Breast Cancer Prediction App

Submitted By:

Simran Kachle

01 April 2023

(Task 0)

1) PROBLEM STATEMENT:

Breast cancer is a major health concern for women worldwide, with early detection being critical for effective treatment and improved patient outcomes. However, the traditional screening approach of testing all women is not cost-effective, and it can result in unnecessary interventions for low-risk patients. Hence, there is a need for a personalized approach to breast cancer screening that can accurately identify women at high risk of developing the disease. The aim of this project is to develop app that can predict the risk of breast cancer in women based on various risk factors, such as age, family history, reproductive history, and lifestyle factors. The model used in app will be validated using a large, independent dataset to ensure its accuracy and generalizability. This will help to ensure that the model can be used effectively in different healthcare settings and populations. Ultimately, the goal of this project is to provide healthcare providers with a reliable and effective tool to identify women who may benefit from additional screening or preventive measures, leading to earlier detection and better outcomes for breast cancer patients.

2) MARKET NEED ASSESSMENT:

- Determine target audience: Identify the demographic and geographic factors that are most likely to use the app such as age, gender, income, education level, and geographic location.

- Conduct market research: Gather data from potential users through surveys, focus groups, or other research methods. Inquire about their knowledge of breast cancer risk factors, interest in using a risk prediction app, and preferences for features or functionalities.

- Evaluate competitors: Analyze the strengths and weaknesses of existing breast cancer risk prediction apps or tools and find ways to differentiate your app from the competition.

- Assess pricing and revenue potential: Determine the pricing and revenue models that will be used, such as fees for app access, collaborations with healthcare providers or insurers, or licensing agreements with research institutions.

3) TARGET SPECIFICATIONS AND CHARACTERIZATION:

- Target Audience:
  The target audience for breast cancer risk prediction app includes women who are interested in knowing their risk of developing breast cancer, as well as healthcare providers who use risk prediction app to make clinical decisions regarding breast cancer screening and prevention.

- Data Collection Models:
  There are several models for collecting data for breast cancer risk prediction. These include:

  - Case-control studies: These studies compare women who have been diagnosed with breast cancer to women who have not. Data is collected on various risk factors for breast cancer, and this data can be used to develop risk prediction models.

  - Family history assessment tools: These tools collect information on the number and ages of relatives with breast cancer, as well as other factors that may be related to breast cancer risk, such as the age at which relatives were diagnosed.

4) EXTERNAL SEARCH (online information sources):

Breast Cancer Risk Assessment Tool

Reduce Your Risk | Breast Cancer UK

Breast Cancer: Symptoms, Stages, Types, and More (healthline.com)

5) BENCH MARKING ALTERNATE PRODUCTS:

- A 2018 study published in the Journal of Medical Internet Research compared the accuracy of six breast cancer risk assessment tools: BRCAPRO, IBIS, BOADICEA, Tyrer-Cuzick, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA), and CancerGene. The study found that the Tyrer-Cuzick Model had the highest accuracy.

- Another 2019 study published in the Journal of the National Cancer Institute compared four risk assessment tools: the Gail model, the Breast Cancer Surveillance Consortium model, the International Breast Cancer Intervention Study model, and the Tyrer-Cuzick model. The study found that the Tyrer-Cuzick Model had the highest sensitivity for predicting breast cancer risk in high-risk women.

- A 2015 study published in the Journal of Clinical Oncology compared the Myriad myRisk Hereditary Cancer Test to traditional genetic testing for BRCA1 and BRCA2 mutations. The study found that the myRisk test had higher sensitivity and specificity for detecting these mutations.

- A 2008 study published in the Journal of the American Medical Association compared the Breast Cancer Risk Assessment Tool to other models, including the Gail Model, the Rosner-Colditz Model, and the CARE Model. The study found that the Breast Cancer Risk Assessment Tool was more accurate than the other models for certain populations.

6) APPLICABLE CONSTRAINTS:

- Need high-quality data on clinical and lifestyle factors.
- Data may be incomplete, inconsistent, or difficult to obtain.
- Model performance may be affected by overfitting, missing data, and bias.
- Collecting and storing sensitive health data requires careful attention to privacy and security considerations.
- Need data encryption, user authentication, and compliance with regulatory requirements.

7) BUSINESS MODEL (Monetization Idea) :

- A business model for an application of breast cancer prediction could involve a combination of revenue streams, such as:

    - Paid downloads or subscriptions: Users could pay to download the app or subscribe to use it on a recurring basis.

