.2 Project time line chart:-**



**6.3 Milestones and Deliverables:-**

**Month 1: Milestones & Deliverables**

|  |  |
| --- | --- |
| **Milestones** | **Deliverables** |
| Study about our model requirement, planning | Analysis Report |
| Understand a project definitions and basic terms and logic for Parameter Evaluation. | Analysis Report |
| Gathering the requirements of the project using different factfinding techniques and finding datasets. | Analysis Report |
| Still Continue with Requirement’s study. | Analysis Report |

**Month 2: Milestones & Deliverables**

|  |  |
| --- | --- |
| **Milestones** | **Deliverables** |
| Study Analysis and find the reason behind cancer | Analysis Report |
| System Design including various diagrams | SRS |
| Image processing using different method | Designing/Coding |
| Create deep learning model | Designing/Coding |
| Try to increase Accuracy | Designing/Coding |
| Study about model like roc auc curve, performance report etc. | Analysis model |
| Test with different images | Testing |

**Month 3: Milestones & Deliverables**

|  |  |
| --- | --- |
| **Milestones** | **Deliverables** |
| Integrating techniques of bootstrap and Flask | Designing/Coding |
| Database creation and testing | Designing/Coding/Testing |
| Client Module of Client Support System | Designing/Coding |
| Webapp Testing | Testing |
| Required changes after testing | Designing/Coding |

7. Coding and Accuracy view

• Deep learning model

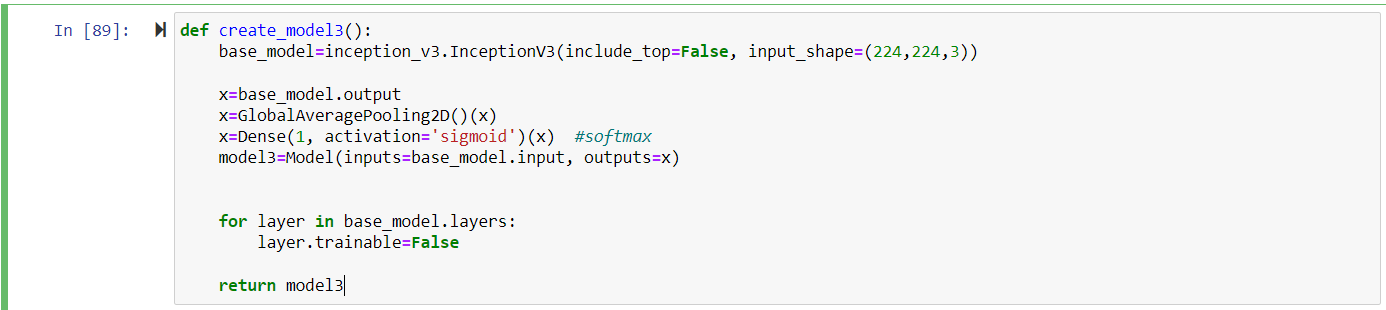
• Image Processing

• Compare Accuracy and Loss

• ROC curve

**7.1 Deep Learning model:-**

🡪 Inception V3

****

**7.2 Image Processing:-**

****

**7.3 Accuracy and Loss:-**

|  |  |  |
| --- | --- | --- |
| **Model** | **Accuracy** | **Loss** |
| **CNN** |  |  |
| **RESNET50** |  |  |
| **VGG16** |  |  |
| **Inception V3** |  |  |
| **CNN-SVM** |  |  |

