

# **Breast Cancer Detection**

## **A MINI PROJECT REPORT**

*Submitted by*

<b>NIDIN</b>	<b>1905034</b>
<b>SANJAY PRATAP T K</b>	<b>1905046</b>
<b>YOKESH R S</b>	<b>1905062</b>
<b>ABUTHAGEER</b>	<b>2005201</b>

*in partial fulfilment for the award of the degree of*

**BACHELOR OF ENGINEERING**

*in*

**COMPUTER SCIENCE AND ENGINEERING**



**COIMBATORE INSTITUTE OF TECHNOLOGY**

*(Government Aided Autonomous Institution Affiliated to Anna University)*

**COIMBATORE-641 014**

**ANNA UNIVERSITY - CHENNAI 600 025**

**APRIL 2021**

## **BONAFIDE CERTIFICATE**

Certified that this mini project “**BREAST CANCER DETECTION**” is the bonafide work of **NIDIN S (1905034), SANJAY PRATAP T K (1905046), YOKESH R S (1905062), and ABUTHAGEER S (2005201)** under my supervision during the academic year 2021-2022.

**Dr.G.Kousalya, M.E.,Ph.D**

**HEAD OF THE DEPARTMENT,**

Department of CSE

Coimbatore Institute of Technology,

Technology,Coimbatore – 641 014.

**Dr.S.P.Abirami, Ph.D**

**ASSISTANT PROFESSOR,**

Department of CSE,

Coimbatore Institute of

Coimbatore – 641 014.

Certified that the candidates were examined by us in the project work viva voce examination held on .....

**Internal Examiner**

**External Examiner**

Place:

Date:

## **TABLE OF CONTENTS**

<b><u>CHAPTER NO.</u></b>	<b><u>TITLE</u></b>	<b><u>PAGE NO.</u></b>
	<b>ACKNOWLEDGEMENT</b>	<b>4</b>
	<b>ABSTRACT</b>	<b>5</b>
<b>1)</b>	<b>INTRODUCTION</b>	
	1.1. INTRODUCTION	<b>6</b>
	1.2. GOOGLE TREND ANALYSIS	<b>7</b>
<b>2)</b>	<b>LITERATURE SURVEY</b>	<b>8</b>

## ACKNOWLEDGEMENT

---

### ACKNOWLEDGEMENT

We express our sincere thanks to our Secretary **Dr.R.Prabhakar** and our Principal **Dr. Rajeswari A** for providing us a great opportunity to carry out our work. The following words are a small part to express our gratitude to them. This work is the outcome of their inspiration and product of plethora of their knowledge and rich experience.

We record the deep sense of gratefulness to **Dr.G.Kousalya**, Head of the Department of Computer Science and Engineering, for his encouragement during this tenure.

We equally tender our sincere thankfulness to our project guide **Dr.S.P.Abirami**, Assistant Professor, Department of Computer Science and Engineering, for his valuable suggestions and guidance during this course.

During the entire period of study, the entire staff members of the Department of Computer Science and Engineering have offered ungrudging help. It is also a great pleasure to acknowledge the unfailing help we have received from our friends.

It is matter of great pleasure to thank our parents and family members for their constant support and co-operation in the pursuit of this endeavor.

## ABSTRACT

---

### ABSTRACT

Globally, breast cancer is the most common cancer among women, and the most likely cause of female cancer deaths. High-income countries (HICs) have made the most progress in improving breast cancer outcomes. Between 1990 and 2014, breast cancer death rates dropped by 34% in the US attributable to the combination of improved earlier detection and effective adjuvant therapies. By contrast, breast cancer is an increasingly urgent problem in low- and middle-income countries (LMICs), where historically low incidence rates have been rising by up to 5% per year. The strides seen in HICs have not been mirrored in LMICs where optimal management strategies from wealthy countries cannot be fully implemented due to significant resource constraints related to limited personal resources, underdeveloped health care infrastructure, lack of pharmaceuticals and cultural barriers.

“Breast Cancer Detection ” is a simple and effective project that uses efficient machine learning algorithms for early detection and diagnosis of breast cancer by classifying either malignant or benign patients

The model primarily uses Wisconsin Dataset for training the model and classifies using Artificial Neural Network technique.

## INTRODUCTION

---

### CHAPTER 1

#### INTRODUCTION

##### 1.1. Introduction:

Breast Cancer is the most affected disease between women all over the world. All most 25% of all cancers with an estimated 1.67 million new cancer cases diagnosed in 2012 and its incidence is increasing day by days. It is also said that Breast cancer is the second leading cause of death for women worldwide.

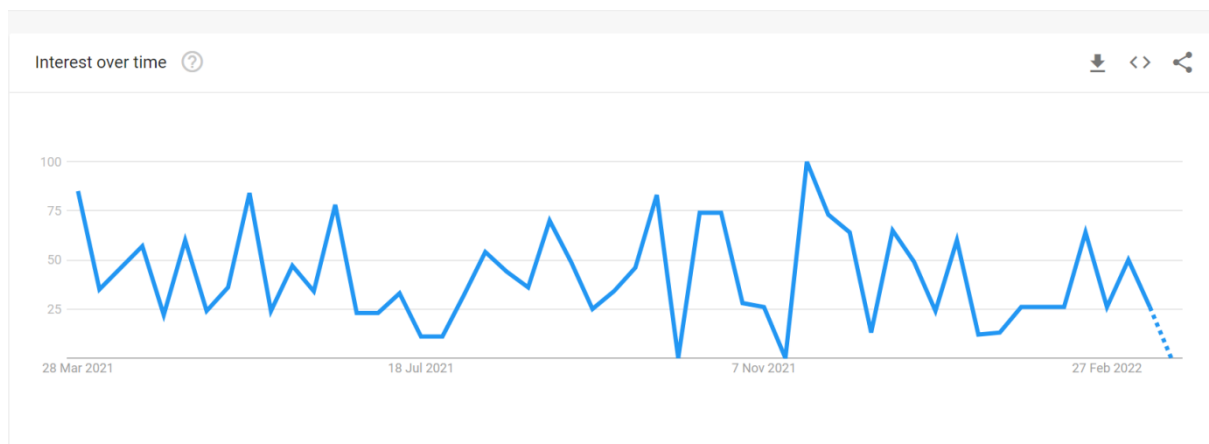
Early detection of cancer can reduce the risk of deaths for cancer patients. To increase the survival rate the early diagnosis of breast cancer and a trustworthy detection model is required.