    - In-app purchases: The app could offer additional features or services for purchase within the app, such as personalized risk assessments or recommendations for screening and prevention.

    - Advertising: The app could display targeted advertising based on user demographics and health profiles.

    - Partnerships: The app could form partnerships with healthcare providers, insurance companies, or other stakeholders to offer value-added services or generate referral revenue.

    - Data licensing: The app could license user data to researchers, healthcare organizations, or other third-party entities for research or commercial purposes.

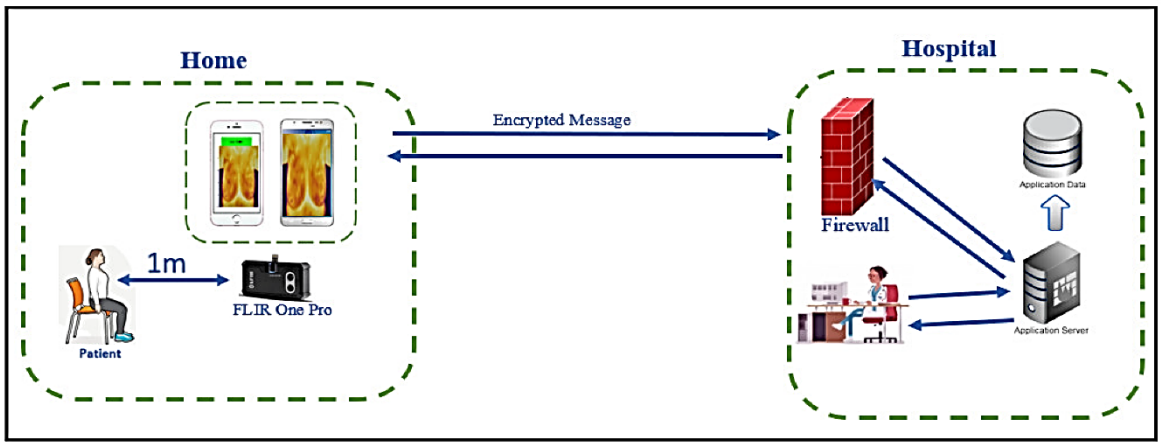
    - The business model should not compromise the privacy and security of user data or conflict with ethical considerations. The app should also provide value to users and contribute to improving breast cancer prevention and management.

8) CONCEPT GENERATION:

Breast cancer prediction application is an idea that emerged due to the need for improved early detection and prevention of breast cancer. Breast cancer affects millions of women worldwide and is the leading cause of cancer-related deaths among women. Traditional breast cancer screening and prevention programs have limitations, and many women at high risk for breast cancer are not identified until the disease has progressed. With the use of machine learning and artificial intelligence, applications can provide personalized risk assessments and recommendations for prevention. These applications can improve early detection and prevention of breast cancer and provide a more patient-centered and collaborative approach to breast cancer management.

9) CONCEPT DEVELOPMENT:

The application can use advanced machine learning algorithms and artificial intelligence to provide a more accurate and personalized risk assessment. The application can analyze large datasets to identify patterns and factors that can contribute to the development of breast cancer, thereby improving the accuracy of the risk assessment. The application may also use cloud services such as Amazon Web Services (AWS) or Microsoft Azure for storage and computation of user data. Additionally, the application may use web or mobile app development frameworks such as React Native or angularjs to build the user interface and provide a seamless user experience.



10) FINAL PRODUCT PROTOTYPE:

- Personalized Risk Assessment: The application asks the user to answer questions about their personal and family medical history, lifestyle habits, and other factors that can affect their risk of developing breast cancer. This enables the application to provide a more accurate and personalized risk assessment.

- Interactive Dashboard: The application provides an interactive dashboard that shows the user's risk level, potential risk factors, and recommendations for risk reduction. This helps users to understand their risk and take steps to reduce it.

- Breast Cancer Screening Reminder: The application provides reminders to users about when they are due for breast cancer screenings, such as mammograms or clinical breast exams. This helps users to stay up to date with their screening schedule and catch any potential cancer early.

- Educational Resources: The application provides educational resources, such as videos or articles, to help users learn more about breast cancer risk factors, screening options, and risk reduction strategies.

