



Feasibility and Usability of Facebook Chatbots as a Breast Cancer Risk Assessment Tool

¹Bryan Augustine C. Cabansay, ²Larmie S. Feliscuzo and ³Elmer A. Maravillas
¹*Cebu Institute of Technology, University Bryan, Tempe, Arizona, United States*
²*IT Department*
³*Department of Computer Science*

Key words: Breast, cancer, cancer risk assessment, chatbot, Facebook Messenger, supervised machine learning, web application

Corresponding Author:

Bryan Augustine C. Cabansay
Cebu Institute of Technology, University Bryan, Tempe, Arizona, United States

Page No.: 28-34

Volume: 19, Issue 1, 2020

ISSN: 1682-3915

Asian Journal of Information Technology

Copy Right: Medwell Publications

Abstract: In Asia, the Philippines has the highest incidence rate of breast cancer. This is one of the most common forms of cancer which can be detected early and when treated properly, be cured. There is also no denying that the Philippines love social media. In fact, the country is the leader in social media usage amounting to 3 h and 57 min of average use. The purpose of this study is to take advantage of Facebook's Messenger platform in creating a chatbot for breast cancer risk assessment and assess its feasibility and usability. In order to choose an accurate model, this study used publicly available data from the Breast Cancer Surveillance Consortium to compare the performance of different classification algorithms. The multilayer perceptron classifier was chosen as the prediction model since it had the best performance and was uploaded to a web application. This web application was then integrated to the Messenger platform in the form of a chat bot. A convenience sample of 28 women tested the application and a usability survey was conducted immediately after. The outcomes are presented and discussions are made to illustrate the significance of the results and how it can spearhead additional or enhance existing public health care services among assigned government agencies.

INTRODUCTION

The American Cancer Society describes cancer as a group of diseases characterized by the uncontrolled growth and spread of abnormal cells and if left uncontrolled, usually results in death. It has caused a great burden globally and is a major cause of grief and loss among families. Unlike other diseases, causes of most cancers are usually unknown but can be attributed to some risk factors. These risks may include modifiable risk

factors like lifestyle (tobacco consumption, bodyweight, etc.) and non-modifiable ones like genetic mutations and hormones or immune related conditions^[1].

Last 2015, cancer was dubbed as the second leading cause of death and was responsible for 8.8 million mortalities globally. Nearly 1 in 6 deaths is attributed to cancer and around 70% of those deaths arise from low and middle-income countries^[2]. In the Philippines, the situation is the same. Cancer or neoplasm is the second leading cause of death for both sexes which consists

10.4% of the total. For men, it is the third major cause of death at 8.8%. For women on the other hand, it is the first ranked cause of death which is 12.5% of the total^[2].

Breast cancer is the most frequent form of cancer among women. In the Philippines, breast cancer is predominant for both sexes combining at 19% in 2015 and is ranked first among women accounting for 33%^[3]. According to the Philippine Society of Medical Oncology, one in every thirteen Filipinas is expected to have it in her lifetime and that the Philippines have the highest incidence rate of breast cancer in Asia^[4].

Like other diseases, early detection is key in increasing the survival of patients. Most high-income countries adopt a form of mass screening using mammography, together with a breast exam from a physician. But for a low-to-middle income country like the Philippines, cost can be a prohibitive factor for this to be a common practice. Country-wide screening for all women is not viable. It would be more economical if a selective approach is implemented. This means that those who need it the most or those who are at risk will be prioritized so that interventions could have a more effective outcome. Examining those who are at-risk usually involves a form of cancer risk assessment which is done to predict a person's tendency of developing cancer. With the current technologies available today, it is not farfetched to say that such health care service can be made accessible to all women from all walks of life.

It does not come as a surprise to say that the Philippines is one of the leading countries in terms of social media usage. According to the consultancy firm We Are Social's^[5] digital report in 2018, Filipinos spend an average of 3 h and 57 min a day on social media. And this is dominated by Facebook at 57%. Currently, the Philippines has an estimated 67 million monthly active Facebook users comprising of 57% declared females^[5]. Due to Facebook's pervasiveness and becoming engrained in the lives of many Filipinos (rich or poor), it provides a significant platform for institutions like government health agencies to reach people who were previously deemed unreachable. There have been a lot of studies incorporating Facebook in health promotion or providing direct health care interventions yielding positive results. It has encouraging implications because mass adoption of Facebook means a highly scalable and economical means of promoting health and wellbeing.

