Towards A Machine Learning based Platform for Diseases Detection : Case of Breast Cancer.

Amen Allah BERREJEB
amenallah.berrejeb@etudiant-enit.utm.tn

Donies HADDAD
donies.haddad@etudiant-enit.utm.tn

Wafa MEFTEH
wafa.mefteh@enit.utm.tn

RISC-ENIT Laboratory, National Engineering School of Tunis, University of Tunis-Manar, Tunisia.

***Abstract*—This research paper focuses on the development of a web application for predicting breast cancer using machine learning techniques. The study involves the analysis of patient and tumor data, leveraging machine learning models to predict breast cancer diagnosis. The primary objective is to create an easy-to-use web application that provides users with a convenient way to assess their risk of developing breast cancer and help both doctor and individuals using different dataset for each one. The paper outlines the methodology employed, including data analysis, feature selection, model training, and web application development.**

***Index Terms*—Breast Cancer, Machine Learning, Prediction, Web Application, Data Cleaning, Web Modeling.**

I. Introduction

Breast cancer is the leading cause of cancer death among women worldwide. It is the leading cause of cancer death among women in developing countries and the second leading cause among women in developed countries [1]–[3]. In developing countries like Tunisia, the incidence of this cancer is still increasing.

The International Organization for research on Cancer: World Health Organisation: shows that there are 3,092 new cases of breast cancer in women and 986 deaths from this disease in 2020 [4], [5]. As statistics show, cancer is an aggressive disease with a low median survival rate. Moreover, the treatment process is lengthy and very expensive due to its high recurrence and mortality rates. Hence, early diagnosis of cancer is essential to improve patient survival.

Advances in statistics and computer engineering over the years have encouraged many scientists to apply computer methods such as multivariate statistical analysis to analyze the prognosis of the disease. Furthermore, artificial intelligence, especially machine learning and deep learning, has found popular applications in clinical cancer and health research in recent years. These methods are nowadays playing a major role in improving the accuracy of predictions of cancer susceptibility, recurrence and survival [6]. This paper presents a machine learning based prototype for disease detection. More specifically, the case of breast cancer.

II. Breast Cancer Dtection

Cancer is one of the most dreaded diseases in the history of mankind. It never ceases to test the ability of science to fight it, since it has a unique characteristic: it comes from ourselves. Breast cancer starts in breast cells. The cancerous (malignant) tumor is a group of cancer cells that can invade neighboring tissues and destroy them. It can also spread (metastasize) to other parts of the body [7]. There are several types of breast cancer. There are in breast cancers in order of frequency: infiltrating ductal carcinomas, infiltrating lobulars, papillae and mucinous. This varies according to their histological type, that is to say from which cells the development ofbreast cancer takes place [8].

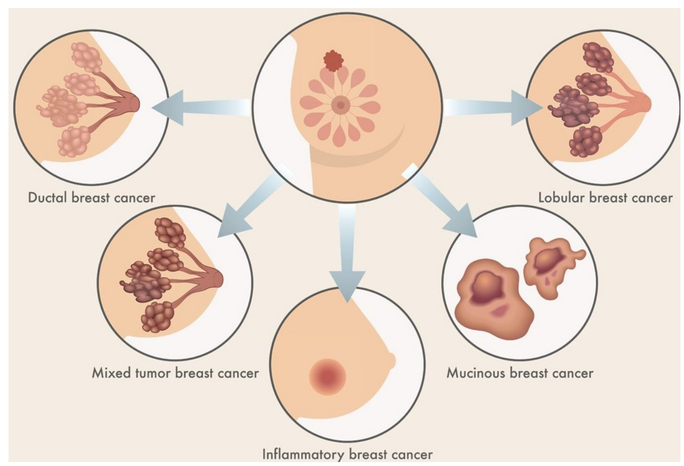

Fig. 1. Types of breast cancer [9].

Breast cancer detection typically involves a combination of screening tests and diagnostic tests. Screening tests are used to identify potential cases of breast cancer in individuals who do not have any symptoms. The most common screening test for breast cancer is mammography, which uses low-dose X-rays to produce images of the breast tissue. Other screening tests include breast MRI (magnetic resonance imaging) and clinical breast exams [10].

III. Machine Learning for Medical Purposes

Machine learning addresses the question of how to build computers that improve automatically through experience. It is one of today's most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the core of artificial intelligence and data science [11]. There are three main types of machine learning: supervised learning, unsupervised learning, and reinforcement learning. Table IV gives a useful summary and a reference in choosing the appropriate machine learning technique for a particular task [11].

