Ilkogretim Online - Elementary Education Online, 2020; Vol 19 (Issue 2): pp.1643-1660
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2020.02.696747

Comparative Analysis For Cancer Prediction Using Genetic Algorithm To Other Machine Learning Algorithms

Rishm Dept. of Computer Science and Engineering Gurukashi university Talwandi sabo Bathinda, India rishm.mehta@gmail.com

Dr. Misha Mittal MM Engg College, Mullana Ambala

Dr. Dinesh kumar Dept. of Computer Science and Enginnering Gurukashi University Talwandi sabo Bathinda, India dineshkumar@gku.ac.in

Dr. Ashwani Sethi Guru Kashi University, Talwandi sabo Bathinda, India

Abstract—Cancer is one of the leading causes of death these days. There are number of reasons which cause cancer but classification and predicting of cancer at the right stage is very important. This research work mainly focuses on early prediction of cancer using different approaches of machine learning in comparison with genetic algorithm to predict cancer at early stage. Genetic Algorithm optimizes the neural network by involving connection weights. The Breast Cancer Wisconsin (Original) Data Set is taken from UCI machine learning repository and it is trained and tested by applying different algorithm like Random forest, Naïve Bayes, Artificial neural network and Genetic algorithm. Random forest is used as an ensemble learning algorithm which is applied on the microarray data of cancer to achieve good accuracy and reliable performance, For predictive modeling Naïve Bayes algorithm is also used based on weight concept by assigning weight on dataset of breast cancer taken from UCI machine learning repository and Artificial neural network uses back propagation algorithm which has different neurons in hidden layer to analyze output by calculating the error using weight adjustment method ,it reduces error between the required and actual output. From experimental analysis genetic algorithm provides topmost accuracy of 97% where as other algorithms like artificial neural network provides accuracy of 96%, Naïve bayes provides accuracy of 94% and Random forest provides accuracy of 95%.

Keywords: Cancer, Genome, Genetic, targeted sequencing, Prediction, Machine Learning

I. INTRODUCTION

In the field of computer science, machine learning is a subset of Artificial intelligence. It is used in many areas such as business, medical sciences, research areas etc. In medical fields machine learning techniques can be best used for prediction, estimation, decision making. In machine learning there are two main learning approaches supervised and unsupervised learning. In supervised machine learning techniques data is divided into inputs or independent variables and outputs or dependent variable. Here the learning process is carried out by predicting values of the output from input values and for this purpose training set of data is employed that guides the process of learning[1]. Three different machine learning algorithms are used here for early stage of cancer prediction like Random Forest, Navie Bayes and Artificial neural network. Genetic algorithm is also used here to predict cancer and then comparison of results is made for both machine learning techniques and genetic algorithm [2]. Machine learning techniques help in extracting important biological information from DNA, RNA. Genetic algorithms have stochastic nature. Genetic analysis has become important in the field of clinical practices and different researches like sequencing of DNA etc.

A. Genetic Algorithm

Genetic algorithms used in machine learning to provide optimized solutions for making decision and obtain better result. These algorithms consider possible solutions from given population. Test methods of genetics have become accessible to be performed in small size labs after Human Genome completion. Genetics and selection are two mechanism of genetic algorithm on which population is based upon and their solution is given in the form of genes. Members from current population and mating two different solutions together produce new solution. This is done to get better solution and others are discarded [5]. It always finds a space for potential solution and is called probabilistic method of search. These algorithms are used in machine learning to find potential solutions that will solve the given. It is motivated from the principals of evolution and natural selection; it's a randomized searching method and also optimized technique.

B. Steps of Basic Genetic algorithms

There are different forms through which implementation of basic genetic algorithms can be done but researchers hardly use simple steps instead they use extensions as outlined below:

- It first uses initialization operator which tell from where to start the search and if new regions of search are need the recombination operator is used.

- Initial populations are considered first in which we have number of chromosomes that are made from genes then fitness function is used on every chromosomes and goodness is identified [6].
- Clustering problem is solved using crossover and mutation
- Elitist operation is applied by genetic algorithm in which new and previous population is compared
- All these steps of selection, crossover, mutation and elitist operation go step by step until we get satisfactory result as shown in Figure.1.