Computer science, especially artificial intelligence, seems to be the breeding ground of innovation. Technologies like machine learning have been successfully used in clinical practice across multiple medical specialties. Machine learning, specifically supervised machine learning (which uses past learned data to generate a prediction) are usually applied in medical decision-making process like the cancer risk prediction^[6].

Chatbots are software used to conduct a conversation using either audio or text. Facebook Messenger provides a rich a platform where organizations can customize one for multitudes of use cases like customer management, query handling, and survey taking. To take advantage of this free technology, this pilot study aims to assess the feasibility and usability of using a Facebook Messenger chatbot to assess breast cancer risk using machine learning models.

MATERIALS AND METHODS

Overall design: The objective of this study was to evaluate the feasibility and usability of using a Facebook Messenger chatbot to assess breast cancer risk using a supervised machine learning algorithm. This study is composed of two phases: the performance evaluation and selection of the model to be used by the chatbot and the usability testing by convenience sampled women. For the first phase, the main goal is to find out which is more accurate or best performing model between the different classification algorithms. The model that produced the best results after three iterations is chosen and is attached to the bot. The second phase involves a survey asking the attitude of the participants towards the chatbot. The answers of all the participants are then presented as percentages.

Supervised classification algorithms: All calculations and testing were performed using the web application Jupyter notebook utilizing the machine learning python library called scikit-learn. Binary classification was performed where the output would only constitute of one result which will have a value ranging from zero and one (0 for negative and 1 for positive breast cancer). Everything was coded using the Python programming language. The algorithms tested were random forest, multilayer perceptron, quadratic discriminant analysis, K-nearest neighbors, stochastic gradient descent, logistic regression, Gaussian Naive Bayes, Bernoulli Naive Bayes, decision tree and AdaBoost classifiers. Recall, precision, area under the Receiver Operating Characteristic curve (ROC) or AUC and F1-score were used as performance measures. Performance evaluation was performed using existing functions found in the scikit-learn library.

Data set: The data used for training and testing the models are retrieved from Breast Cancer Surveillance Consortium (<http://bcsc-research.org/>) risk factor data set which is composed of thirteen columns (see Table 1 for descriptions). To improve model performance, data that have missing or unknown values are excluded together

Table 1: Description of variables in the data set

Variable/column name	Description	Coding
years	Calendar year of observation	
age_group_5_years	Age (years) in 5-year groups	1 = Age 18-29 2 = Age 30-34 3 = Age 35-39 4 = Age 40-44 5 = Age 45-49 6 = Age 50-54 7 = Age 55-59 8 = Age 60-64 9 = Age 65-69 10 = Age 70-74 11 = Age 75-79 12 = Age 80-84 13 = Age ≥85
race_eth	Race/ethnicity	1 = Non-Hispanic white 2 = Non-Hispanic black 3 = Asian/Pacific Islander 4 = Native American 5 = Hispanic 6 = Other/mixed 9 = Unknown
first_degree_hx	History of breast cancer in a first degree relative	0 = No 1 = Yes 9 = Unknown
age_menarche	Age (years) at menarche	0 = Age ≥14 1 = Age 12-13 2 = Age <12 9 = Unknown
age_first_birth	Age (years) at first birth	0 = Age <20 1 = Age 20-24 2 = Age 25-29 3 = Age ≥30 4 = Nulliparous 9 = Unknown
BIRADS_breast_density	BI-RADS breast density	1 = Almost entirely fat 2 = Scattered fibroglandular densities 3 = Heterogeneously dense 4 = Extremely dense 9 = Unknown or different measurement system
current_hrt	Use of hormone replacement therapy	0 = No 1 = Yes 9 = Unknown
menopause	Menopausal status	1 = Pre- or peri-menopausal 2 = Post-menopausal 3 = Surgical menopause 9 = Unknown
bmi_group	Body mass index	1 = 10-24.99 2 = 25-29.99 3 = 30-34.99 4 = 35 or more 9 = Unknown
biophx	Previous breast biopsy or aspiration	0 = No 1 = Yes 9 = Unknown;
breast_cancer_history	Prior breast cancer diagnosis	0 = No; 1 = Yes; 9 = Unknown
count	Frequency count of this combination of covariates	Numerical

with BI-RADS breast density variable since the study participants are less likely to know this information. The columns year and count were also not included. The year column represents the calendar year when the observation

was made and the count column indicates the number of times the same data combination appeared. Instead of involving the column count, this variable served as the total number of times the row of data appeared in the final

