The International Journal of Engineering and Science (IJES)
|| Volume || 11 || Issue || 4 || Series I || Pages || PP 01-09 || 2022 ||
ISSN (e): 2319-1813 ISSN (p): 20-24-1805

Case-Based Reasoning Recommender System for Diagnosing Breast Cancer Disease: Confusable Disease Using Genetic Algorithm

Jatto Joshua O.[1], Okengwu Ugochi A.[2],
M.Sc. Student Department of Computer Science,University of Port Harcourt, East/West Road, PMB 5323 Choba Rivers State, Nigeria[1]
Senior Lecturer Department of Computer Science,University of Port Harcourt, East/West Road, PMB 5323 Choba Rivers State, Nigeria[2]

----------ABSTRACT----------
Most countries are still struggling and suffering to have a better health delivery system. The disease commonly found among ladies is breast cancer regarded as a confusable disease and results of researchers reveals that if the breast cancer disease is detected at early stage, the chances of overcoming the disease will be very high compared to when the disease is been treated or detected at later stage. A lot of ladies have lost their lives regards to the confusability of the disease (breast cancer).In this paper the Genetic and SVM Algorithms are employed to detect and ascertain the level or stages of human confusable breast cancer disease. The genetic algorithm (GA) performed better than the Support Vector Machine (SVM) model in terms of detection accuracy, precision and recall. The system was successfully trained and tested compared to 92.40% and 71.99% accuracy level with GA and SVM respectively.

KEYWORDS: *Recommendation Engine, Confusable Disease, Genetic Algorithm, SVM, Case-Based reasoning.*

Date of Submission: 01-04-2022 Date of Acceptance: 10-04-2022

I. INTRODUCTION

Confusable diseases are sicknesses/diseases that share common manifestations/symptoms and as such become difficult for the clinical specialist to effectively analyze them [10]. The most common way of diagnosing confusable diseases is generally perplexing particularly when there is no computerized framework to back it up [1],[10]. Clinical Diagnostic framework is exceptionally essential in creating precise and fast analysis of disease. Especially, in the present time, there are a few diseases whose manifestations are very comparative in beginning stage, and yet, introductory level determination is additionally needed to be precise. All in all, there is need for automated analysis frameworks to have exact finding results.Medical uncertainty is an inherent phenomenon in medical science; it is what fuels medical research, prompts patients to seek medical attention and stimulate medical intervention notwithstanding, it poses challenges in diagnostic decision making [12]

In clinical science, determination of a disease is extremely convoluted, and many tests should be done on patients to get a close to exact analysis. This has brought about modernized symptomatic apparatuses, expected to help the doctor in settling on essential clinical choices and subsequently an early determination. A significant region for such electronic devices is in the space of confusable infections. In this review, we proposed a genetic diagnosis system for the breast cancer disease in our health sector. Diagnosis of diseases must be handled with care since it's the initial stage of therapeutic actions towards eventual taking care of the disease; a mistake at this initial stage is disastrous and such adequacy must be ensured [1],[7],[10]. Diagnosis is the act of investigating a disease attributed to a particular symptoms and health conditions [10],[12] Clinical diagnosis in African content is been carried out with the symptoms of disease as a platform for diagnosing human disease in achieving optimal solution. Several factors such as lack of good and portable water supply, wide spread of infections in rural communities with modesty in waste management system, insufficient healthcare equipments/facilities and self-treatment for clinical diseases all contribute to the issues faced in African settings.[19], stated that "the administration of efficient healthcare services has been a major challenge in developing countries due to inadequate healthcare delivery personnel and the inappropriate diagnostic techniques often adopted". Medical Diagnosis can be defined as the process of finding out which disease or condition explains a person's symptoms and signs.

The proposed and adopted model is the GA. GA is a technique used in solving the problem of both constrained and unconstrained optimization associated with natural selection and biological evolution [16]. The

GA repeatedly modifies the individual population. The GA has promising implications in various medical specialties including radiology, radiotherapy, oncology, pediatrics, cardiology, etc.

The Fitness function in GA is recorded to be higher in value, attainable, and also equates to be a better solution in any problem domain. If the fitness is chosen poorly or defined inaccurate, the GA might be unable have a solution, or may find wrong solution. However, a major drawback of GA is that its convergence is very much dependent on its initial solution. This is also because; since GA mapps to a class of non-deterministic algorithms which the optimal solution a user gets from GA may varies for the algorithm at every same input data. So it is rather sub-optimal then. GA systems are systems constructed by genetic programming that imitate the process of natural evolution, to identify its structure and parameter. Therefore, soft computing and genetic programming (GP) techniques have been widely used to identify structure and parameters of genetic systems.