TABLE I
TASK-BASED COMPARISON OF REINFORCEMENT, UNSUPERVISED, AND SUPERVISED LEARNING TECHNIQUES

Category	Task	Algorithms
Reinforcement Learning	Control	Q-learning, SARSA
Unsupervised Learning	Clustering	K-means, DBSCAN
Supervised Learning	Classification	Decision trees, SVM

Table II compares different classification algorithms in terms of their advantages and disadvantages. It was helpful for choosing the most suitable algorithm for a given classification problem.

TABLE II
COMPARISON OF DIFFERENT CLASSIFICATION ALGORITHMS FOR BOTH SUPERVISED AND UNSUPERVISED LEARNING

Algorithm	Advantages	Disadvantages
Support Vector Machines	Effective in high dimensional spaces, can handle non-linear decision boundaries and can be used for both classification and regression	Training can be time-consuming and complex ,can be sensitive to the choice of kernel function and not suitable for large datasets.
Logistic Regression	Can handle non-linear decision boundaries with polynomial features and can be used for classification tasks	Assumes a linear relationship between features and target variable , not suitable for many features and Can be affected by outliers
k-Nearest Neighbors	Simple and easy to implement and can handle non-linear decision boundaries	Can be computationally expensive during testing ,sensitive to the choice of k value and Performance can be affected by irrelevant or noisy features
K-Means Clustering	Simple and fast clustering algorithm that can handle large datasets	Requires the number of clusters to be specified and can be sensitive to initial centroid placement
Decision Trees and Random Forest	Easy to interpret and visualize and can give good results even with very large datasets	Can overfit on noisy or complex data , performance can be affected by irrelevant features and Random Forest can be computationally expensive during training
Hierarchical Clustering	Does not require the number of clusters to be specified and can create a dendrogram to visualize the clustering structure	Can be computationally expensive and can create long chains of clusters in certain situations
Principal Component Analysis	Can reduce the dimensionality of large datasets while preserving the most important information	May not be suitable for nonlinear relationships between variables and can result in loss of interpretability

Machine learning plays an important role in the medical field. It is an artificial intelligence (AI) technique that allows computers to learn from data, without being explicitly programmed to perform a specific task. It's used in many medical diagnostic and treatment applications. Here are some examples of machine learning applications in this field:

- **Diagnosis:** To analyze clinical data such as symptoms, lab tests, and diagnostic images. It can also be used to predict patients' risk of disease, using risk factors such as age, medical history, lifestyle habits, family history and genetic information.
- **Cancer Detection:** To analyze mammogram to detect signs of cancer.
- **Personalized Treatment:** To develop personalized treatment plans for patients using clinical data, medical history, lifestyle habits and genetic information.
- **Detection of Rare Diseases:** To help diagnose rare diseases by analyzing patients' clinical and genetic data. Machine learning models can also be used to identify patients who may benefit from experimental treatments.
- **Health Monitoring:** To monitor the health status of patients, using wearable sensors, health tracking devices and other technologies.

This preliminary study has provided a foundation for understanding the key components of our breast cancer prediction project. We have discussed the medical background of breast cancer and discuss its various types, the complex process of cancer detection, and the potential for machine learning to enhance diagnostic accuracy and improve patient outcomes. Moving forward, we will provide an overview of related works and our specific contributions to the field.

IV. OVERVIEW ABOUT RELATED WORKS

Several research works underline the potential of machine learning in improving the accuracy and efficiency of breast cancer detection. They discuss different approaches to developing machine learning prototypes, including the use of deep learning techniques, ensemble machine learningand automated detection using machine learning techniques.The authors used various datasets of mammograms to train and test their models, and achieved high levels of accuracy ranging from 90% to 98.1%.
The figure 2 illustrates number of publications on AI in breast cancer diagnosis between 2012–2022.

Nassif AB et al. [13] present a systematic literature review of the literature on the use of artificial intelligence techniques for breast cancer detection. The authors analysed scientific literature on the use of different artificial intelligence techniques, such as deep learning, neural networks, decision trees and classification methods, for the detection of breast cancer. The results of the study showed that the use of artificial intelligence to detect breast cancer is a growing area of research. The authors identified several data sets used for model training, such as the Breast Cancer Wisconsin Diagnostic Data Set (BCWD) and the Digital Database for Screening Mammography (DDSM). The results also showed that the performance of

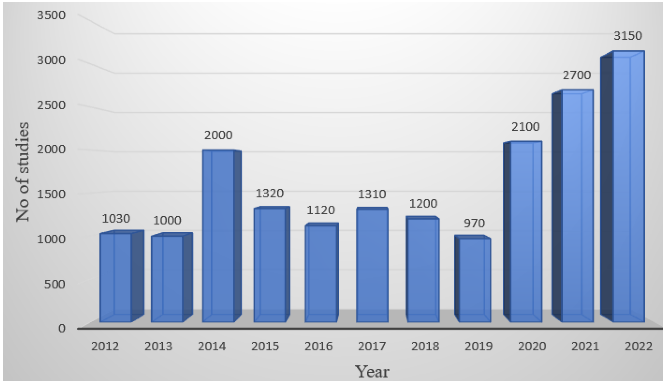

Fig. 2. Publications on AI in breast cancer diagnosis between 2012–2022.

artificial intelligence methods varied considerably depending on different factors, such as data quality, data set size, the choice of characteristics and the choice of hyperparameters.