**7.4 ROC curve:-**

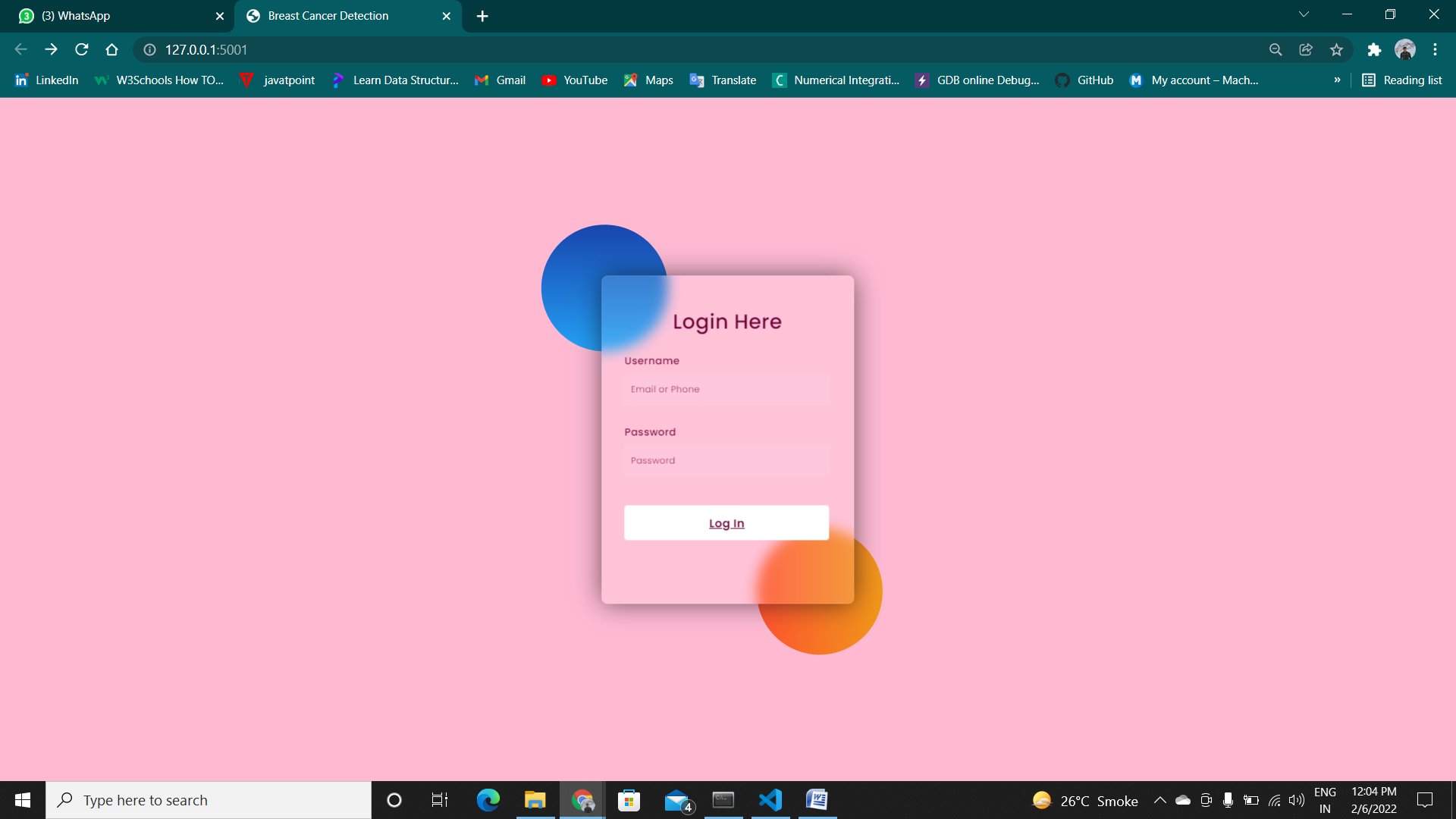
|  |  |
| --- | --- |
| **Model** | **ROC** |
| **Resnet50** |  |
| **VGG16** |  |
| **Inception V3** |  |