II. RELATED WORKS

The two major motivations in medical science are to prevent and diagnose diseases [10],[12]. Diagnosis of disease must be done with care since it is the first stage of therapeutic actions towards eventual management of the disease; a mistake at this stage is disastrous, and such, adequate care must be ensured [1],[7] Diagnosis becomes difficult in medical domain due to influence of medical uncertainties that arises from confusability in disease symptomatic presentation between two diseases. This confusability of these diseases stems from the overlaps in the disease symptomatic presentation and has led to misdiagnosis with various degrees of associated costs and in worst cases led to death [10]. In recent times, the negative effect of medical uncertainties has attracted attention about the emerging realities of this period where evidence based and patient- centered care has brought to fore the limitation of scientific knowledge.

[13] looked at a cloud-based Model for predicting and Diagnosing Breast Cancer. The study developed a cloud-based intelligent BCP-T1FSVM (Breast Cancer Prediction Model using Type-1 Fuzzy and Support Vector Machine) with 2 variations/models like BCP-T1F and BCP-SVM. The developed BCP-T1F-SVM system has employed two main soft computing algorithms.) Furthermore, the developed BCP-T1F-SVM expert system was specifically used to identify the different types and stages of cancer a person is suffering from. In the central role of the breast cancer disease, the BCP-SVM was used to diagnose the breast cancer at early stage; the BCP-T1F was also used to diagnose the breast cancer disease at the early stage, taking the stages and levels of the breast cancer disease into consideration. The evaluation and calculation used in the work shows that the BCP-SVM performed better and higher than that of the BCP-T1F in diagnosing the breast cancer disease.

[18] developed Fuzzy cluster means algorithm for the diagnosis of confusable disease. They built an expert system driven by the fuzzy cluster means (FCM) algorithm, the system accepted symptoms as input and provides the degree/level of membership of each disease in any confusable disease set. The model specifically was designed to diagnose patient infected with the alcoholic liver as a confusable disease. The data on alcoholic liver disease were collected and used in the development of knowledge base. Fuzzy logic and FCM algorithm was propelled as the inference engine. The system was implemented with CLIPS expert system shell and Java as the front end platform while Microsoft Access was used as the database application. The system gives a measure of each disease within a set of confusable disease. The system had a classification accuracy of 60%.

[12] adopted a hybrid methodology for differential diagnosis of confusable diseases, which was an application of an intelligent model, based on fuzzy logic technology was harnessed in the work to facilitate successful differential diagnosis of confusable tropical disease. The research specifically focuses on hepatitis, malaria typhoid fever and urinary tract infection as confusable diseases. Data from patients who were already diagnosed conventionally (and confirmed by laboratory tests) of hepatitis, malaria, typhoid fever and urinary tract infection were collected and the results show that the diagnosis carried out using the fuzzy system compares favourably with diagnosis arrived at conventionally by experienced physicians.

[1] researched on a framework for Early Differential Diagnosis of Tropical Confusable Diseases using the fuzzy Cognitive Map Engine. The study developed a soft-computing framework for the differential diagnosis of tropical diseases. It focuses on the development of a fuzzy cognitive map (FCM) model for early differential diagnosis of tropical diseases. [7] carried out a research using Analytic Hierarchy Process (AHP) model for Diagnosis of Tropical Confusable Diseases. The model mined the experiential knowledge of medical experts to develop an AHP model for the diagnosis of seven confusable diseases that are prevalent in tropical (mostly developing) countries. The system also takes cognizance of some risk factors that could pre-dispose an individual to infection. In addition, it recognizes the sematic causative relationships among symptoms, and can account for comorbidity in the seven diseases under consideration. The system was implemented in an Android environment; it recognizes the need for a friendly and simple user interface for the medical practitioner.

[11] presented Case-based Recommender Systems: A Unifying View. The work fostered a unifying structure to demonstrate case-based reasoning recommender system (CBR-RSs). CBR-RSs have complex designs and practices that is based on the CBR problem solving procedure/methodology in various manners.

The objective of the created structure is to delineate both the normal provisions of the different CBR-RSs just as the focuses were these frameworks take various arrangements. The proposed structure was determined by the analysis of some certain frameworks and techniqueincluding nine unique recommendation functionalities. A definitive objective of the system is to facilitate the assessment and the comparison of CBR-RSsand to provide a device to recognize open area for additional research/input.

[4] proposed a framework that permits CBR to make decisions about item/product similarities to work on the nature of their suggestions, recommendation and subsequently this kind of approach has demonstrated to be exceptionally effective in numerous online business settings, particularly when the requirements and inclinations of clients/users are badly characterized.

[8] investigated on CBR Systems inspired by social insects. As indicated by them, Case-based recommender systems that can learn about user preferences over time and automatically suggest products that fit these preferences" and presented a system called CASIS. In CASIS they combined case-based reasoning approach with a metaphor from colonies of social insects namely the honey bee dance. In nature, this metaphor is used to indicate the best nectar source among honey bees. Similarly they used it to retrieve the most similar case to the user's query. Their results show that this combination is effective when used in the retrieval step of the recommendation cycle: the most similar case is found by the "bees".