Babarenda Gamage TP et al. [15] presents an automated workflow for analyzing breast cancer using computational biomechanics. The proposed workflow consists of three main stages: image segmentation, finite element modeling, and mechanical analysis. In the first stage, the breast MRI image is segmented to obtain the geometry and material properties of the breast tissue. In the second stage, the finite element model is created to simulate the mechanical behavior of the breast tissue. In the third stage, the mechanical analysis is performed to analyze the deformation and stress distribution of the breast tissue under different loading conditions. The authors tested their workflow on a dataset of 20 breast MRI images and compared the results with manual segmentation and analysis. The results showed that the automated workflow was accurate and efficient in analyzing the breast tissue mechanics. They demonstrated the potential applications of their workflow in breast cancer diagnosis and treatment. They suggested that the workflow could be used to identify early-stage breast cancer, assess tumor response to treatment, and optimize treatment planning. They highlights the potential of using computational bio-mechanics to improve breast cancer diagnosis and treatment. The proposed workflow may help in the development of personalized treatment plans and improve patient outcomes.

Assiri AS et al. [16] present a study on the use of ensemble machine learning techniques for breast cancer classification. The authors used four different machine learning algorithms: k-nearest neighbor (KNN), support vector machine (SVM), decision tree (DT), and random forest (RF) to classify breast cancer data. They also applied three ensemble learning techniques: bagging, boosting, and stacking to improve the accuracy of classification. The authors used the Wisconsin Breast Cancer Dataset (WBCD), which contains 699 instances with 11 features, to train and test their models. They evaluated the performance of each model using various metrics such as accuracy, precision, recall, and F1 score. The results showed that the ensemble learning techniques significantly improved the accuracy of the classification models. The best performance was achieved using the stacked ensemble model, which had an accuracy of 97.57%, precision of 97.51%, recall of 97.49%, and F1 score of 97.5%. Overall, the article demonstrates the effectiveness of using ensemble machine learning techniques for breast cancer classification. The authors suggest that the proposed methodology could be useful in assisting medical professionals in the early detection and diagnosis of breast cancer. The process map for the approach suggested in this article is illustrated in Figure 3.

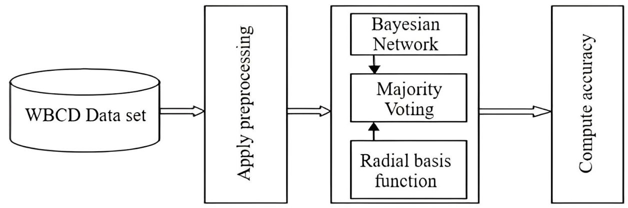

Fig. 3. Process Flow Diagram of the Proposed Approach [12].

Raman et al. [13] proposes a methodology for detecting, segmenting, and classifying tumors in digital mammograms using machine learning and problem-solving approaches. The proposed methodology consists of three stages: tumor detection, segmentation, and classification. In the first stage, an adaptive thresholding algorithm is used to identify the possible tumor regions in the mammogram. In the second stage, an edge detection algorithm is applied to segment the tumor region from the background. In the third stage, a feature extraction algorithm is used to extract features from the segmented tumor region, which are then used to classify the tumor as benign or malignant using a machine learning algorithm. The authors implemented a prototype of their methodology and tested it on a dataset of 322 mammograms. The results showed that the proposed methodology achieved an accuracy of 93.62% in tumor detection, 91.86% in tumor segmentation, and 93.47% in tumor classification. To conclude, the article demonstrates the potential of using machine learning and problem-solving approaches to improve the accuracy of tumor detection, segmentation, and classification in digital mammograms. The proposed methodology and prototype implementation may be useful in assisting radiologists in the early detection and diagnosis of breast cancer. The figure 2.4 illustrates the overall block diagram of the method for classifying weights.