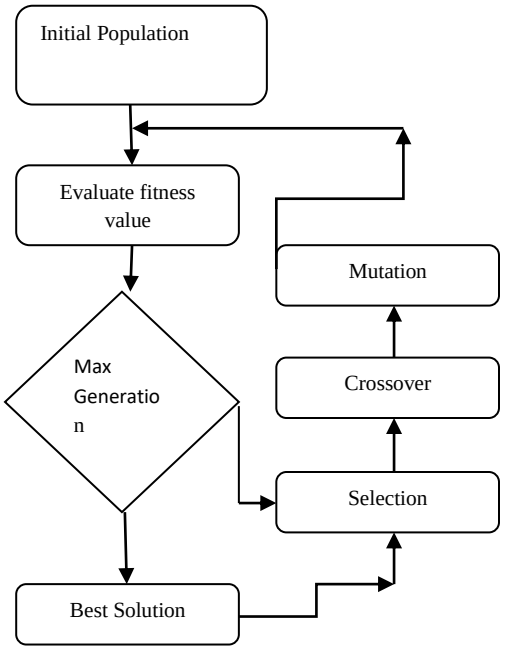


Fig.1 Flow diagram of GA

C. Genetic Algorithm Framework

There exists number of different variations when work with genetic algorithm and genetic programming. It is a meta-heuristic algorithm which is used to generate solutions with high quality, for optimization or for search problems [7]. Here given is simple form.

Algorithm1: Genetic Algorithm Framework

Begin
N=0

1. Set random initialization of population _set p(n)

2. Identification of fitness value from population_set. p(t)

   While n=n+1 **do**
3. Root element selection from parent_set p(n)
4. Apply crossover operation on parent_set to generate children (n+1)
5. Perform mutation operation (n+1)
6. Identify the fitness of population determine(n+1)

   end while

   until best individual found in the population_set

This process is called "generational" as it is partitioned into different generations which lead change in entire population. Also offspring's main motive is to bring few changes in the string at each step and generate "mating pool" which uses recombination.

D. Optimization of Genetic Algorithm

The research uses genetic algorithm that is based on reconfiguration of gene to solve large problems.

Coding proposal: The research uses genetic algorithm that is based on reconfiguration of gene to solve large problems. Fitness function confirmation [7]: Error functions reciprocal is used to access individuals. Fitness function F= 1/Ei2

Where Ei is error function of network

$$Ei=1/2\sum_{k=1}^{m} \quad \sum_{j=1}^{n3} \quad (Cjk - Cj)2 \qquad (1)$$

In the above equation 1 is denoted as sample quantity of cluster, cj as actual output, Cjk as ideal output.

II. Literature Review

Mehrnaz Ronagh,et al.(2019)[1] has proposed a technique of novel hybrid optimization for solving the problem of inverse scattering which make use of binary particle swarm optimization or binary genetic algorithm(BGA). This study is done to develop a algorithm for diagnosing the malignant tumor that is present in the breast. In this paper, for the calculation of electromagnetic scattering a tomography method is used for finite difference time domain method. Results shows that proposed algorithm is 4 times better than regular binary genetic algorithm and also FDTD/hybrid method easily construct heterogeneous and breast tissues so that an output can provide quantitative image. Proposed technique has reconstructed heterogeneous breast tissues that provide quantitative image also its capable of detecting the size and location of tumor.

Pragya Chauhan, et al. (2018) [2] has proposed a new technique to genetic algorithm based method which has helped to overcome the problem of average method of classical weight. Genetic algorithm based weighted average ensemble method is used with different type of

models for prediction. Breast cancer is such a problem that can be solved using different machine learning models like decision tree, random forest, support vector machine etc. To improve the accuracy of breast cancer prediction there is need to combine multiple models together. In this research comparison has been made between different genetic techniques: Particle swarm optimization, Differential evolution and the result show that genetic algorithm shows better result for weighted average method also comparison is made between classical_ensemble and genetic algorithm based weighted average method in which genetic algorithm based weighted average method performs better.

T. Cithambaram, et al. (2017) [3] Brain tumor is a major problem that need to be cured by classifying it into two categories primary and secondary tumor. Primary tumor is one that originates inside the brain and secondary tumor is one that originates and spread from other parts. The aim of this research was to develop interaction CAD system which helps in classifying different tumor. Research is done on the dataset of 350 post con MR images and publically available post contrast 280 images. Genetic algorithm is used to select optimal features. Implementation of two hybrid machine learning models named GA using support vector machine and GA with neural network. SVM enhancing overall accuracy by 79.3% to 91.7% and ANN provides accuracy from 75.6% to 94.9%.