[6] presented CBR-S for personalized finance advisory. They further developed a business-based model to help brokers, individuals and bankers in predicting investment values termed as "wealth management system" to meet up with their investment goals and objectives but required some deep machine learning techniques to accurately predict the target variable with large volume of unstructured and structured data, which to a great extent require a profound information on the monetary space, a pattern in the space is the abuse of suggestion innovations to help monetary consultants and to work on the viability of the cycle.

[21] carried out a conversional web-to-case-based recommender model to aid the discovery of metadata because to search for online-resources in the study of big data was a problem.They further delineated that finding assets of interest in an enormous asset escalated climate is a difficult issue. In their paper they introduced the exploration on resolving this issue through the improvement of a recommender framework to help with metadata revelation. Their recommender approach utilizes Conversational Case Based Reasoning (CCBR), with semantic web markup language giving a standard structure to case presented. They presented their underlying endeavors in planning, designing and developing cosmology for an Earthquake Simulation Grid, to utilize these to direct case recovery, talk about how these are taken advantage of in a prototype application, and recognize future strides for the methodology.

[3] adopted a cased-based reasoning technique to express recommendation on similar items for customers with market basket analysis. According to them, "the most broadly utilized recommendation techniques will in general prescribe comparative things to those enjoyed/liked by a client/user before or products/items that comparable clients preferred dependent on the theory that clients have a steady conduct after some time and that comparable clients share comparative preferences.

III. METHODOLOGY

In this paper, we took on the Object Oriented Analysis and Design Methodology(OOADM).The Object Oriented Analysis and Design (OOD) is the most common way of utilizing an Object situated strategy to plan a registering framework or application. This procedure empowers the execution of a product arrangement dependent on the ideas of items. Object Oriented Analysis and Design (OOAD) technique fills in as a feature of the Object Oriented Programming (OOP) measure or lifecycle. The item arranged methodology consolidates information and cycles known as techniques into single elements called objects.

Dataset: The Datasets used in this research was gotton from [4][17]. This dataset contained a total of 568 items. 70% of the Dataset was used for training which amount to 398, while 30% was used for testing which amount to 171. The future attributes of the proposed system dataset includes: Mean radius as the mean of distances from the center to points on the perimeter. Mean texture is the standard deviation of gray-scale sample values in the test result. Mean perimeter is the mean size of the patient core tumor. Mean area and smoothness is the mean local variation in radius lengths. Mean compactness is the mean of perimeter^2/area-1.0. mean concavity point, mean symmetry, mean fractal dimension, radius error, tecture error, perimeter error, area error, smoothness error, compactness error, concavity error, concave points error, fractal dimension error, worst radius, worst texture, worst perimeter, worst smoothness, worst compactness, worst concavity, worst concave points and worst fractal dimension.

Table I: Breast Cancer Dataset [4],[17]

	Mean radius	Mean texture	. . .	Worst symmetry	Worst fractal dimension
0	17.99	10.38	. . .	0.4601	0.1189
1	20.57	17.77	. . .	0.2750	0.0890
2	19.69	21.25	. . .	0.3613	0.0876
3	11.42	20.38	. . .	0.6638	0.1730
4	20.29	14.34	. . .	0.2364	0.0768
. . .	. . .	. . .	. . .	. . .	. . .
564	21.56	22.39	. . .	0.2060	0.0712
565	20.13	28.25	. . .	0.2572	0.0664
566	16.60	28.08	. . .	0.2218	0.0782
567	20.60	29.33	. . .	0.4087	0.1240
568	7.76	24.54	. . .	0.2871	0.0704

The Case-Based Reasoning (CBR)

A case-based thinking is a strategy for tackling another issue by recollecting a past, comparative circumstance and reusing same data with the information on that circumstance. The idea of CBR accepts that comparable issues have comparative solutions. We employed the CBR to classify data/input preprocessing and formalization operations into mild, moderate, severe with critical cases of the breast cancer with genetic algorithm. The CBR system rules are built with the most common test that medical professionals uses to interprete the stages with TNM system. The reasoning system uses test diagnosis and scans sourced from lab scientistswith attributes of turmor(T), size_of_node(N) and metastasis(M) showing the spread of cancer to other parts in the human body. However, the main reasoning task is embodied in the CBR engine formulated with rules.

The stages of breast cancer disease

Stage 0: Theseare the mild-cases of breast-cancer where the size has spread from 1-4 axillary lymph nodes**.** The stage 0 referred to a non-invasive cancer which abnormal cells are deposited at the lining of the breast-milk duct. In stage-0 cancer; the breast cells have not spread at the outer part of the ducts or lobules with the surrounding breast tissues. The ductal Carcinoma in Situ is very early and is treatable, but if left undetected can spread to the breast tissues..

Stage-1: The stage-1 it is evident but contained only the area where the first abnormal cells began to develop. The breast-cancer detected is referred to as early stage and can effectively be treated. The stage-1 is grouped into stage IA and IB. It's treatable.

Stage 2: The stage 2 breast cancer called the moderate cases: where the lymph nodes size spread from 4-9 auxiliary lymph-nodes. The growth has only spread to the nearby lymph nodes. This stage has 2-groups: Stage 2A and 2B. The difference is the turmor-size and the lymph-nodes.