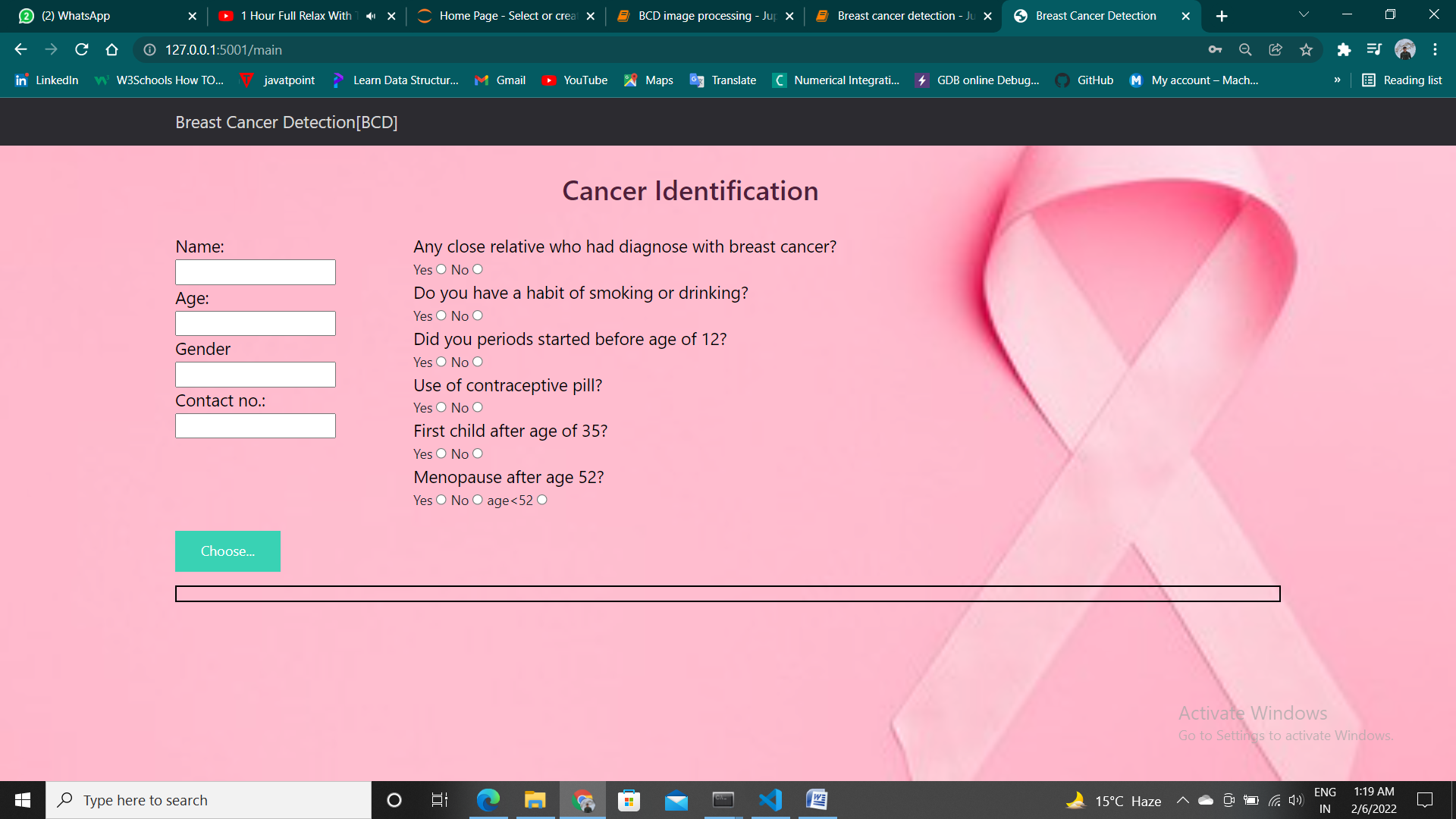
8. Web App view and Test Table

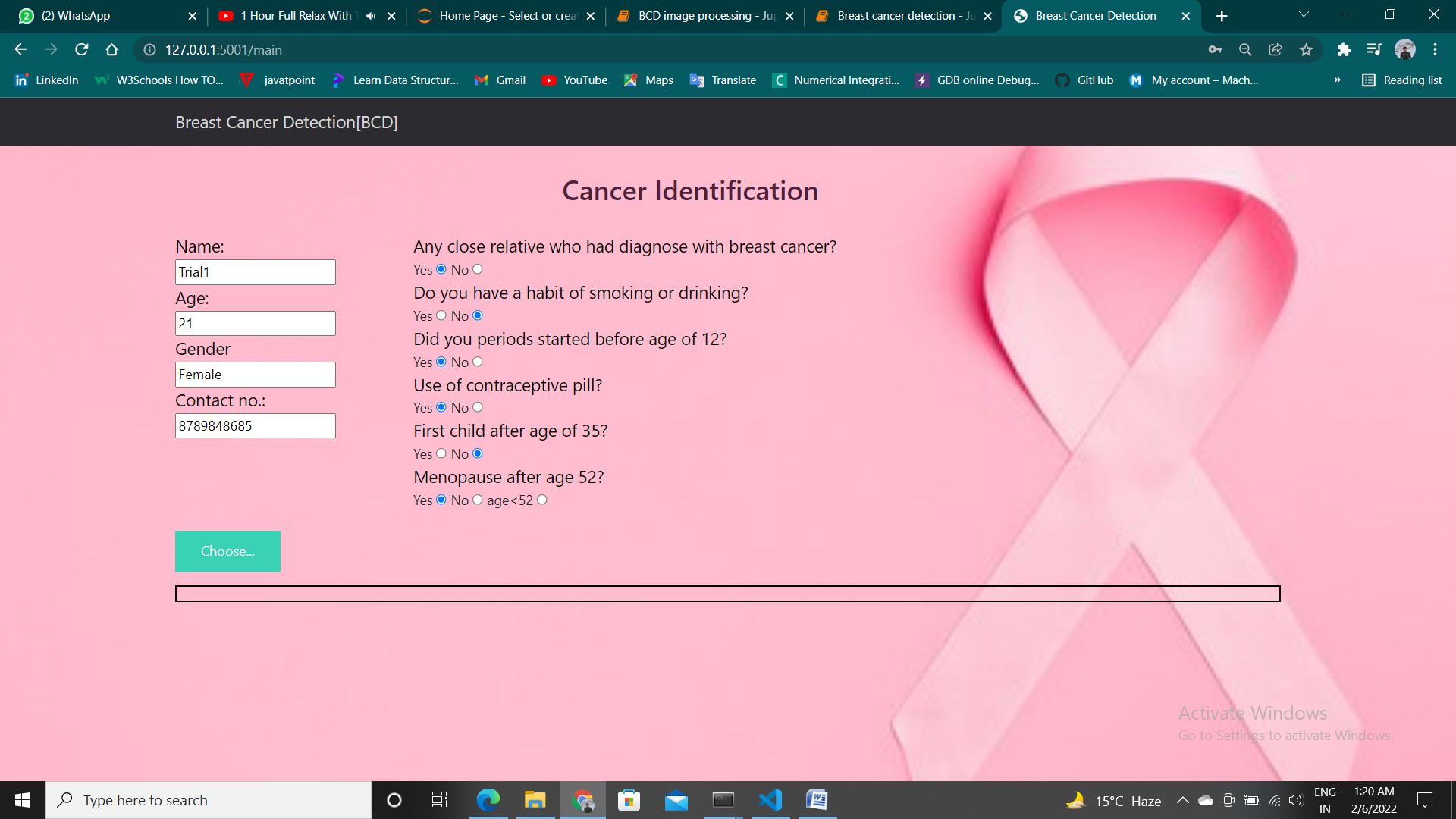
• View

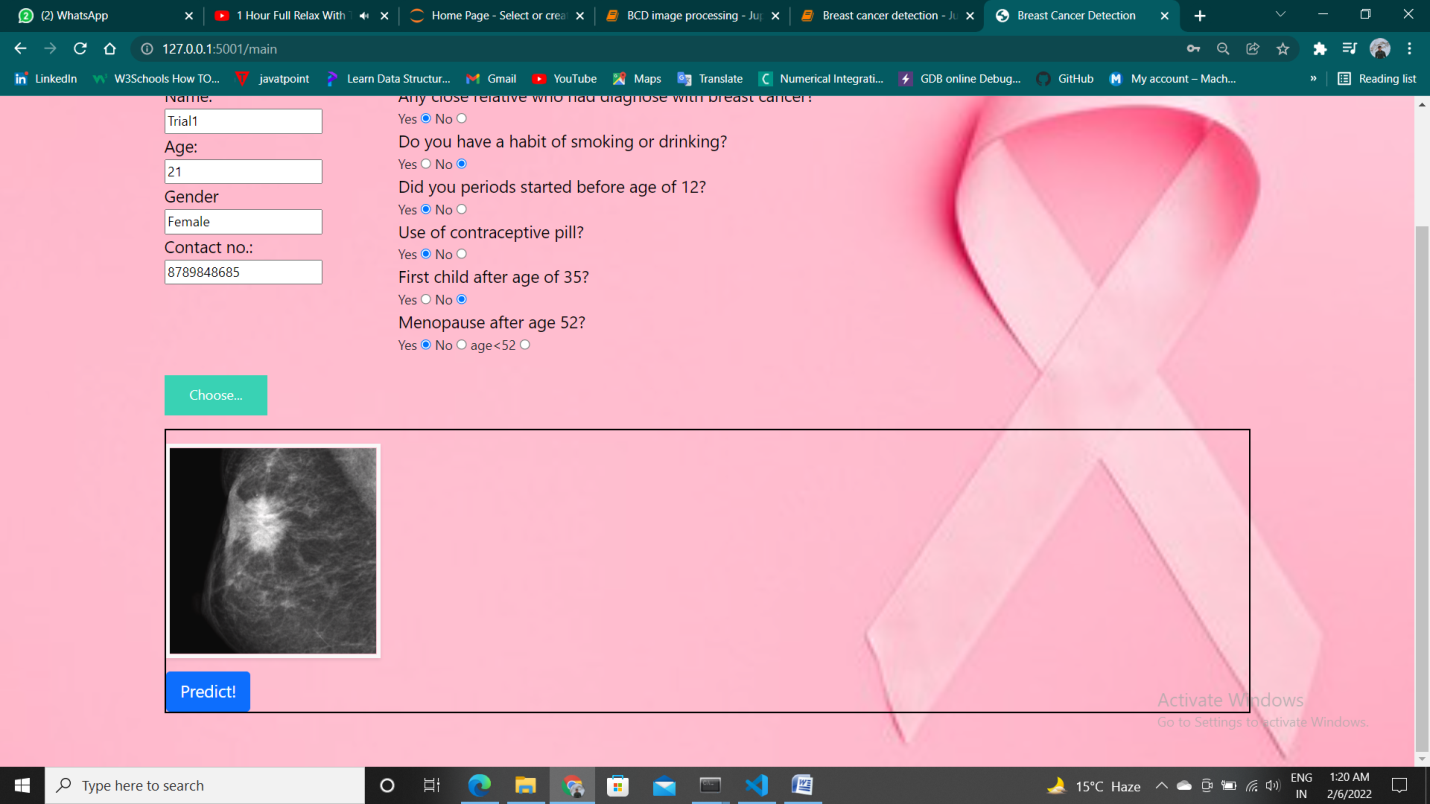
• Test Table

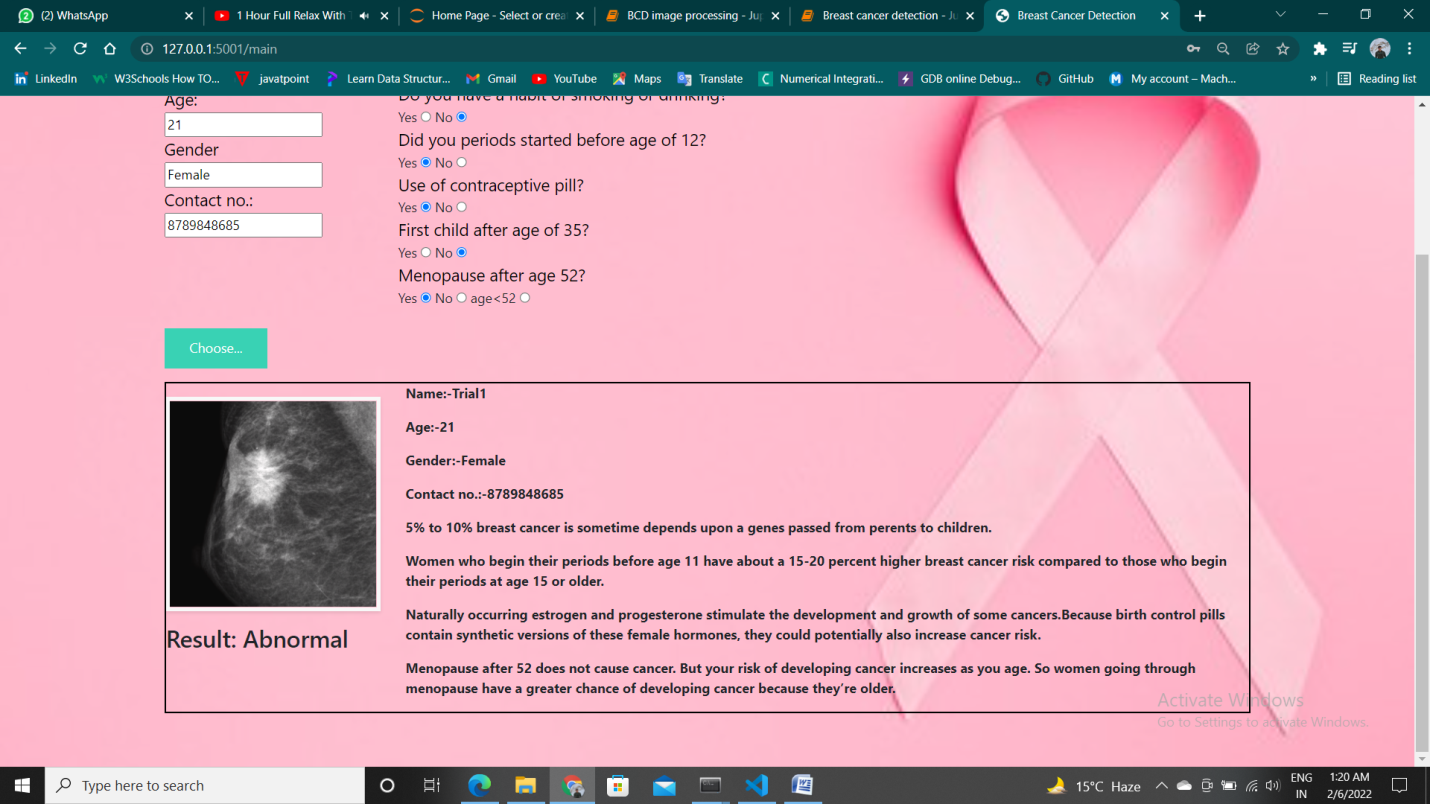
**8.1 Web App View:-**

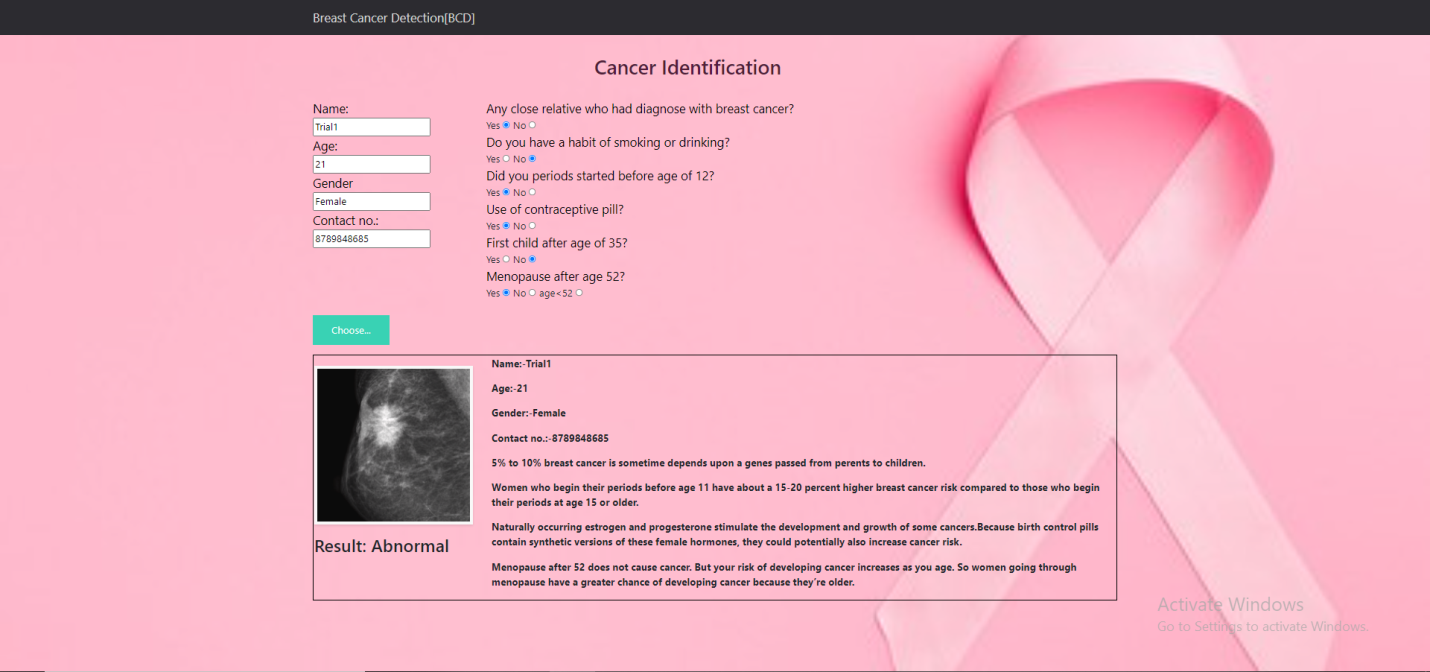
****

****

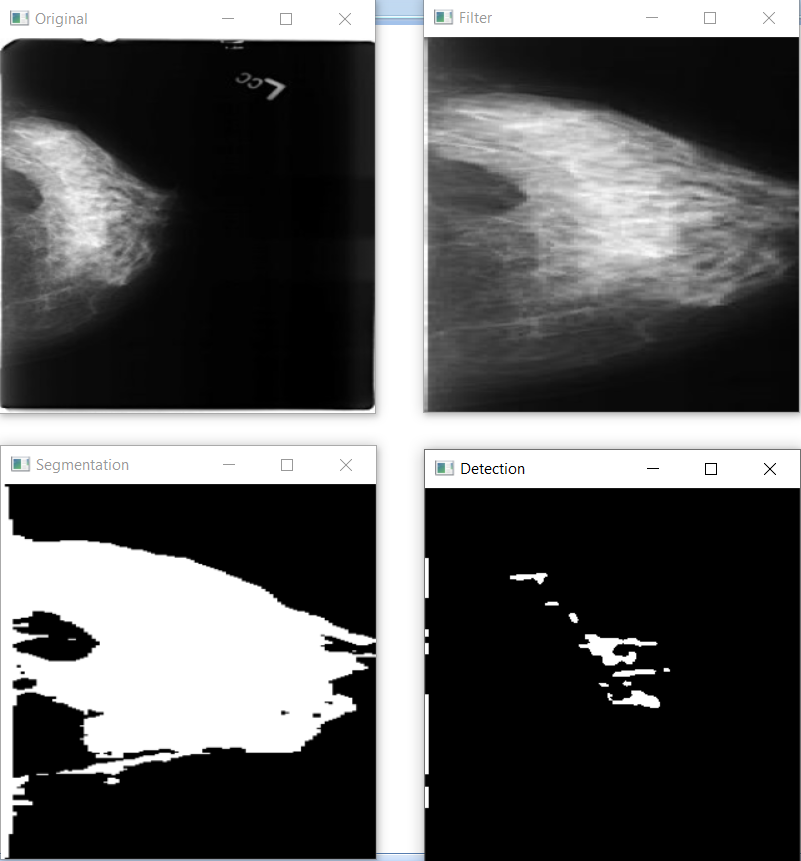
****

****

****

****

**Image processing**

****

**8.2 Test Table:-**

Test Case -1

Test Case ID: - ta\_id1

Test Priority (Low/Medium/High): - High Test

Title: - Verify Login with Valid User name and password

Description: - Test the Login Page

Pre-conditions: -User has valid Username and Password

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Steps** | **Test Step** | **Test data** | **Expected Result** | **Actual Result** | **Status** |