Y. Edrees Almalki1 et al. [18] give a study on the application of hybrid machine learning techniques for the early detection of breast cancer in Saudi Arabian women using mammographic images. The authors used a dataset of 310 mammograms from Saudi Arabian women to train and test their machine learning model. They used two feature extraction methods: texture features and shape features, to extract features from the mammograms. They then used a hybrid machine learning algorithm, which combines k-nearest neighbor (KNN) and artificial neural network (ANN) algorithms, to classify the mammograms as either benign or malignant. The authors evaluated the performance of their model using various

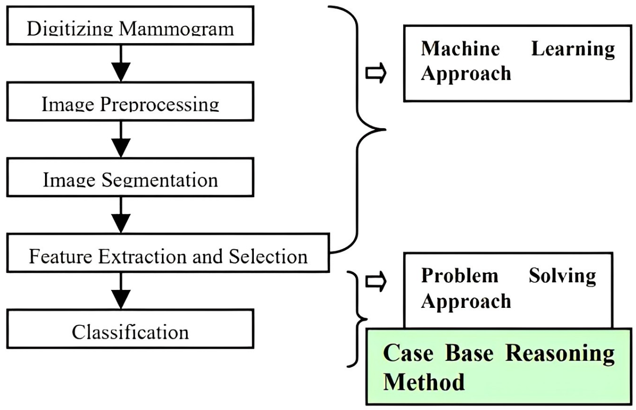

Fig. 4. The overall block diagram of the method for classifying weights [13].

metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). The results showed that the hybrid machine learning algorithm outperformed the individual KNN and ANN algorithms in terms of accuracy, sensitivity, specificity, and AUC-ROC. The best performance was achieved using the hybrid algorithm with texture and shape features, which had an accuracy of 98.1%, sensitivity of 97.1%, specificity of 99.0%, and AUC-ROC of 0.998. Overall, the study demonstrates the potential of using hybrid machine learning techniques and mammographic images for the early detection of breast cancer in Saudi Arabian women. The proposed approach is based on three models to get the results of the whole model. These models have different structures with different layers of parameters, functionality, learning rates, steps and filter sizes. This last may help in improving the accuracy of breast cancer diagnosis and assist radiologists in making better clinical decisions.

Digital Mammography DREAM Challenge [19] was a collaborative effort among several institutions and research organisations to develop machine learning algorithms for breast cancer detection. The challenge used a large dataset of digital mammograms to train and test the algorithms, and the winning solutions were able to achieve high levels of accuracy in detecting breast cancer.

The goal of the Digital Mammography (DM) DREAM Challenge was to apply an open science, crowd-sourced approach to develop and assess algorithms for risk stratification of screening mammograms that can be used to improve breast cancer detection. These algorithms have the potential to improve the interpretation of other tumour images, affecting a wide range of cancerous patients. The DM Challenge encouraged teams to apply deep learning approaches to a large set of mammographic images of over 640,000 images of 80,000 women. Dozens of teams participated in the Challenge, leading to the development of many innovative approaches to cancer detection and the setting of public standards and benchmarks.

iCAD SecondLook Digital [20] is a computer-aided detection (CAD) system that uses machine learning algorithms to analyze digital mammograms for signs of breast cancer. The system has been approved by the FDA and is being used in clinical practice to help radiologists detect breast cancer at an early stage. SecondLook is based on sophisticated patented algorithms that analyze the data, automatically identifying and marking suspicious regions in 2D mammography images. The solution provides the radiologist with a "second look" which helps the radiologist detect actionable missed cancers earlier than screening mammography alone. SecondLook detects and identifies suspicious masses and micro-calcifications utilizing image processing, pattern recognition and artificial intelligence techniques. Information from thousands of mammography images are incorporated into these algorithms enabling the product to distinguish between characteristics of cancerous and normal tissue.

Breast Cancer Risk Assessment Tool [21] (also known as the Gail Model) is a machine learning-based tool used to estimate a woman's risk of developing breast cancer. The tool uses various risk factors such as age, family history, and personal health history to predict a woman's likelihood of developing breast cancer over a certain period of time. The breast cancer risk assessment tool allows health professionals to estimate the risk of invasive breast cancer in a woman over the next five years and up to age 90 (lifetime risk). The tool uses a woman's personal medical and reproductive history and the history of breast cancer in her first-degree parents (mother, sisters, daughters) estimate the absolute risk of breast cancer — its likelihood of developing invasive breast cancer within a defined age range.

CancerSEEK [22] is a blood test that uses machine learning algorithms to detect multiple types of cancer, including breast cancer. The test analyzes various biomarkers in the blood to identify cancer at an early stage when it is more treatable.

CancerSEEK involves "the use of a single, non-invasive, multi-analyte test that simultaneously assesses the presence of mutations and eight biomarkers of cancer-associated proteins in the blood," says co-developer Anne Marie Lennon. PCR multiplex analysis of cell-free circulating tumour DNA (TDNA) detects mutations at 2,001 genomic positions in 16 genes, while protein biomarker levels are evaluated using immunoassays.