Stage 3: The stage-3 or severe-cases of breast-cancer is when it extends beyond the immediate region of the turmor and have invaded nearly lymph-nodes and muscles. This stage is grouped into sub-classes: Stage 3A, 3B & 3C. The difference is determined by t turmor-size and whether the cancer has spread to the lymph nodes and surrounding tissues.

Stage 4: The stage-4 or critical(worst case) breast cancer is not curable but usually tratable and current advances and medical technology shows the women with such issue of living longer by managing the disease as a chronic-illness. The cancer extended to the other parts or area like the bones, brain, liver and lungs.

Procedure for CBR System

The case-based reasoning is developed with the IF...THEN...ELSEIF statement using the GA. This clearly defined the stages of breast cancer as added advantage over the existing methods as the procedures used by the CBR engine.

Table II: Procedure for CBR System

```
If breast cells are found in the lining of the breast milk duct:
Cases= "mild"  # stage 0&1
elsif breast cells has grown  and extended to the nearby lymph nodes:
        Case= "Moderate" # Stage_2
elsif breast cells have invaded the lymph nodes and extends beyong the regional turmor:
        Cases= "Severe" # stage_3
elsif breast cells have spread to the bones, brain, liver and lungs:
        Cases= "Critical" # stage_4
Endif
```

The Genetic Algorithm (GA)

GA is one of the development methods in the field of AI motivated by human genetic cycles of passing qualities starting with one age then onto the next. It is utilized for advancement reason and different fields of study. We propose to prepare and test to learn the exactness of anticipating the objective variable utilizing the proposed GA in SVM. This will assist with further developing the exactness level via preparing GA with SVM for better element choice. We characterized an underlying populace about Breast cancer for every patients having their own arrangement of chromosomes and a capacity was called to group whether a patient have the confusable disease (Breast disease) or not. The chromosome were considered as more fit when it contains more survival points as chromosome 1 is viewed as more fit than chromosome 2.We selected those patients with no cancer for mating to produce off-springs and replaced the off-springs with those having the cancer disease from the population and repeated the process while fitness of fitted individuals population which are not high enough. We selected the fitted chromosomes from the population and produce off-springs. Initialization step was first created for initializing the population randomly given data from numpy pool of library with chromosoms in python. The function is used to compute the fitness for each chromosome as set of hyper-parameters which define a proposed problem with solution that the genetic algorithm is trying to solve. The Selection step considers the best fitted chromosomes as parents to pass the genes for generationsto come and create new population which retrunsthe variable called n_paraents. The Cross-over module is created to form new set of chromosome in combination with parents and add them to generate a new population set. The mutation process was initiated to alter 1 or more gene values of a chromosome in the new population set generated. This helps in getting more of the diverse opportunity obtained in the population for the generation to come. The training and testing of GA was done with feature initialization, fitness function, selection, cross-over and mutation having the best chromosoms and score values. The result of case-based reasoning is feed to the genetic algorithm with features and rules of mild, moderate, severe and critical cases of breast cancer.

Algorithm 1: Support vector machine(SVM)

Step	**Processes involved**
1	Start
2	Find candidate_SV with closest pair from classification (SV=>support vector)
3	If there are violating points then
4	Find violating_points
5	Compute the candidate_SV= candidate_SV + voilating_points)
6	If there is any $\alpha_p < 0$ due to the addition of c to S that gives negative then
7	Candidate_SV = candidate_SV/p
8	Repeat module to prune all the data points
9	end_if
10	end_if

Algorithm 2: The proposed Genetic Algorithm(GA)

Step	**Processes involved**
1	Start
2	Iniatialize(K=0) and create a population(P_k): of n randomly-generated individuals
3	Comput fitness(i) for each iϵP_k
4	**Selection**, Make selection with members of P_k and insert into P_{k+1}
5	**Crossover**:select and pair members of Pk; produce offspring, insert the offspring into P_{k+1}
6	**Mutation**: Select μ x n members of Pk+1; invert a randomly-selected bit for each
7	**Evaluate** Pk+1; Compute fitness(i) for each iϵP_k
8	K=k+1 // increment k by 1
9	Repeat while fitness of fitted individual in P_k is not high enough
10	**Return** fitness individuals from P_k

Evaluation Metrics: Several evaluation metrics are employed to ascertain the performance or success rate of learning algorithms. In this work, we employed detection accuracy, the confusion metrics, precision, recall and f1-score as diagnostic tools to measure the accuracy of both SVM and the proposed genetic algorithm. The detection accuracy is used measure the correct classifications from the overall number of cases. The precision: is tool adopted to measure positive classifications, Recall is a metric used to measure the false negative classifications and f1-score to measure the true positive and false positive regardless of false negative and false positive classification. Table 8 shows the diagnoses accuracy of breast cancer measure in percentage for genetic algorithm and SVM model obtained from the same testing dataset. The GA produced 92.40% metrics of accuracy higher than the existing SVM with 71.99% metrics of accuracy.