Table 2: Average performance of models

Models	Classification	Precision	Recall	F1-Score	AUC
Random forest classifier	0	0.94	0.81	0.87	0.878495
	1	0.83	0.95	0.89	
Multilayered perceptron	0	0.96	0.79	0.87	0.881853
	1	0.83	0.97	0.89	
Quadratic discriminant analysis	0	0.94	0.81	0.87	0.878624
	1	0.83	0.95	0.89	
K-nearest neighbors classifier	0	0.89	0.82	0.85	0.860194
	1	0.84	0.9	0.87	
Stochastic gradient descent classifier	0	0.95	0.81	0.87	0.879885
	1	0.83	0.95	0.89	
Logistic regression CV classifier	0	0.95	0.8	0.87	0.880254
	1	0.83	0.96	0.89	
Gaussian naive bayes	0	0.89	0.84	0.87	0.870499
	1	0.85	0.9	0.87	
Bernoulli naive bayes	0	0.89	0.84	0.86	0.865261
	1	0.85	0.9	0.87	
Decision tree classifier	0	0.93	0.81	0.87	0.875833
	1	0.83	0.94	0.88	
AdaBoost classifier	0	0.97	0.79	0.87	0.881141
	1	0.82	0.97	0.89	

dataset. The breast cancer history which has the binary of data of “yes” or “no”, column was treated as the dependent variable. The variables age in five-year groups, race or ethnicity, breast cancer history of a first degree relative, age (years) at menarche, age (years) at first birth of child, use of hormone replacement therapy, menopausal status and body mass index composed the independent variables. All data were scaled using sci-kit Standard Scaler and were rearranged randomly before sampling.

Chatbot: The chatbot was coded in Python using the Django web framework and was deployed in Heroku cloud platform. The bot was attached to a Facebook page (via webhook) where the users can interact using the Messenger app. A dummy user was created specifically for this study and was utilized by each subject to allow interaction with the chatbot. The application was in development mode the entire duration of the study.

Participants: The study used convenience sampling to recruit female participants. All subjects were required to interact with the chatbot and provide feedback immediately after by answering the usability questionnaire. Full written consent was required before allowing the subjects to participate. To maintain the privacy and confidentiality, this study strictly adhered to the Republic Act No. 10173 (Act), also known as the Data Privacy Act of 2012 with respect to the data that it will collect, record, organize, update, use, consolidate or destruct from the study participants. The personal data obtained was stored in a password enabled file and was not made public to anyone outside of the authors. Google forms and traditional questionnaires were used as means to conduct and collect the survey.

Usability questionnaire: The survey is made up of two major parts. The first involves demographics and consists of the age, educational attainment, marital and employment statuses of the participants. The rest of the survey questionnaire used a 7-point Likert scale and focused on three aspects: usefulness, ease of use and satisfaction. The questions were based on and were selected based on its applicability with the software. Participants were also encouraged to provide suggestions or opinions if they have any.

RESULTS AND DISCUSSION

Model selection: There was a final total of 6,318,638 rows produced when additional rows were generated according to the variable 'count'. This data was reduced to 1,131,052 records after removal of unneeded columns and rows where there was a missing or unknown value. In the end, there were a total of 70,991 records that had breast cancer and 1,060,061 records that did not. Since the dataset was highly imbalanced, under sampling was performed on the majority class and the data was split into training (80%) and test (20%) datasets. After three iterations of training and testing, the multi-layered perceptron (MLP) model performed the best followed by the AdaBoost classifier. For the no cancer or zero classification, MLP had average scores of 0.96 in precision, 0.8 in recall and 0.87 as the F1 score. For the positive classification, it had average scores of 0.83 in precision, 0.97 in recall and 0.89 as the F1 score. Both MLP and AdaBoost had almost identical precision, recall, and F1 scores but just slightly differed in ROC AUC. MLP had an 0.880 ROC AUC while AdaBoost had 0.879. Table 2 shows the performance of the rest of the models.

Chatbot setup: In this study, the chatbot served both as a data collection tool (by asking questions and saving the answers) as well as a risk assessor (by using the model to predict the dependent variable). After the model selection, the resulting data structure and scaler were serialized and were saved as pickle files. These pickle files were used when the users finished answering all the questions and it was time for classification. The chatbot took advantage of Messenger’s different user interfaces. For multiple choice questions, users could press buttons that represents their answers. For answers requiring numbers, users were asked to input text directly. Figure 1 shows a flowchart of how the chatbot functions.