Hologic Genius 3D Mammography [23] is a digital mammography system that uses machine learning algorithms to detect breast cancer. The system creates 3D images of the breast tissue, which can help improve the accuracy of breast cancer detection compared to traditional 2D mammography. The new 3Dimensions system is designed to provide better 3D images to radiologists, a more comfortable mammography experience for patients and improved workflow for technologists.

V. BreastCancerCheck - Towards a Machine Learning based Platform for Cancer Detection

The existing solutions for breast cancer detection vary in terms of their approach and features. We propose a platform that aims to fill the gap in this field by providing an easy-to-use solution for early breast cancer detection. We suggest a Streamlit web application that helps doctors and make it simple for patients to predict the likelihood of breast cancer

based on a number of risk factors. It allows doctors to input tumor information for various features and get a prediction of whether the tumor is malignant or benign.The doctor can input the tumor information for the specified features. The user is prompted to input personal information and blood test results, which are used to make a prediction about the likelihood of breast cancer.If the prediction is positive, the user is advised to contact a doctor for further evaluation and guidance.
To implement the ML component , we used two datasets one containing personal data like Age and several blood test results about individuals, including whether they are breast cancer patients or not, and the other containing extracted from medical MRI images about made for doctors, indicating whether a tumor is malignant or benign. Based on the labeled data (0 for malignant and 1 for benign in the doctor dataset, and 1 for healthy and 2 for patients in the patient dataset),we need supervised learning algorithms so we trained four algorithms from the desired category on our data: kNN, SVM, Logistic Regression, and Random Forest (or Gradient Boosting).
For each dataset, we began the machine learning process by importing the dataset into our coding environment. After this, we performed exploratory data analysis to understand the relationships between the different columns in the 2 datasets. Then, we trained four supervised learning models on the data to identify which model performs best at classifying breast tumors.
To choose the best model we calculated metrics including accuracy which allowed us to evaluate the models' performance. Based on these metrics, we selected the model that provided the most accurate predictions for the 2 datasets and we deployed the best model as shown in fig 5.

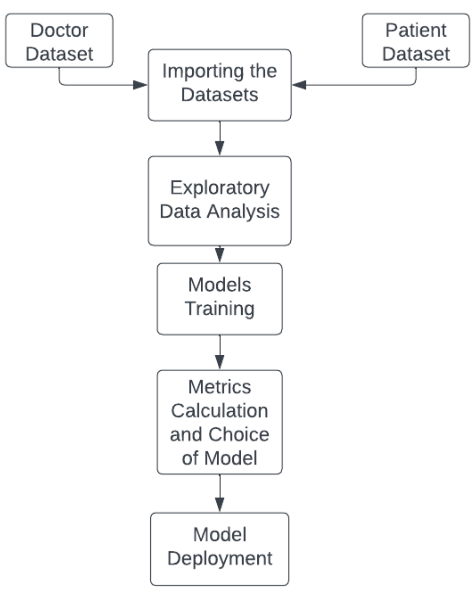

Fig. 5. Phases of our ML Component.

To select the best model, we evaluated each algorithm's accuracy and computational time. After evaluating the metrics for each model, we chose the one that achieved the highest accuracy while still being computationally efficient. Our model selection process ensures that we have a highly accurate and efficient model that can effectively classify breast cancer cases. The figure 6 shows how the model was selected.

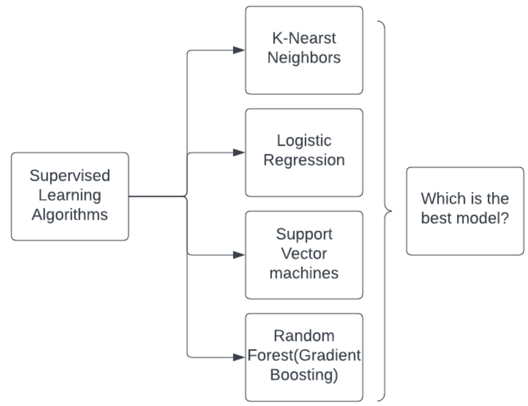

Fig. 6. Model selection.

The first dataset used in this project is the Wisconsin Breast Cancer Dataset (WDBC) The dataset contains : • 569 samples of breast mass lesions from patients in Wisconsin • 33 columns (diagnosis , radius mean ,etc...)