IV. RESULTS AND DISCUSSION

We are discussing in detail about the results of SVM and GA obtained from the proposed system training and testing dataset as presented below in figures and tables. The bar charts, tables and graphs are used for presentation of results and discussion.Table 1 shows summary of the Breast Cancer Dataset obtained from ice.uci.edu [4].This dataset contained a total of 568 items. 70% of the Dataset was used for training which amount to 398, while 30% was used for testing which amount to 171. The future attributes of the proposed system dataset includes: Mean radius as the mean of distances from the center to points on the perimeter. Mean texture is the standard_deviation of gray-scale sample values in the test result. Mean perimeter is the mean size of the patients core tumor. Mean area and smoothness is the mean local variation in radius lengths. Mean compactness is the mean of perimeter^2/area-1.0. mean concavity point, mean symmetry, mean fractal dimension, radius error, tecture error, perimeter error, area error, smoothness error, compactness error, concavity error, concave points error, fractal dimension error, worst radius, worst texture, worst perimeter, worst smoothness, worst compactness, worst concavity, worst concave points and worst fractal dimension.

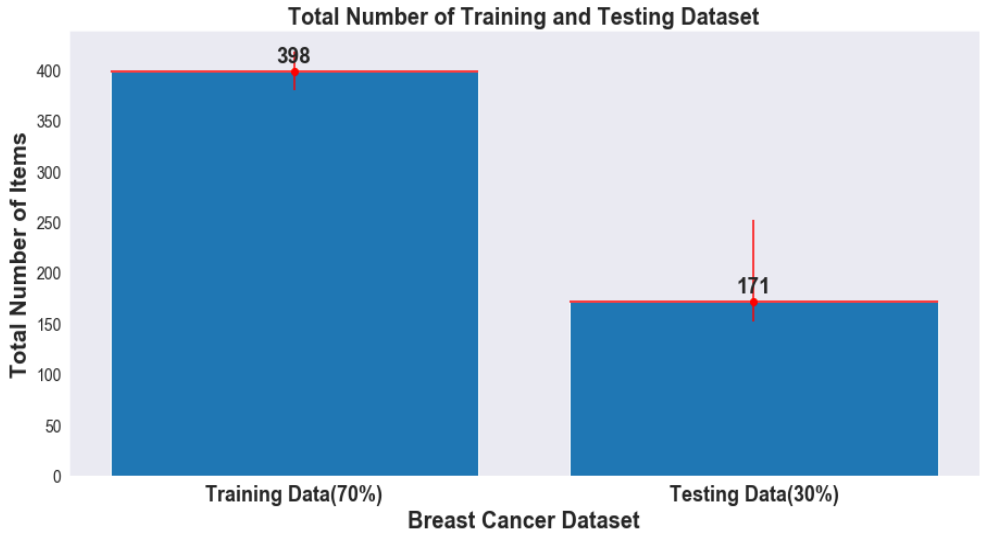

Fig 1: Breast Cancer training and testing dataset

Fig. 1 graph showing the total no. of training and testing dataset. The total dataset is divided into (70% of 569 = 398 items) as training and (30% of 569 = 171) testing dataset. This was done with the import sklearn model selection library in Python to include the train_test_split with input/target features.

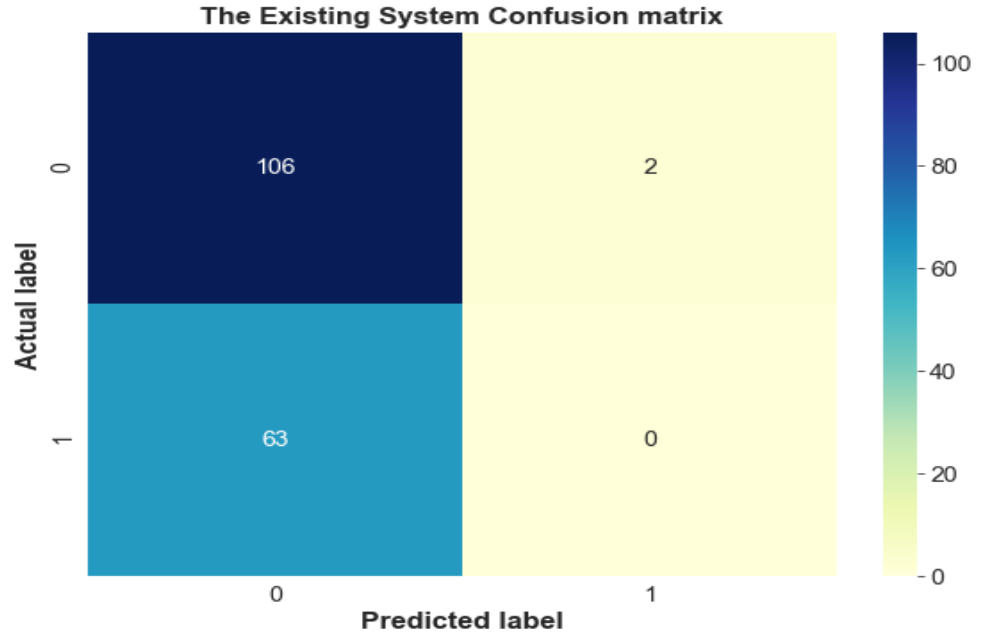

Fig 2: The SVM system confusion matrix

Fig. 2 depicts the confusion matrix of the SVM model with the correct predictions recorded on the main-diagonal and wrongly predicted values on the secondary diagonal called the off-diagonal elements. The total no. of correct predictions = 106+0=**106** and wrongly predicted = 63+2= **65** as shown in figure 4.2 above.