Usability: The demographic characteristics of the 28 study participants are shown in Table 3. All participants were Filipino women that had a mean age of 36.75 ± 13.64 years. Majority of the respondents had a college degree (75%) and were currently employed (57%). User response was very positive as shown in both Table 4 and Fig. 2. Users all agreed that the chatbot was fun to use and easy to understand, use and learn. Most of the participants were satisfied with using it and they did not notice any inconsistencies. The results also show that most subjects did not need any written instructions when using it and thought that the chatbot was effortless. There were six subjects that expressed their opinions and almost all of them were positive. One participant wrote “The chatbot is easy to use especially for women who really gives importance to their health status. I suggest to have it shared through social media so all women can take advantage”. The Philippines is an archipelago in Southeast Asia consisting of thousands of different islands. It has a population of 104.9 million as of 2017 and is the thirteenth most populous country in the world^[7]. Even though the government has made significant strides towards development especially in health, there are still problems with disease prevention and control. Last 2016, the government appropriated PHP 123 billion to the Department of Health. Public health composed 75.4% of this budget which included a sizable amount to subsidies. Disease prevention and control only accounted for 6.5%^[7].

As discussed earlier, breast cancer is a major problem for Filipino women. As per^[8], early detection is critical in order to improve outcome and survival. Countries like the Philippines that have limited resources where most women are diagnosed late should prioritize early diagnostic programs. Breast cancer can be detected by using either mammography, clinical breast exam, or breast self-exam. Mammography is a diagnostic procedure involving X-ray and has been shown to reduce breast cancer deaths. Clinical breast exam requires the presence of a trained health care professional and is a promising

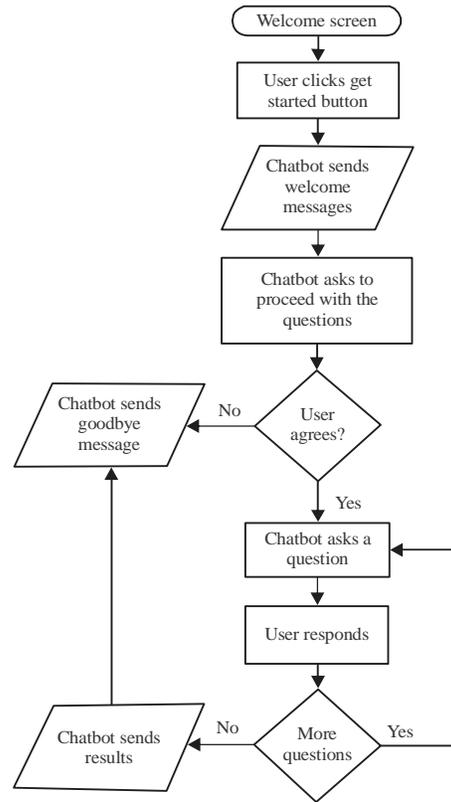


Fig. 1: Flow chart of how the chatbot works

approach for low resource health care settings but is still being investigated. These screening procedures require substantial investment and costs to the government, especially if implemented in the public health care setting. In a country like the Philippines where screening programs barely exist^[7], careful thought should be given before proceeding with this type of service. An alternative to this mass screening of the entire population is selective screening.

Selective screening involves inclusion of only those of the high-risk group of a population. This greatly reduces cost and assumes that there is a subpopulation with a high risk of the disease and that these people can be identified^[9]. One method of identifying these high-risk individuals is by using the chatbot developed in this study. The application of this chatbot is not only limited to risk assessment. This can also become a tool for data mining, health information dissemination, etc. Because of the pervasiveness of the Messenger app in the Philippine society, public health stakeholders should piggyback on this app’s reach and accessibility and use it to their advantage. This technology can incredibly cut costs and handle health care delivery problems like logistics. This study had several limitations that should be considered in