- Risk Reduction Strategies: The application provides personalized recommendations for risk reduction strategies, such as lifestyle changes or medications that may help to reduce breast cancer risk.

- Connection to Healthcare Providers: The application provides a way for users to connect with healthcare providers, such as through

telemedicine or in-person appointments, to discuss their risk and receive personalized recommendations for risk reduction.

- Overall, the final product prototype for a breast cancer risk prediction application combines personalized risk assessment, education, and support to help users reduce their risk of developing breast cancer and catch potential cancer early through screening reminders.

10.1) Diagram:

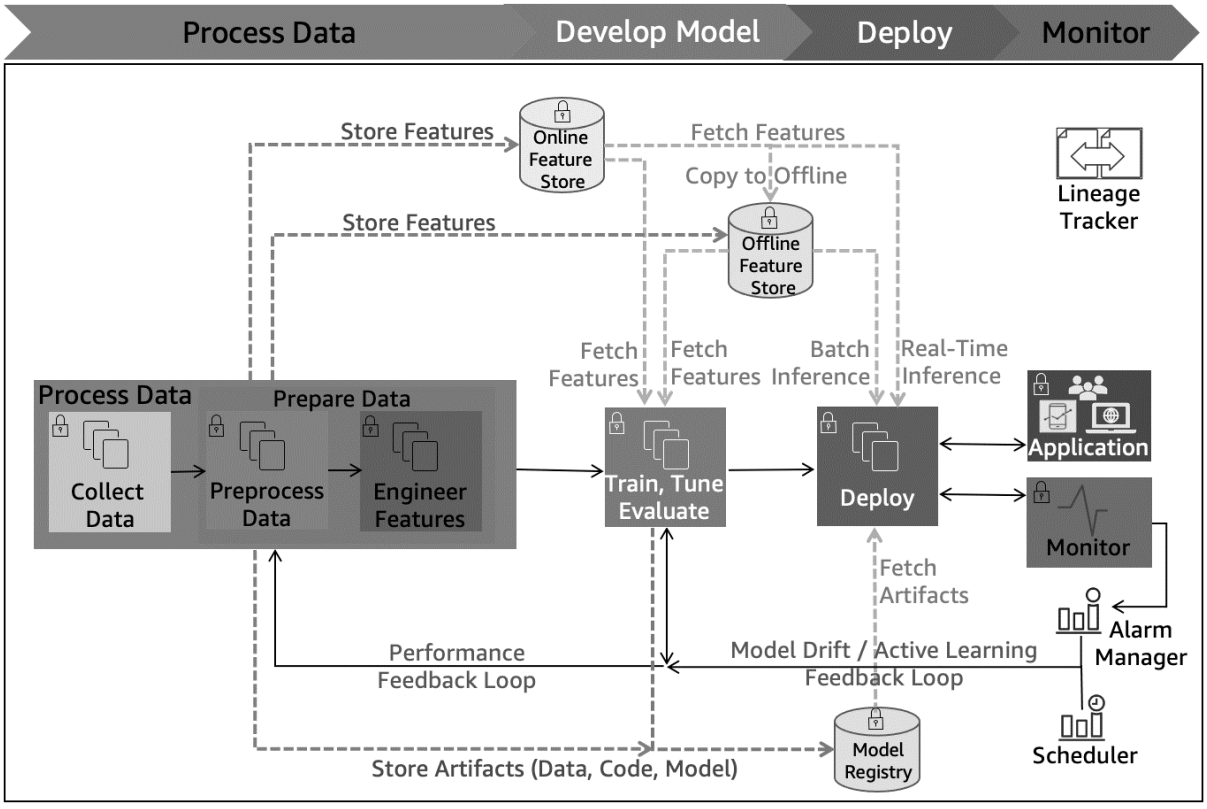


Fig. Product Prototype

11) PRODUCT DETAILS:

11.1) How does it work?

The breast cancer prediction application collects user data such as personal and family medical history, lifestyle habits, and other factors that can affect their risk of developing breast cancer. This data is then analyzed using machine learning algorithms to identify risk factors for breast cancer. Based on the user data and risk factors identified, the application provides a personalized risk assessment and recommendations for risk reduction strategies. The application also provides screening reminders, educational resources, and connections to healthcare providers to help users reduce their risk of developing breast cancer.

11.2) Data Sources

The breast cancer prediction application may source data from various sources such as the user's electronic health records, self-reported data from the user, and publicly available data on breast cancer risk factors.