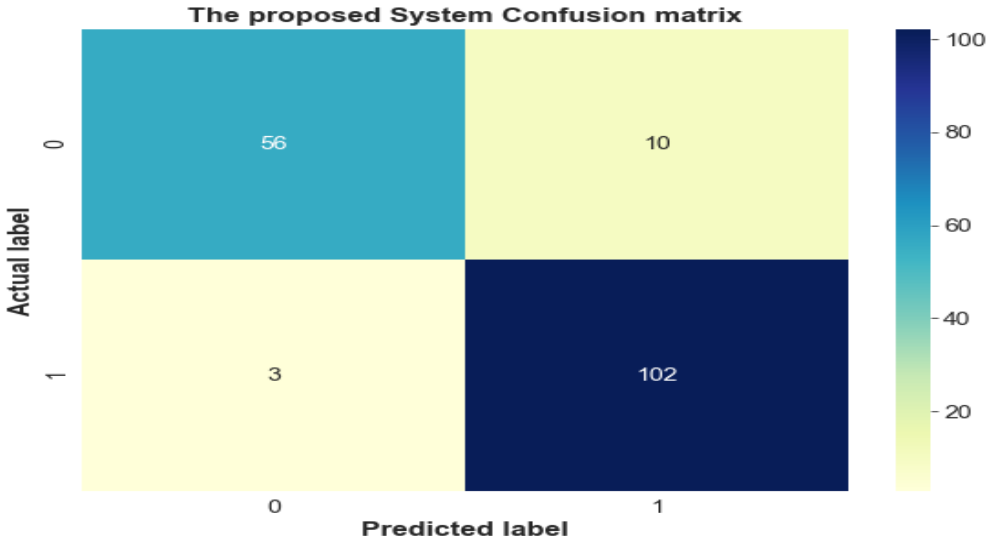

Fig 3: The GA confusion matrix

Fig. 3 depicts the confusion matrix of the (GA) model. The main diagonal elements or values showing the total number of correctly predicted values that are equal to the actual or true values. The secondary diagonal showing the wrongly predicted values which falls-off the main diagonal called the off-diagonal elements. The higher the diagonal values the better the predictions. The total number of correct predictions = 56 + 102 = 158 and wrong predictions = 10 +3 = 13

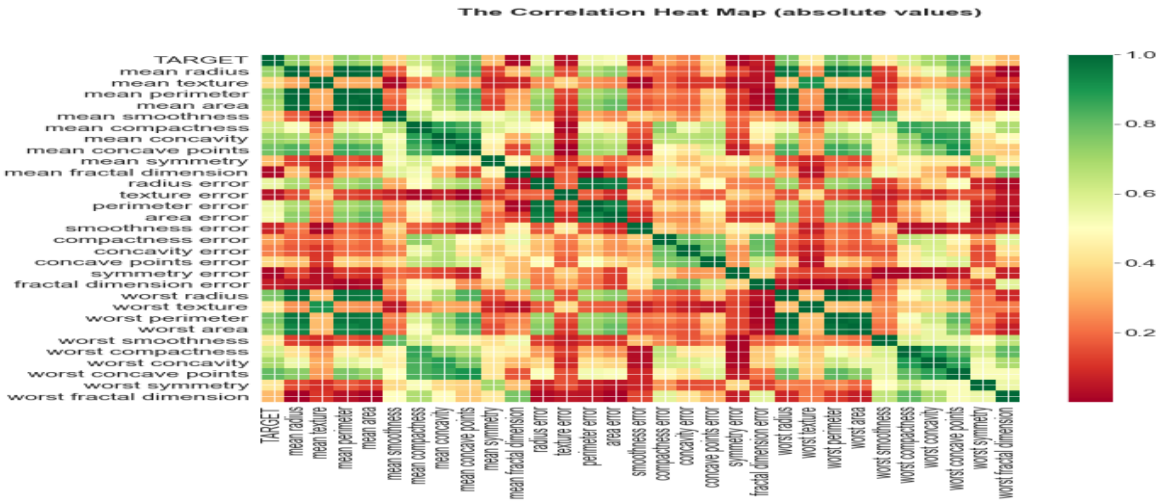

Fig 4: The correlation graph of proposed system heat map

Fig. 4 is the correlation matrix used to measure the relationship between breast cancer feature variables. The matrix depicts the linear-correlation for all possible pairs of attributes in breast cancer dataset. There is strong-relationship in the main diagonal and other pairs as shown in the correlation heat map. There is strong-correlation between the mean radius and mean perimeter, mean area and mean perimeter.