A.H Beg, et al. (2016) [4] discussed that Clustering is a method used when we want to group similar records in one place called cluster and dissimilar in another but major difficulty comes in clustering when we do not know in advance that how many clusters will be needed in advance and also most popular method of clustering that is K-mean stucks at local optima. All these limitations are looked after by Genetic Algorithms that are based on clustering techniques. This paper has discussed different clustering techniques using genetic algorithm and applied in various fields .Genetic algorithm based techniques have been reviewed that make use of real life applications like Gas company, cellular network segmentation , medical fields etc. The review shows that two third of the techniques did not require users that will define number of clusters.

Ming Chen, et al. (2008) [5] has inferred that Data Mining is a way of examining a data that already exists and helps in generating new information from that data. Classification sometimes gets difficult in data mining. Classification is the major problem in data mining .In this paper, main focus is on gene reconfiguration by using optimizing classifiers of neural network that use Genetic Algorithm. Shift reverse logic crossover operation is used and improved genetic algorithm is used. BP algorithm is used it has slow speed and run into local minima easily buy genetic algorithm used overcome this problem of local minima. In this research first Simple Genetic Algorithm (SGA) is used for network structure and

then for gene reinforcement genetic algorithm is used .The result shows that improved genetic algorithm will improve classifying veracity.

III. MACHINE LEARNING ALGORITHMS

Machine learning Algorithms are applied by researchers for the prediction of cancer survivability rate. These algorithms are well accepted and work well for proper diagnosis and predication. An ensemble of Random Forest, Naïve Bayes and Artificial Intelligence can be used for predicting cancer at early stage.

A. Random Forest

It is an ensemble technique which integrates the series of K_learning model for the creation of better prediction model. Random forest algorithm uses decision tree classifier for randomization [3]. This helps in minimization of bias, tolerates outliers and also helps in avoiding over fitting. In this paper, random forest has been used for the prediction of cancer . The independent training sample $\mathcal{L}p$ = {(X1,Y1), ..., (Xp,Yp)} are identically distributed [0,1]l × R valued over random variables (l ≥ 2). It consists of random based regression tree.
{ rp(x, γq, $\mathcal{L}p$), q ≥ 1}

Where γ1, γ2…are identically independent.When random tree is merged it gives aggregate regression as given in equation 2

$$rp(X, \mathcal{L}p) = E_{\gamma}[r_p(X, \gamma, \mathcal{L}p)] \quad (2)$$

where E γ is random parameter, $\mathcal{L}p$ is data set

Algorithm 2: Random Forest

1. Select coordinates at each node X= (X(1)……X(l)), with probability of qpk ∈ (0,1)
2. Select the coordinate to find the midpoint at each side of node
3. Random tree rp(X, γ) provides output above Ym, having Xm vector of random partition:

$$rp(X, \gamma) = \frac{\sum_{m=1}^{p}[xm \in Ap(x\,\gamma)]}{\sum_{m=1}^{p}[Xm \in Ap(x\,\gamma)}\, 1Hp(x, y)$$

Where Hp(X, γ) is given in equation below

$$Hp(X, \gamma) = \sum_{m=1}^{p} 1[xm \in Ap(x\,\gamma)] \neq 0$$

4. Random forest regression takes form $rp(X,\gamma)=E\gamma[rp(X, \gamma)] = E\left\{\frac{\sum_{m=1}^{p}[ym1[\, xm \in Ap(x\,,\gamma)]}{\sum_{m=1}^{p}[Xm \in Ap(x\,\gamma)}\, 1Hp(x, y)]\right\}$

B. Naïve Bayes

It is a decision support model which uses Bayesian Network and Bayes theorem with independent assumptions among predictors.

For predictive modeling Naïve Bayes is the technique that analyze the data within each attribute and for the events that evaluate conditional probability [5]. During training probability for each class is calculated to estimate the modeling of instances in training dataset so that class labels can be assigned to problem instance Naive Bayes provides a method for the calculation of posterior probability P(c/x), P(c), P(x) and P(x/c). NB classifier defines the effect of predictor(x) on class (c) which is independent of other predictor values. Equation 3, 4 shows the class conditional independence.

$P(c/x) = P(x/c)P(c)/P(x)$ (3) $P(c/X) = P(x1/c)^* P(x2/c)^* \ldots\ldots ^*P(xn/c)^* P(c)$ (4)

Here, P(c/x) is target class posterior probability; P(c) is prior probability, P(x/c) Probability of predictor class, P(x) Predictor prior probability.