The goal of the research is to identify and classify Malignant and Benign patients by proposing a simple and efficient model for early detection of breast cancer with minimal error percentage.

This model uses an Artificial neural network and produces Accuracy of 98% and f1 score as 98% - Benign and 97% - Malignant .

## 1.2. Google trend analysis:

The diagram shown below displays the google trend analysis about the domain of this project.



## LITERATURE SURVEY

---

### CHAPTER 2

#### LITERATURE SURVEY

##### **2.1 Breast Cancer Detection using K-nearest Neighbor Machine Learning Algorithm (2016)**

**Authors:** Moh'd Rasoul Al-hadidi, Abdulsalam Alarabeyyat, Mohannad Alhanahnah.

##### **Description**

In this paper, we proposed a new method to detect the breast cancer with high accuracy. This method consists of two main parts, in the first part the image processing techniques are used to prepare the mammography images for feature and pattern extraction process. The extracted features are utilized as an input for a two types of supervised learning models, which are Back Propagation Neural Network (BPNN) model and the Logistic Regression (LR) model. The LR model resulted in Accuracy of 45% with 750 features and BPNN with 93.7 % which is comparatively extremely good but with only 240 features.

##### **Advantages**

1. They have used image processing to obtain the features for training the prediction model and used have BPNN and Logistic regression.

##### **Disadvantages**

1. The accuracy rate of the model greatly affects with the increase in the features and the maximum accuracy they could bring out was 93.7% only.



## **2.2 Efficient Approaches for Accuracy Improvement of Breast Cancer Classification Using Wisconsin Dataset (2017)**

**Authors:** Shajib Ghosh ,Jubaer Hossain, Dr.Shaikh Anowarul Fattah , Dr. Celia Shahnaz, Asir Intisar Khan

### **Description**

This paper deals with different statistical and deep learning analysis of Wisconsin Breast Cancer Database for improving the accuracy in detection and classification of breast cancer based on different attributes. Applying Naïve Bayes, SVM, Logistic Regression, KNN, Random Forest, MLP and CNN classifiers, higher accuracy is obtained which is up to 98% to 99%.

### **Advantages**

1. The proposed goes through a lot of classification models and compares the accuracy for each.

### **Disadvantages**

1. Harder to optimize the network as convergence time significantly increases.

## 2.3 Breast Cancer Diagnosis Using deep learning Algorithm (2018)

**Authors:** Naresh Khuriwal , Dr. Nidhi Mishra

### **Description:**

In this paper they've proposed the deep learning method convolutional neural network that mostly used for classification of images dataset. It is basically divided into three parts first collection of dataset and applied pre-processing algorithm for scaled and filter data then split dataset for training and testing purpose and generate some graph for visualization data. They've concluded deep learning technology is a good way for diagnosis breast cancer with Wisconsin Breast Dataset. This database provides 569 rows and 30 features in the dataset. In this paper they've **just used 11 features** for diagnosis .After the implementation they've achieved 98.67% accuracy.

### **Advantages:**

1. Uses Convolutional Neural network faster and better Accuracy

### **Disadvantages:**

1. Uses only 11 features of given 30 features.

## **2.4 Breast Cancer Detection Based on Deep Learning Technique (2019)**

**Author:** Nur Syahmi Ismail and Cheab Sovuthy.

### **Description:**

In this paper, they've used deep learning technique using VGG16 and ResNet50 network to classify between normal tumour and abnormal tumour using IRMA dataset and have implemented for normal and abnormal breast cancer detection. The classification methods were evaluated using three performance evaluations which are precision, recall, and accuracy rate. The best result of classification accuracy was VGG16 with 94%. Compared to ResNet50 with 91.7% in term of accuracy.

### **Disadvantage:**

1. Usage of already available Model
2. Poor accuracy.

## **2.5 Breast Cancer Prediction Using Machine Learning (2020)**

**Author:** Ramik Rawal.

### **Description:**

In this paper they have used four algorithms SVM, Logistic Regression, Random Forest, and KNN to predict the outcome the breast cancer. It broadly speaks about three domains. First domain is prediction of cancer before diagnosis, second domain is prediction of diagnosis and treatment and third domain focuses on outcome during treatment. The accuracy obtained by SVM (97.13%) is better than the accuracy obtained Naïve Bayes and k-NN that have an accuracy that varies between 95.12 % and 95.28 %. It was also clear with SVM having highest value of correctly classified instances and the lower value of incorrectly classified instances than the other classifiers.

### **Advantage:**

1. Better preprocessing Technique
2. Average accuracy of different models was high.

### **Disadvantage:**

1. Efficiency of benign class and malignant class got the highest for two different algorithms such as SVM for benign class and K-NN for malignant class