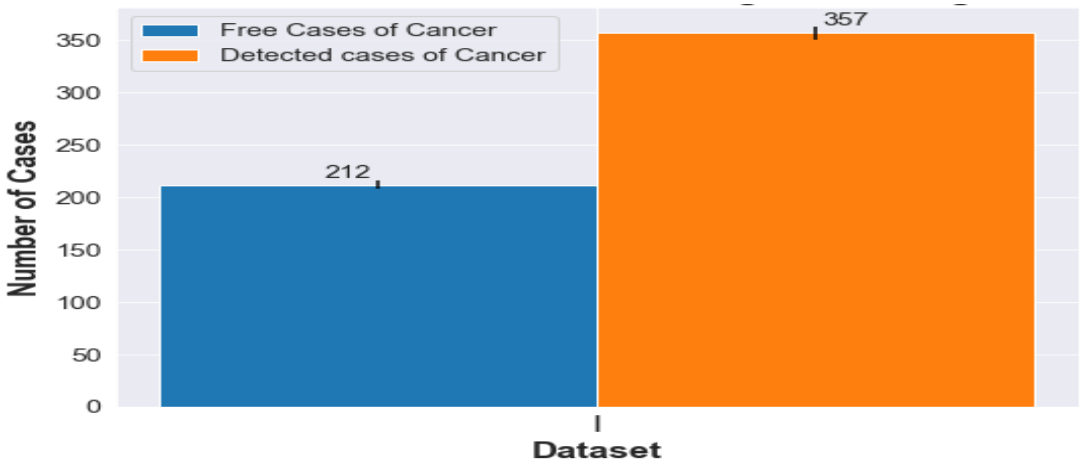

Fig 5: The breast cancer detection chart

Fig. 5 contain total no. of breast cancer detected cases and those free from the disease as gotten from the same dataset by the GA after training stage. From the result 357 of the total dataset was detected to have and suffering from breast cancer and while 212 items of total dataset was detected to free from breast cancer..

Table III: The classification report of SVM

	Precision	Recall	F1-score	Support
0	0.64	0.98	0.77	108
1	0.60	0.05	0.09	63
Accuracy			0.64	171
Macro_avg	0.62	0.51	0.43	171
Weighted_avg	0.62	0.94	0.52	71

Table III shows the classification report of the existing SVM measured in terms of precision, recall and f1-score for predicting the target. In precision; the free case of breast cancer produced 0.64 and those having the breast cancer disease gave 0.60. For recall 0.98 free case and 0.05 having the disease, For f1-score 0.77 free case and 0.09 having the disease as classified by the existing system

.

.**Table IV:** The classification report of genetic algorithm(GA)

	Precision	Recall	F1-score	Support
0	0.93	0.86	0.90	66
1	0.92	0.96	0.94	105
Accuracy			0.92	171
Macro_avg	0.93	0.91	0.92	171
Weighted_avg	0.92	0.92	0.92	171

Table IV is the classification report of the proposed GA in terms of precision, recall and f1-score for predicting the target variable. The free case of breast cancer produced 0.93 and those having the breast cancer disease gave 0.92 classifications in precision. Recall 0.86 free case and 0.96 have disease. For f1-score 0.90 free case and 0.94 having the disease as classified by the proposed system model.

Table VI: The Diagnostic Metric of GA and SVM

	ALGORITHM	ACCURACY
0	Genetic Algorithm	92.40
1	SVM	71.99

Table VI shows the diagnoses accuracy of breast cancer disease measure in percentage for GA and SVM model obtained from the same testing dataset. The GA produced 92.40% metrics of accuracy higher than the existing SVM with 71.99% metrics of accuracy. The Proposed System GA performs better than the Existing System SVM in terms of speed and accuracy.

V. CONCLUSION

Python was used as a tool to evaluate the performance of the existing SVM and the proposed GA in tern of accuracy and speed using evaluation metrics to ascertain the success rate of the learning algorithms. The result analysis encompasses the end-users who tested the model, diagnosed Confusable Disease: breast cancer issue and the fatality rate. Parameters for analysing the results encompassed the speeds in validating a registered

patient, symptoms response classification, symptoms processing speed, and result authentication. The result values for the Existing model and mentioned parameters are 23, 31, 28 and 12 respectively, while that for the proposed model is 17, 14, 10 and 5 respectively. The proposed GA model perform better than the existing model in terms of detection accuracy, precision, F1-Score and recall. The GA produced 92.40% matric of accuracy higher than the SVM with 71.99% after training. Therefore the CBR-RS-GA gives higher accuracy value than that of exiting BCP-SVM.