Algorithm 3: Naïve Bayes classifier

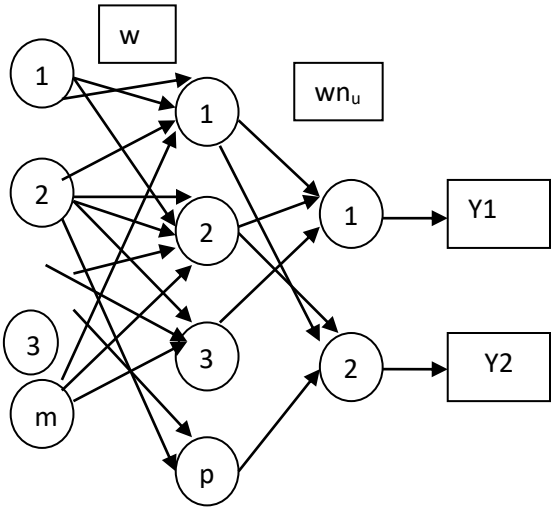
Begin

1. Take two independent events A and B : P (A,B) = P(A)P(B)
2. Evaluate result through : $P(y|x1,\ldots\ldots xn)=P(x1|y)P(x2|y)\ldots.P(xn|y)\ P(y)/\ P(x1)P(x2)\ldots P(xn)$
3. Equation of step 2can be expressed using $P(y|x1\ldots\ldots xn)=P(y)\ \Pi\ ni=1P(xi|y)/\ P(x1)p(x2)\ldots.\ P(xn)$
4. Remove the term as denominator remains constant by $P(y|x1\ldots\ldots xn)\ \alpha\ P(y)\ \Pi\ ni=1P(xi|y)$
5. Create classifier model by finding probability of set of inputs for all values of variable y :
   Y= argmaxy $P(y)\ \Pi\ ni=1P(xi|y)$
   Calculate P(y) and P(xi| y)

C. Artificial Neural Network

Artificial neural network is based on Artificial intelligence that relates to the human brain function. These networks interpret nonlinear data and biological data is usually nonlinear. The benefit of using neural networks in such situations is that it has a good tolerance capacity for noisy data and also easily classifies patterns that are not even trained [21]. In this research, back propagation method is used to diagnose cancer. It has three layers input layer, hidden layer, and output layer. All these layers perform a different function like the input layer receives input and send it to the hidden layer; the hidden layer adds computational weight and provides net input which is applied to the activation function which provides actual output. Artificial neural network has been used in various areas like in prediction, economical modeling, medical applications, bioinformatics, etc.

- Back-propagation: In medical science, for prediction bases back- propagation (multi layer neural network) network is used. Here network travel from input neuron to output neuron as shown in Fig.2.



Input Neuron Hidden Neurons Output

Fig.2.Three-Layer Artificial Neural [21]

Network representation with Sx_m Input Neuron, W_{ih} Hidden Neurons and W_{ho} Output Neurons[21].

Algorithm 4: Artificial Neural Network

1. Create weights arbitrarily for multilayer neural network.
2. Assume E is training example and C as actual output
3. Calculate feed-forward propagation for both hidden and output layer.
   h_i, i= 1, 2, 3… p (Hidden layer neurons)

   $h_i = W_{inbias} + \sum_{i=1}^{m} u_i W_{ih}$
4. Sigmoidal function as activation function: $F(x)=1/(1+e^{-x})$
5. Calculate hidden node activation values : $u_i=f(h_i)=1/(1+e^{-hi})$
6. Calculate both input weight and output node value O_j ,j=1,2,3….n
   $Oj=Woutbia=s + \sum_{j=1}^{n}$ hiwh,o
7. Activation value of output node :$U_0=f(oj)=1/(1+e-0j)$
8. After pass one , calculate the error using weight adjustment method of back propagation to reduce error between wanted and actual output.