| 1 | Provide valid User name | User-ID = XYZ@gmail.com | User ID should be valid | User ID is valid | PASS |
| 2 | Provide valid Password | Password = Abc@123 | Password should be valid | Password Accepted | PASS |
| 3 | Click On login button | **--** | **--** | **--** | PASS |

Test Case -2

Test Case ID: - ta\_id2

Test Priority (Low/Medium/High): - High Test

Title: - Print Entered Data

Description: - Test show details

Pre-conditions: - User must enter correct details

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Steps** | **Test Step** | **Test data** | **Expected Result** | **Actual Result** | **Status** |
| 1 | Enter Detail | **--** | **--** | **--** | PASS |
| 2 | Click Predict Button | **--** | **--** | **--** | PASS |
| 3 | Show Enter Details | **--** | Showed Entered Details | Showed Entered Details | PASS |
| 4 | Show Enter Details | **--** | Showed Entered Details | Showed Blank Details | Fail |

Test Case -3

Test Case ID: - ta\_id3

Test Priority (Low/Medium/High): - High Test

Title: - Print Result

Description: - Show classified result

Pre-conditions: - User must Upload correct image

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Steps** | **Test Step** | **Test data** | **Expected Result** | **Actual Result** | **Status** |
| 1 | Upload Image | Check image extension | **--** | **--** | PASS |