11.3) Algorithms

The breast cancer prediction application uses machine learning algorithms such as Gradient Boosting Trees and random forests to analyze the user data and identify risk factors for breast cancer. These algorithms are trained on historical data to learn the patterns and relationships between the different variables that affect the risk of breast cancer. The application also uses statistical methods to validate the accuracy and reliability of the predictions.

Random Forest (RF)

As a non-parametric approach, the RF uses the classification method. For each set of data, the RF performs categorization at high speed and applies a large number of decision trees. In each tree, there is a random number of input variables, then all the trees are combined for a better inference from the variables.

Gradient Boosting Trees (GBT)

This algorithm is one of the reinforcement gradient algorithms with a very good performance in classification and performs the best classification for each of the data. In this method, the trees are trained one after another; each subset tree is taught primarily with data erroneously predicted by the previous tree. This process continuously reduces the model error since each model is sequentially improved against the weaknesses of the previous model.

Multi-Layer Perceptron (MLP)

As a deep artificial neural network, the MLP is composed of an input layer for receiving the signal, an output layer used for prediction, and in between those two, some hidden layers are acting as the computation engine. The MLP is trained by a backpropagation algorithm, which is part of the supervised networks. In this network, data are driven from input nodes to output nodes. If there is an error in the output, this error must be somehow returned from the output to the input, and this corrects the weights. The most commonly used method for this is the post-diffusion algorithm .

Genetic Algorithm (GA)

As a subset of the evolutionary computing algorithm, GA is directly associated with artificial intelligence and used for solving optimization problems through the evolution process. To obtain the best answer, the GA applies the best survival rule to a series of problems for patterning the best solution for problems. In each generation, the optimal solution is achieved

based on a natural biological process and by selecting the best chromosomes for creating the subsequent generation to solve the problem optimally .

11.4) Frameworks and Software

The breast cancer prediction application may use software such as Python, and tensorflow to implement the machine learning algorithms and frameworks. The application may also use cloud services such as Amazon Web Services (AWS) or Microsoft Azure for storage and computation of user data. Additionally, the application may use web or mobile app development frameworks such as React Native or angularjs to build the user interface and provide a seamless user experience.

12) CONCLUSION:

The breast cancer risk prediction application has the potential to be a valuable tool for improving breast cancer prevention and early detection efforts. By leveraging machine learning algorithms, the application can provide a more accurate and personalized risk assessment for users based on their individual risk factors. Additionally, the application can provide recommendations for risk reduction strategies and screening reminders to help users stay up to date with their breast cancer screenings. The application also has the potential to provide educational resources and connections to healthcare providers to further support breast cancer prevention and early detection efforts.However, the success of the application will depend on its ability to collect and analyze accurate and relevant data while maintaining user privacy and security. The application will also need to be user-friendly and accessible to a wide range of users, including those with varying levels of technological literacy.

.

13) REFERENCES:

1. Hidayanto, A. N., Ayuningtyas, D., & Budi, I. (2020). Machine learning approach for breast cancer risk prediction. International Journal of Advanced Science and Technology, 29(9), 2498-2506.

2. Wenbin Yeu,Zidong Wang,Hongwei Chen,Annette Payne & Xiaohui Liu.Machine Learning with Applications in Breast Cancer Diagnosis and Prognosis.

3. Reza Rabiei,[1] Seyed Mohammad Ayyoubzadeh,[2] Solmaz Sohrabei,[3*] Marzieh Esmaeili,[2] and Alireza Atashi. Prediction of Breast Cancer using Machine Learning Approaches .Journal of Biomedical Physics and Engineering

4. Yala, A., Lehman, C., Schuster, T., Portnoi, T., Barzilay, R., & Mitchell, S. (2019). A deep learning mammography-based model for improved breast cancer risk prediction. Radiology, 292(1), 60-66.

5. A Breast Cancer Smartphone App to Navigate the Breast Cancer Journey: Mixed Methods Study.
    a. JMIR Formative Research - A Breast Cancer Smartphone App to Navigate the Breast Cancer Journey: Mixed Methods Study

6. National Cancer Institute. Breast Cancer Risk Assessment Tool.
    a. Breast Cancer Risk Assessment Tool

7. American Cancer Society. Breast Cancer Risk and Prediction
    a. Breast Cancer Risk Factors and Prevention Methods