Error=0.5*(Odesired- Oactual)2
Output neuron δ 0 = (Ci-U0) u0 (1=u0)
Sigmoidal derivative u0 (1=u0)

9. Update network weight at each layer
   Hidden to output layer Wi*j=Wi,j+ pδ0 ui
   Hidden to input layer Wi*j=Wi,j+ pδi ui

   P is represented as learning rate

IV. MOTIVATION

Machine Learning techniques help in identification of cancer. A framework of identifying cancer helps doctors to save person from increasing effect of cancer by predicting it at early stage. To improve the performance of the cancer prediction framework, genetic algorithm and other machine learning techniques can be implemented on the dataset to achieve optimum accuracy of the result from these algorithms. In this research, the use of genetic algorithm in comparison with random forest, naïve bayes and artificial neural network is made on the basis of various performance parameters.

V. RESEARCH METHODOLOGY

The methodology of given technique is shown in Figure.3. The process of implementation is divided into three parts:

- Data randomization
- Model selection, training, and testing
- Checking the performance using various performance parameters

Here, data from the dataset is divided into two sets training and testing set. By using different algorithms of machine learning and genetic algorithm on the same dataset performance is evaluated using different performance parameters like accuracy, precision, sensitivity, specificity and Fpr and best algorithm for the cancer prediction is selected on the basis of accuracy estimation. Following steps given in the Fig.3 are followed to find the accuracy.

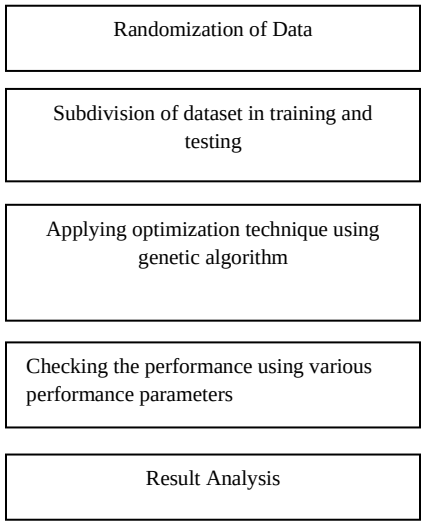


Fig.3. Methodology Used

VI. PARAMETERS FOR COMPARISON

- Accuracy: The percentage of total number of predictions that were correctly predicted
  Accuracy = TP + TN /TP + TN + +FP+FN
- Precision: The percentage of positive values that were correctly predicted.
  Precision =TP/TP + FN
- Sensitivity: It specifies what percentage of positive cases are covered by algorithm
  Sensitivity (Sens) =TP/(TP + FN)
- Specificity :The percentage of unpredicted values which are correctly predicted
  Specificity (Spec) =TN/TN + FP
- Fpr : Combination of precision and recall
  F1 score = (2 Precision*Recall) / (Precision + Recall)

VII. DATASET DESCRIPTION

In the research here, Breast Cancer Wisconsin (Original) Data Set is taken from UCI machine learning repository. Table 1 shows the core attributes. Breast cancer Wisconsin dataset consists of data with Multivariate characteristics in which number of instances are 699 and number of attributes are 10. Dataset also describes the malignant class and benign as positive and negative (1 and 0) respectively.

TABLE I. Dataset Description [25]

Attributes	Data Description
Sample code	Id number
Thickness of clump	Range(1-10)
Cell Size uniformity	Range(1-10)
Cell shape	Range(1-10)
uniformity	Range(1-10)
Marginal Adhesion	Range(1-10)
Cell size Epithelial	Range(1-10)
Bare Nuclei	Range(1-10)
Bland Chromatin	Range(1-10)
Normal nucleoli	Range(1-10)
Mitoses	Benign=0
Cancer	Malignant=1

VIII. RESULT

A. Random Forest

Random forest generates the classification for the cancer prediction with various parameters like true positive, true negative, false positive , false negative as mentioned in the confusion matrix given below:

- Confusion matrix: The confusion matrix for random forest is given below with certain values like: True negative value=106, False positive value=5, False Negative value= 2 and True positive value=58.

  [[106 5]
  [2 58]]

- Random Forest Performance measure

TABLE II. Random Forest Performance measure

	Random Forest
Accuracy	95%
Precision	95%
Sensitivity	97%
Specificity	90%
Fpr	5%

Table 2. Show the results for Random Forest for different performance parameters. Accuracy, Precision, Sensitivity, Specificity and Fpr are different parameters that are used to measure performance for both training and testing dataset values and the Accuracy for the Random forest is evaluated as 95%.