The second dataset contains informations about patient and their blood test results to detect breast cancer.
The subsequent aim of our study involves the compilation of a comprehensive evaluation of various machine learning models, focusing on their performance in terms of accuracy. This evaluation will be presented in 2 tabular formats, where the models' scores will be recorded and subsequently arranged in descending order, allowing for a clear comparison and identification of the top-performing models. Based on the results of both cross-validation accuracy and accuracy scores , we noticed that the Random Forest and Descision Tree

TABLE III
MODEL PERFORMANCES ON THE WISCONSIN DATASET

Model	Cross validation Accuracy	Accuracy
Random Forest	71.55 %	100 %
Logistic Regression	64.66 %	94.73 %
Linear SVC	56.90 %	91.04 %
KNN	47.41 %	94.73 %

TABLE IV
MODEL PERFORMANCES ON THE PATIENT DATASET

Model	Cross validation Accuracy	Accuracy
Descision Tree	91.96 %	100 %
Logistic Regression	94.20 %	78.45 %
Linear SVC	89.81 %	61.21 %
KNN	47.41 %	94.73 %

models are overfitting the training data in both datasets (100 % accuracy) so w chose the second in the ranking which is Logistic Regression model and we saved the model in two separate .pkl file using dump function from the pickle library. During our deep research in this theme, we observed that existing solutions in the field of breast cancer prediction primarily remained theoretical and were not effectively utilized by patients and doctors. To bridge this gap, we embarked on a mission to bring together the latest research and technology in a more accessible manner for ordinary individuals. We used as example figure 7 and 8 which describes the action of our streamlit app.

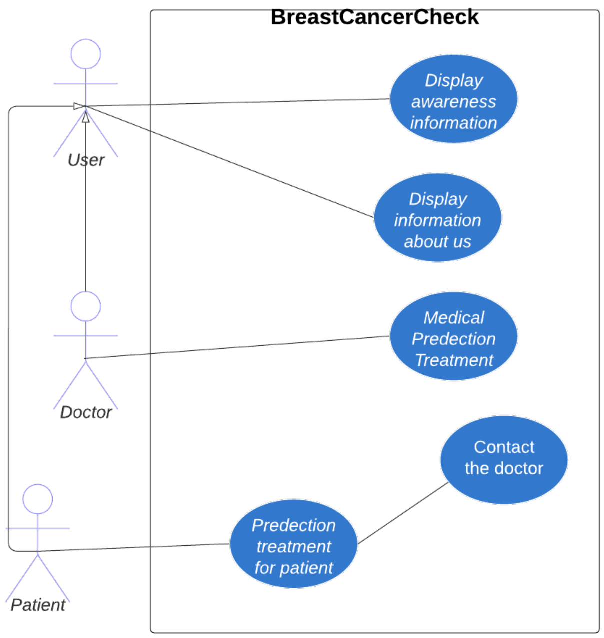

Fig. 7. Actions of our app

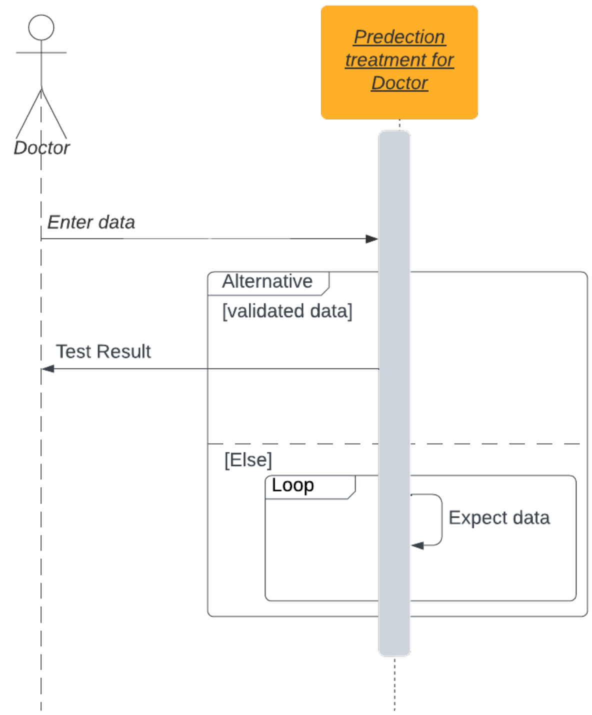

Fig. 8. Doctor Sequence Diagram

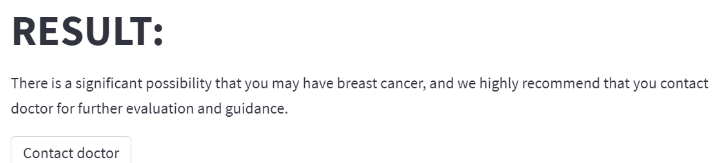

Fig. 9. Breast Cancer detection result

Our breast cancer detection web application was developed using Visual code and Streamlit library in python. And as mentioned, we chose Streamlit because it is a library that allows you to create user-friendly web applications in Python with great ease. This application uses machine learning algorithms to predict the presence or absence of breast cancer in patients. These algorithms were trained using data from previous patients. The prediction results are then displayed .