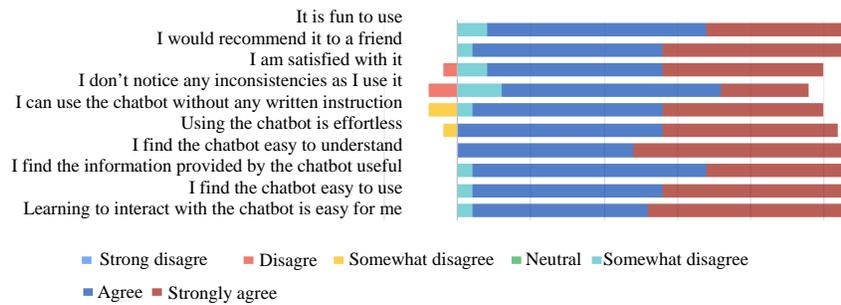


Fig. 2: Usability questionnaire response from users

Table 3: Participant's baseline demographic

Characteristic	n	Percentage
Total sample	28	100
Demographic characteristics		
Mean age, years (m±SD)	36.75±13.64	
Education		
Secondary education	7	25
College degree	21	75
Marital status		
Single	13	46
Married	14	50
Widowed	1	4
Employment		
Homemaker	7	25
Employed	16	57
Unemployed	5	18

Table 4: Summary of participant responses to the usability questionnaire for using the chatbot

Survey item	Strongly disagree		Somewhat disagree		Somewhat agree		Strongly agree
	0	1	0	1	0	1	
Learning to interact with the chatbot is easy for me	0	0	0	1	1	12	14
I find the chatbot easy to use	0	0	0	1	1	13	13
I find the information provided by the chatbot useful	0	0	0	1	1	16	10
I find the chatbot easy to understand	0	0	0	1	0	12	15
Using the chatbot is effortless	0	0	1	1	0	14	12
I can use the chatbot without any written instructions	0	0	2	1	1	13	11
I don't notice any inconsistencies as I use it	0	2	0	2	3	15	6
I am satisfied with it	0	1	0	2	2	12	11
I would recommend it to a friend	0	0	0	1	1	13	13
It is fun to use	0	0	0	1	2	15	10

future efforts. Models were not optimized to their fullest and were used as is or out of the box. Future research should focus on developing the most appropriate model that will have optimal performance and what sampling method would suit best. The small sample size and lack of diversity of the study subjects should also be mentioned.

Almost all of the study participants have a college degree or were able to finish secondary education. It would have been better to also include subjects that had other educational attainments other than the two that were present in this study. There is also a concern about the acceptability of sending information to Facebook Messenger. With the recent controversies surrounding

Facebook, studies should be performed to assess the attitude of people towards the use of messaging platforms in health care delivery and other government related services.

Other than being a risk assessment tool for breast cancer, future studies can also be done to test the other fields where a Messenger chatbot could be of use. And it should not only be limited to medicine either. Data mining, information dissemination, etc. are potential good candidates. Chatbots, like WoeBot, have already been successfully used as to provide therapeutic care and was proven to significantly reduce anxiety and depression among people aged 18-28 years old^[10]. The potential is there and should be investigated further.

REFERENCES

01. ACS., 2018. Cancer facts & figures 2018. American Cancer Society, Atlanta, Georgia, USA. <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html>
02. WHO., 2018. Cancer. World Health Organization, Geneva, Switzerland. <https://www.who.int/news-room/fact-sheets/detail/cancer>
03. Laudico, A.V., M.R. Mirasol-Lumague, V. Medina, C.A. Mapua, F.G. Valenzuela and E. Pukkala, 2015. Philippine cancer facts and estimates. Philippine Cancer Society, Inc., Manila, Philippines. http://www.philcancer.org.ph/wp-content/uploads/2017/07/2015-PCS-Ca-Facts-Estimates_CAN090516.pdf.
04. Anonymous, 2018. Breast cancer in the Philippines. Philippine Council for Health Research and Development, Taguig, Philippines.
05. We Are Social, 2018. Global digital report 2018. We Are Social, London, USA. <https://digitalreport.wearesocial.com/>
06. Barlow, W.E., E. White, Ballard-Barbash, R. Vacek and P. Titus-Ernstoff *et al.*, 2006. Prospective breast cancer risk prediction model for women undergoing screening mammography. J. Nat. Cancer Institute, 98: 1204-1214.
07. Dayrit, M.M., L.P. Lagrada, O.F. Picazo, M.C. Pons and M.C. Villaverde, 2018. The Philippines Health System Review. World Health Organization, Geneva, Switzerland, ISBN: 9789290226734,.
08. WHO., 2020. Breast cancer. World Health Organization, Geneva, Switzerland. <https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>
09. Hakama, M., E. Pukkala and P. Saastamoinen, 1979. Selective screening: Theory and practice based on high-risk groups of cervical cancer. J. Epidemiol. Community Health, 33: 257-261.
10. WLI., 2020. Frequently asked questions. Woebot Labs Inc., San Francisco, California. <https://woebot.io/faqs>.