PCET's Pimpri Chinchwad University, Pune

Department of Computer Science and Engineering

SY B Tech (B)

Breast Cancer Detection

Using Machine Learning

Presented by

Prashant Gupta (B-17)

SOE23201010438

Guided by

Dr. Sachin Jadhav

Introduction

Breast Cancer Overview:

- Breast cancer is one of the most common cancers affecting mens and womens worldwide.
- Early detection is crucial for improving survival rates and treatment outcomes.

Challenges in Detection:

- Traditional diagnostic methods, such as mammography, have limitations, including false positives and negatives.
- Manual analysis of medical images is time-consuming and subject to human error.

Data Utilized:

- Medical data, including 29 features of tumor for identification of cancer .
- Breast Cancer Wisconsin data from biopsy samples (Kaggle).

Objectives of ML Model:

- Classify tumors as benign or malignant.
- Assist radiologists in interpreting medical reports.

Benefits:

- Improved diagnostic accuracy and reduced human error.
- Ability to process and analyze large volumes of data efficiently.
- Support personalized treatment plans based on predictive analytics.

LITERATURE REVIEW

Overview:

- Breast cancer is a significant global health concern.
- Mammograms are widely used but have limitations (false positives/negatives).
- Machine Learning (ML) improves accuracy in detection.

ML Techniques Used:

- Logistic Regression
- Decision Trees
- Random Forest

Identified Gaps:

- Heavy reliance on imaging datasets.
- Lack of accessible, user-friendly detection systems.

METHODOLOGY

MODEL DEVELOPMENT:

- Developed a breast cancer detection model using python and trained it on the dataset.
- The model was serialized into a .pkl file to save its state.
- Integrated this model into the backend of a Flask based web app for easy interaction.

Flask Based Web Application:

- The web app allows for manual entry of 29 features or the upload of a PDF file to extract these features.
- Based on the 29 features, the model predicts whether the cancer is malignant or benign.

DESIGN & IMPLEMENTATION

ML Algorithms Implemented:

- Logistic Regression (98% accuracy)
- Decision Tree Classifier (100% accuracy, risk of overfitting)
- Random Forest Classifier (99% accuracy)

Data Processing:

- Exploratory Data Analysis (EDA)
- Encoding & Scaling (StandardScaler applied)
- Train-Test Split (80-20% division)

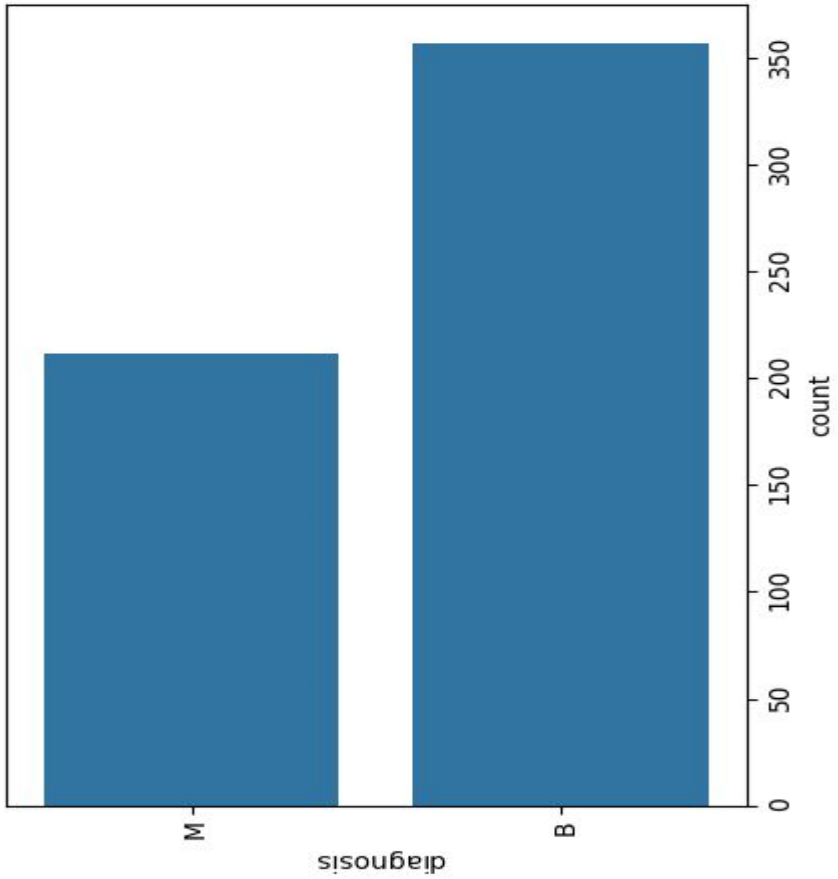
System Implementation:

- Flask-based web application for interactive use.
- Supports manual feature input and PDF upload for automated processing.

RESULT & PERFORMANCE

ALGORITHMS & ACCURACY:

- Logistic Regression: 98.90%
- Decision Tree: 100%
- Random Forest: 99.78%



IMPACT, BENEFITS & CHALLENGES

Impact & Benefits:

- **Enhanced Accuracy:** ML reduces human error in cancer diagnosis.
- **Timely Interventions:** Faster and more reliable screening process.
- **Resource Optimization:** Helps medical professionals focus on complex cases.

Challenges:

- **Data Extraction Issues:** Variability in PDF formats can lead to misinterpretation.
- **Overfitting Concerns:** High training accuracy may not generalize well.
- **Integration Complexity:** Seamless backend-frontend communication needs optimization.

Conclusion

- Machine Learning significantly improves breast cancer detection.
- Implementation of a web-based ML tool enhances accessibility.
- User friendly web application for early detection of cancer.
- Addressing existing challenges can further refine model effectiveness.

Future work:

- Plan to introduce an image upload feature for automatic cancer detection from images (eg. mammogram images).

References:

- Udemy: Breast Cancer Detection Using Machine Learning.
- Research works from Research Gate, IEEE Xplore, SciHub.
- Various online resources and datasets.

PCET's Pimpri Chinchwad University, Pune

THANK YOU

A Breast Cancer Detection model and web application using ML

Presented by
Prashant Gupta
CSSB017

Guided by
Dr. Sachin Jadhav