REFERENCES

[1]. Akinwuwesi, B. A., Faith, M. E. U., Taiwo, A., Flora A., Stephen F., Moses O., and Joseph, O.(2016).A framework for Early Differential Diagnosis of Tropical Confusable Diseases using the Fuzzy Cognitive Map Engine. International Conference on Computational Science and Computational Intelligence (CSCI) /10.1109/CSCI.2016.306

[2]. Amer A. (2018), A Collaborative Filter Recommender System for Course Selection, International Journal of Advanced Computer Science and Applications, 7(3), 166 – 175

[3]. Anna, G., Andrew, M. S., Pang, W. K., Zhenghao, C., Maneesh, B., Bipin, S. and Andrew, Y. N. (2015).A Case-Based Recommendation Approach for Market Basket Data, Journal of Software Development (JSD), 4(7), 95 – 99.

[4]. Blake C. L. and Merz C. J. (1988) UCI repository of machine learning databases, http://www.ics.uci.edu/~mlearn/mIrepository.html. University of California, Irvine Department of information and Computer Sciences

[5]. Barry S. (2017), Case-Based Recommendations, an International Publication for University College,Dublin, https://www.researchgate.net/publication/220800382

[6]. Cataldo M., Pavlos L., Nikolaos N., Jamie A., Marijke T., Basma E. R and Ismail R. (2015), Case-Based Recommender Systems for Personalized Finance Advisory, International Journal of Computer Applications (IJCA), 4(7), 15 – 21

[7]. Chukwudi, N., Fikru, D., Boluwaji, A., Moses, O., Faith, M. U. (2017). Analytic Hierarchy Process (AHP) Model for Diagnosis of Tropical Confusable Diseases. International Conference on Computational Science and Computational Intelligence (CSCI)/10.1109/CSCI.2017 .306/18342651

[8]. Daniela, S. M., Ashalatha, E., Zaki T., Abdus A., Aderemi J. J., and Bethany L. (2014), Cases-Based Recommender System inspired by Social Insects, an online article Published at https://www.researchgate.net/publication,228851535

[9]. Dietmar J. (2017), Fast Computation of query relaxations for knowledge-based recommenders, InternationalJournal of Computing Technology (IJCT), 2(9), 47 – 51

[10]. Ejaita, O. A., and Asagba P. O. (2017), " An Improved Framework for Diagnosing Confusable Diseases Using Neutrosophic Based Neural Network." Paper https://digitalrespiratory.unm.edu/nss_journal/vol16/iss1/7

[11]. Fabiana, L., Francesco R., Daneil, E. Stefana, T., Ayman A., and Debajyoti L. (2019), "Case-Based Recommender Systems: A Unifying View," an International Conference: Paper, https://www.researchgate.net/publication/200051089

[12]. Faith, M., Uzoka, E., Okure U. O., Udoinyang G. I. and Joseph.O.(2014). A hybrid methodology for differential diagnosis of confusable diseases. Health Information & Technology Conference.

[13]. Farrukh, K., Muhammah, A. K., Sagheer A., Atifa A., Shahan, Y. S., Abdul H. K., Muhammad A. S.andMuhammed H.(2020), " Cloud-Based Breast Cancer Prediction Empowered with Soft Computing Approaches." International Journal of Healthcare Engineering, Paper https://doi.org/10.1155/2020/8017496

[14]. Giovanni S. (2015), Case-Based Recommender Systems for Personalized Finance Advisory, InternationalJournal of Computer Applications (IJCA), 4(7), 15 – 21

[15]. Guang Z. (2016), Recommender System Application Developments: A Survey, International Journal of Computer Application (IJCA), 2(4), 30

[16]. Harik G.R., Lobo, F. G., Sastry, K. (2006). Linkage Learning via Probabilistic Modeling in the ExtendedCompact Genetic Algorithm (ECGA).Scalable Optimization via Probabilistic Modeling.

[17]. Studies In Computational Intelligence, 33.pp.39-61. Doi:10.1007/978-3-540-34954-93. ISBN978-3-540-34953-2.

[18]. Hussein A. A.(2002) An evolutionary artificial neural networks approach for breast cancer diagnosis. Artificial Intelligence in medicine.

[19]. James, G. G., Okpako A. E., and Ndunagu, J. N. (2017). Fuzzy cluster means algorithm for the diagnosis of confusable disease. ISSN Print, http://www.publication.ncs.org/.23 (1), 2006-5523.

[20]. Jan, A. (2019), The architecture of Intelligent Case-Based Reasoning Recommender System (CBR-RS) Recommending Preventive/Corrective Procedures in the Occupational Health and Safety Management System in an Enterprise

[21]. Jieli E. J., Donghwoon, K., Hyunjoo K., Jinoh K., Sang, C. S., Ikkyun, K. and Kuinam, J. K. (2015).Personalized recommendation System: An application case based reasoning. In Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in System recommendation. ACM Press, New York, NY, USA, 401–410.

Jatto Joshua O, et. al. "Case-Based Reasoning Recommender System for Diagnosing Breast Cancer Disease: Confusable Disease Using Genetic Algorithm." *The International Journal of Engineering and Science (IJES)*, 11(4), (2022): pp. 01-09.