VI. Conclusion

This research paper has contributed to the field of breast cancer detection by developing a machine learning-based prototype that addresses the accessibility and usability challenges faced by non-technical users. The presented solution shows promise in aiding early detection and potentially improving patient outcomes. Further research and development are recommended to refine and enhance the prototype, with the ultimate goal of making a meaningful impact in the fight against breast cancer Building a web application for breast cancer detection is an iterative process that requires continuous development to add functionality and enhance the user experience. However, there may be missing features due to time or resource constraints. In our breast cancer detection project, we have identified several missing features that could be added to improve the user experience and quality of service. These include implementing a login functionality for different user types, extracting medical imaging data directly from radiology equipment, incorporating calendar functionality for appointment management, and integrating a chatbot or smart assistant for immediate user support.
We believe these features would greatly enhance the application and provide a more comprehensive solution for early breast cancer detection. In the next steps of our development, we plan to implement these missing features within the next 6 months, utilizing web development best practices and modern tools. Our goal is to continuously improve the user experience

and provide an effective solution in the fight against breast cancer.

Acknowledgment

We would like to express our gratitude to all individuals and organizations who contributed to the completion of this research project. Their support and assistance have been invaluable in our journey towards developing a machine learning-based prototype for breast cancer prediction.

References

[1] Zemni I, Kacem M, Dhouib W, Bennasrallah C, Hadhri R, Abroug H, Ben Fredj M, Mokni M, Bouanene I, Belguith AS. Breast cancer incidence and predictions (Monastir, Tunisia: 2002-2030): A registry-based study. PLoS One. 2022 May 26;17(5):e0268035. doi: 10.1371/journal.pone.0268035. PMID: 35617209; PMCID: PMC9135193.

[2] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015 Mar 1;136(5):E359-86. doi: 10.1002/ijc.29210. Epub 2014 Oct 9. PMID: 25220842.

[3] Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer. 2013 Mar 1;132(5):1133-45. doi: 10.1002/ijc.27711. Epub 2012 Jul 26. PMID: 22752881.

[4] Mallekh R, Khiari H, Hammami O, Hsairi M. Burden of cancer attributable to insufficient physical activity in Tunisia. East Mediterr Health J. 2023 Feb 26;29(2):110-118. doi: 10.26719/emhj.23.009. PMID: 36880492.

[5] Sharma R, Aashima, Nanda M, Fronterre C, Sewagudde P, Ssentongo AE, Yenney K, Arhin ND, Oh J, Amponsah-Manu F, Ssentongo P. Mapping Cancer in Africa: A Comprehensive and Comparable Characterization of 34 Cancer Types Using Estimates From GLOBOCAN 2020. Front Public Health. 2022 Apr 25;10:839835. doi: 10.3389/fpubh.2022.839835. PMID: 35548083; PMCID: PMC9082420.

[6] Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. Cancer Lett. 2020 Feb 28;471:61-71. doi: 10.1016/j.canlet.2019.12.007. Epub 2019 Dec 10. PMID: 31830558.

[7] Irvin Jr, W. J., & Carey, L. A. (2008). What is triple-negative breast cancer?. European journal of cancer, 44(18), 2799-2805.

[8] Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. N Engl J Med. 2009 Dec 17;361(25):2449-60. doi: 10.1056/NEJMra0804588. PMID: 20018966; PMCID: PMC2843693.

[9] Anderson, W. F., Chu, K. C., & Chang, S. (2003). Inflammatory breast carcinoma and noninflammatory locally advanced breast carcinoma: distinct clinicopathologic entities?. Journal of clinical oncology, 21(12), 2254-2259.

[10] Carter, H. B. (2013). A merican U rological A ssociation (AUA) Guideline on prostate cancer detection: process and rationale. BJU international, 112(5), 543-547.

[11] Grosan, C., & Abraham, A. (2011). Intelligent systems (Vol. 17, pp. 261-268). Berlin: Springer.

[12] Assiri AS, Nazir S, Velastin SA. Breast Tumor Classification Using an Ensemble Machine Learning Method. J Imaging. 2020 May 29;6(6):39. doi: 10.3390/jimaging6060039. PMID: 34460585; PMCID: PMC8321060.

[13] Raman, Valliappan & Putra, Sumari & Mandava, Rajeswari. (2010). A Theoretical Methodology and Prototype Implementation for Detection Segmentation Classification of Digital Mammogram Tumor by Machine Learning and Problem Solving Approach. International Journal of Computer Science Issues. 7.