| 2 | Upload Image | Check image extension | **--** | **--** | FAIL |
| 3 | Click Predict Button | **--** | **--** | **--** | PASS |
| 4 | Show Predicted Result | **--** | Showed Correct Classified Result | Showed Correct Result | PASS |
| 5 | Show Predicted Result | **--** | Showed Correct Classified Result | Showed False result | FAIL |

Test Case -4

Test Case ID: - ta\_id4

Test Priority (Low/Medium/High): - High Test

Title: - Image processing

Description: - Shows Processed images

Pre-conditions: - User must Upload correct image

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Steps** | **Test Step** | **Test data** | **Expected Result** | **Actual Result** | **Status** |
| 1 | Upload Image | Check image extension | **--** | **--** | PASS |
| 2 | Upload Image | Check image extension | **--** | **--** | FAIL |
| 3 | Click Predict Button | **--** | **--** | **--** | PASS |
| 4 | Show Processed Image | **--** | Showed Four Images | Showed Correct Images | PASS |
| 5 | Show Processed Image | **--** | Showed Four Images | Responding error | FAIL |

9. Conclusion

• Conclusion

• Bibliography

**9.1 Conclusion:-**

In our project BCD system takes the mammogram to detect a cancer with less time which helps doctors and patient to diagnose and detect breast cancer at a beginning stage and to get proper treatment within a safe time, which will help them to recover quickly after diagnosis.

**9.2 Bibliography:-**

1. Scikit learn: - <https://scikit-learn.org/stable/>

2. Html CSS: - <https://www.w3.org/standards/webdesign/htmlcss>

3. JS and Bootstrap: - <https://www.javascript.com/> , <https://getbootstrap.com/>

4. Flask: - <https://flask.palletsprojects.com/en/2.0.x/>

5. Dataset: - <https://www.kaggle.com/>

6. Open CV: - <https://opencv.org/>

7. Blogs:- <https://my.clevelandclinic.org/health/diseases/3986-breast-cancer> , <https://breastcancernow.org/information-support/have-i-got-breast-cancer/breast-cancer-causes> ,

<https://www.komen.org/breast-cancer/risk-factor/age-at-first-childbirth/>

8. Existing project paper: - <https://www.researchgate.net/publication/258439384_Mammogram_Images_Thresholding_For_Breast_Cancer_Detection_Using_Different_Thresholding_Methods>