# Cancer Prediction

(A Breast cancer prediction)

## A Project Work Synopsis

*Submitted in the partial fulfillment for the award of the degree of*

## BACHELOR OF ENGINEERING
## IN

## CSE – Big Data and Analytics

### Submitted by:

**Aditya Yadav**
**20BCS3865**

**Pritiranjan Biswal**

**20BCS3896**

**Shubham Joshi**
**20BCS3906**

**Lisa Singh**
**20BCS3911**

### Under the Supervision of:

**Neha Sharma**



## CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

### PUNJAB
**11/2022**

# Table of Contents

# ABSTRACT

This paper presents a research paper for cancer prediction (especially breast cancer). This research paper will cover the analysis we have done after gathering lot of databases.

Breast cancer is a menacing cancer, primarily affecting women. Continuous research is going on for detecting breast cancer in the early stage as the possibility of cure in early stages is bright. There are two main objectives of this current study, first establish statistics for breast cancer and second to methodologies which can be helpful in the early stage detection of the breast cancer based on previous studies

There are several steps performed after which we analyzed the data.

First of all, we gathered data from different sources. Extracting data from different databases enhances the chances of accuracy and promotes less error.

Secondly, we used that extracted data to visualize it and draw some common information.

In third step, we used tools like Weka to visualize the graphs to conclude points like various symptoms experienced.

Lastly, all these information lead us to write research paper.

# INTRODUCTION

*Cancer* is a disease in which cells in the body grow out of control. Breast cancer is a disease in which cells in the breast grow out of control. There are different kinds of breast cancer. Breast cancer can occur in women and rarely in men.

Symptoms of breast cancer include a lump in the breast, bloody discharge from the nipple and changes in the shape or texture of the nipple or breast.

Its treatment depends on the stage of cancer. It may consist of chemotherapy, radiation, hormone therapy and surgery.

## Symptoms:

Different people have different symptoms of breast cancer. Some people do not have any signs or symptoms at all.

Some warning signs of breast cancer are—

- New lump in the breast or underarm (armpit).
- Thickening or swelling of part of the breast.
- Irritation or dimpling of breast skin.
- Redness or flaky skin in the nipple area or the breast.
- Pulling in of the nipple or pain in the nipple area.
- Nipple discharge other than breast milk, including blood.
- Any change in the size or the shape of the breast.
- Pain in any area of the breast.

## Stages of Cancer:

The earliest stage breast cancers are stage 0 (carcinoma in situ). It then ranges from stage I (1) through IV (4). As a rule, the lower the number, the less the cancer has spread. A higher number, such as stage IV, means cancer has spread more. And within a stage, an earlier letter means a lower stage.

## How is the stage determined?

The staging system most often used for breast cancer is the American Joint Committee on Cancer (AJCC) **TNM system**. The most recent AJCC system, effective January 2018, has both clinical and pathologic staging systems for breast cancer:

- The **pathologic stage** (also called the **surgical stage**) is determined by examining tissue removed during an operation.

- Sometimes, if surgery is not possible right away or at all, the cancer will be given a **clinical stage** instead. This is based on the results of a physical exam, biopsy, and imaging tests. The clinical stage is used to help plan treatment. Sometimes, though, the cancer has spread further than the clinical stage estimates, and may not predict the patient's outlook as accurately as a pathologic stage.

In both staging systems, 7 key pieces of information are used:

- **The extent (size) of the tumor (T):** How large is the cancer?  Has it grown into nearby areas?

- **The spread to nearby lymph nodes (N):** Has the cancer spread to nearby lymph nodes? If so, how many?

- **The spread (metastasis) to distant sites (M):** Has the cancer spread to distant organs such as the lungs or liver?

- **Estrogen Receptor (ER) status:** Does the cancer have the protein called an estrogen receptor?

- **Progesterone Receptor (PR) status:** Does the cancer have the protein called a progesterone receptor?

- **HER2 status:** Does the cancer make too much of a protein called HER2?

- **Grade of the cancer (G):** How much do the cancer cells look like normal cells?

**Details of the TNM staging system**

Numbers or letters after T, N, and M provide more details about each of these factors. Higher numbers mean the cancer is more advanced. The categories below use the pathologic (surgical) definitions.

T categories for breast cancer

T followed by a number from 0 to 4 describes the main (primary) tumor's size and if it has spread to the skin or to the chest wall under the breast. Higher T numbers mean a larger tumor and/or wider spread to tissues near the breast.

**TX:** Primary tumor cannot be assessed.

**T0:** No evidence of primary tumor.

**Tis:** Carcinoma in situ (DCIS, or Paget disease of the breast with no associated tumor mass)

**T1** (includes T1a, T1b, and T1c)**:** Tumor is 2 cm (3/4 of an inch) or less across.

**T2:** Tumor is more than 2 cm but not more than 5 cm (2 inches) across.

**T3:** Tumor is more than 5 cm across.

**T4** (includes T4a, T4b, T4c, and T4d)**:** Tumor of any size growing into the chest wall or skin. This includes inflammatory breast cancer.

# METHODOLOGY

There are several steps performed after which we analyzed the data.

- First of all, we gathered data from different sources. Extracting data from different databases enhances the chances of accuracy and promotes less error.

- Secondly, we used that extracted data to visualize it and draw some common information. In this step, we analyzed various charts and graphs of few decades to predict the data for next few years. Like: the age at which it is most likely to occur, common cause, best cure.

- In third step, we used tools like Weka to visualize the graphs to conclude points like various symptoms experienced.

In this synopsis the breast cancer prediction-based research paper, is created to predict breast cancer by analyzing the data of previous few decades. Data collected from various source made it less noisy and reliable.

# REFERENCES

[1] https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/stages-of-breast-cancer.html

[2] https://www.cancer.net/cancer-types/breast-cancer/stages

[3] https://www.mskcc.org/cancer-care/types/breast/diagnosis/stages-breast

[4] https://www.komen.org/breast-cancer/diagnosis/stages-staging/

[5] https://my.clevelandclinic.org/health/diseases/3986-breast-cancer

[6] https://www.cancer.gov/types/breast/patient/breast-treatment-pdq

[7] https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470

[8] https://breastcancernow.org/information-support/have-i-got-breast-cancer/signs-symptoms-breast-cancer

[9] https://www.cancercenter.com/cancer-types/breast-cancer/symptoms

[10] https://www.webmd.com/breast-cancer/understanding-breast-cancer-symptoms