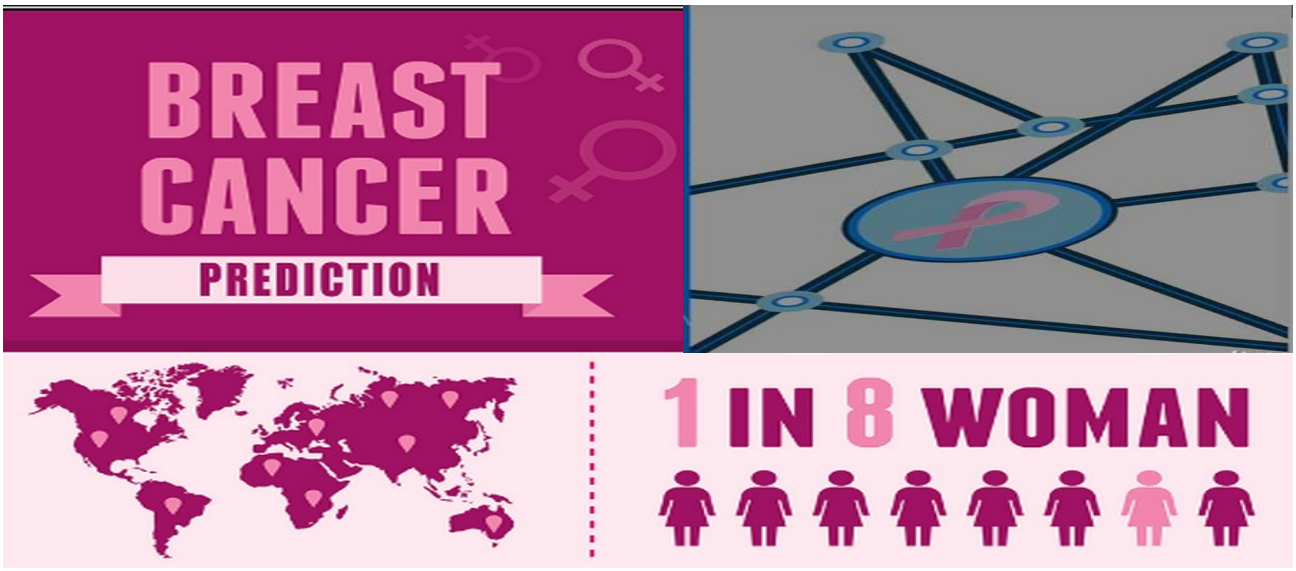
Breast Cancer Prediction &Classification



Project OverView

Breast cancer, affecting 2.2 million individuals annually, is the most prevalent cancer worldwide. Early diagnosis significantly improves survival rates, and a key challenge lies in distinguishing between benign and malignant tumors.

Machine learning (ML) techniques have demonstrated the potential to dramatically enhance diagnostic accuracy, with research revealing a 97% success rate compared to 79% by experienced doctors.

Our upcoming project focuses on leveraging ML to classify breast tumors, aiming to increase diagnostic precision. By integrating advanced computational models, we seek to contribute to more effective interventions and align with the broader trend of incorporating technology to improve healthcare outcomes.

Research Questions

Predicting if the cancer diagnosis is benign or malignant based on several observations/features

Datasets to be used

Analysis Breast Cancer Prediction Dataset | Kaggle

Breast Cancer Dataset (kaggle.com)

Project Requirements:

Tools and Library (Min 2 should be used)

***Scikit-learn and/or another machine learning library.**

***Pandas/Python/ Matplotlib, HTML/CSS/Bootstrap, JavaScript**

*** JavaScript Plotly JavaScript Leaflet**

***Tableau SQL Database MongoDB Database,* Google Cloud SQL Amazon AW**

Rough Breakdown of Tasks:

- Data and data delivery (Khadija)
- Back end (ETL) and visualizations (Khadija&Yosieph)
- GitHub repository, data analysis Report and the README file (Khadija)
- Creating the presentation (Yosieph