[14] Nassif AB, Talib MA, Nasir Q, Afadar Y, Elgendy O. Breast cancer detection using artificial intelligence techniques: A systematic literature review. Artif Intell Med. 2022 May;127:102276. doi: 10.1016/j.artmed.2022.102276. Epub 2022 Mar 5. PMID: 35430037.

[15] Babarenda Gamage TP, Malcolm DTK, Maso Talou G, Mîra A, Doyle A, Nielsen PMF, Nash MP. An automated computational biomechanics workflow for improving breast cancer diagnosis and treatment. Interface Focus. 2019 Aug 6;9(4):20190034. doi: 10.1098/rsfs.2019.0034. Epub 2019 Jun 14. PMID: 31263540; PMCID: PMC6597517.

[16] Assiri AS, Nazir S, Velastin SA. Breast Tumor Classification Using an Ensemble Machine Learning Method. J Imaging. 2020 May 29;6(6):39. doi: 10.3390/jimaging6060039. PMID: 34460585; PMCID: PMC8321060.

[17] Raman, Valliappan & Putra, Sumari & Mandava, Rajeswari. (2010). A Theoretical Methodology and Prototype Implementation for Detection Segmentation Classification of Digital Mammogram Tumor by Machine Learning and Problem Solving Approach. International Journal of Computer Science Issues. 7.

[18] Y. Edrees Almalki1, A. Shaf, T. Ali, M. Aamir, S. Khalid Alduraibi et al., "Breast cancer detection in saudi arabian women using hybrid machine learning on mammographic images," Computers, Materials & Continua, vol. 72, no.3, pp. 4833–4851, 2022.

[19] Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, Lotter W, Jie Z, Du H, Wang S, Feng J, Feng M, Kim HE, Albiol F, Albiol A, Morrell S, Wojna Z, Ahsen ME, Asif U, Jimeno Yepes A, Yohanandan S, Rabinovici-Cohen S, Yi D, Hoff B, Yu T, Chaibub Neto E, Rubin DL, Lindholm P, Margolies LR, McBride RB, Rothstein JH, Sieh W, Ben-Ari R, Harrer S, Trister A, Friend S, Norman T, Sahiner B, Strand F, Guinney J, Stolovitzky G; and the DM DREAM Consortium; Mackey L, Cahoon J, Shen L, Sohn JH, Trivedi H, Shen Y, Buturovic L, Pereira JC, Cardoso JS, Castro E, Kalleberg KT, Pelka O, Nedjar I, Geras KJ, Nensa F, Goan E, Koitka S, Caballero L, Cox DD, Krishnaswamy P, Pandey G, Friedrich CM, Perrin D, Fookes C, Shi B, Cardoso Negrie G, Kawczynski M, Cho K, Khoo CS, Lo JY, Sorensen AG, Jung H. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. JAMA Netw Open. 2020 Mar 2;3(3):e200265. doi: 10.1001/jamanetworkopen.2020.0265. Erratum in: JAMA Netw Open. 2020 Mar 2;3(3):e204429. PMID: 32119094; PMCID: PMC7052735.

[20] Cole EB, Zhang Z, Marques HS, Edward Hendrick R, Yaffe MJ, Pisano ED. Impact of computer-aided detection systems on radiologist accuracy with digital mammography. AJR Am J Roentgenol. 2014 Oct;203(4):909-16. doi: 10.2214/AJR.12.10187. PMID: 25247960; PMCID: PMC4286296.

[21] Garcia-Closas M, Chatterjee N. Assessment of breast cancer risk: which tools to use? Lancet Oncol. 2019 Apr;20(4):463-464. doi: 10.1016/S1470-2045(19)30071-3. Epub 2019 Feb 21. PMID: 30799258; PMCID: PMC8211385.

[22] Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, Douville C, Javed AA, Wong F, Mattox A, Hruban RH, Wolfgang CL, Goggins MG, Dal Molin M, Wang TL, Roden R, Klein AP, Ptak J, Dobbyn L, Schaefer J, Silliman N, Popoli M, Vogelstein JT, Browne JD, Schoen RE, Brand RE, Tie J, Gibbs P, Wong HL, Mansfield AS, Jen J, Hanash SM, Falconi M, Allen PJ, Zhou S, Bettegowda C, Diaz LA Jr, Tomasetti C, Kinzler KW, Vogelstein B, Lennon AM, Papadopoulos N. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018 Feb 23;359(6378):926-930. doi: 10.1126/science.aar3247. Epub 2018 Jan 18. PMID: 29348365; PMCID: PMC6080308.

[23] Practice Bulletin Number 179: Breast Cancer Risk Assessment and Screening in Average-Risk Women. Obstet Gynecol. 2017 Jul;130(1):e1-e16. doi: 10.1097/AOG.0000000000002158. PMID: 28644335.