- Bar Graph for Random Forest

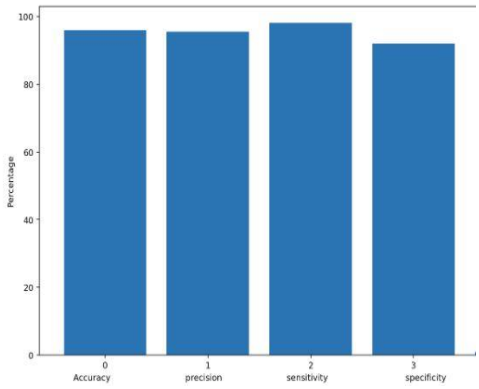


Fig.4. Random forest performance parameter Bar Graph

Fig. 4 shows the bar graph for random forest which provides the percentage level of various performance parameters.

B. Naïve Bayes

Naïve Bayes also generates the classification for the cancer prediction with various parameters like true positive, true negative, false positive , false negative as mentioned in the confusion matrix given below:

- Confusion matrix: The confusion matrix for random forest is certain values like: True negative value=99, False positive value=2, False Negative value= 8 and True positive value=62.

  [[99 8]
  [2 62]]
- Naive Bayes Performance measure

TABLE III. Naive Bayes Performance Evaluation

	Naïve Bayes
Accuracy	94%
Precision	95%
sensitivity	97%

specificity	90%
Fpr	5%

Table3 show the results for Naïve Bayes for different performance parameters. Accuracy, Precision, Sensitivity, Specificity and Fpr are different parameters that are used to measure performance for both training and testing dataset values. The Accuracy for the Naïve bayes comes out to be 94%.

- Bar Graph for Naïve Bayes

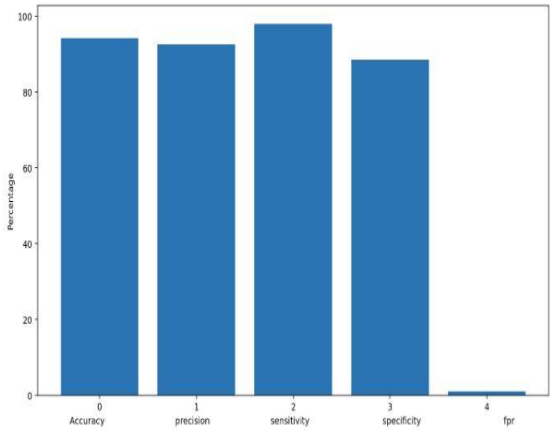


Fig.5. Naïve Bayes performance parameter Bar Graph

Fig. 5 shows the bar graph for Naïve Bayes which provides the percentage level of various performance parameters.

C. Artificial Neural Network

- Line Graph Values for Artificial neural Network Accuracy is given in the Table 4.

TABLE IV. Artificial neural Network Accuracy Evaluation

	Artificial neural Network
Training Accuracy	96%
Validation Accuracy	94%

Table 4 shows the results for Artificial neural network for its training accuracy and validation accuracy. The training accuracy for ANN is evaluated as 96% and Validation accuracy is evaluated as 94% as shown in line graph given below:

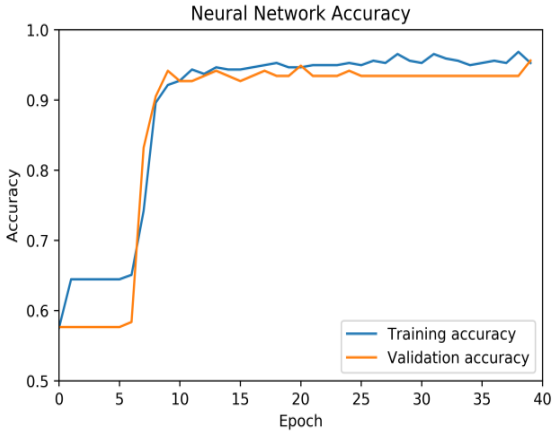


Fig.6 Neural Network Accuracy Line Graph

Fig. 6 shows the line graph for Neural network accuracy which depicts the values in different colors for both training and validation accuracy where training accuracy is shown in blue color and validation accuracy in orange color.

- Loss Line Graph Values for Artificial neural Network

  Loss line graph values for artificial neural network are provided in Table5.

TABLE V. Artificial neural Network Loss Evaluation

	Artificial neural Network
Training Loss	0.16
Validation Loss	0.18

Table 5 shows the results for artificial neural network for its training loss and validation loss. The training loss for ANN is evaluated as 0.16 and Validation loss is evaluated as 0.18. These values are graphically represented using line graph as shown in Fig.7 below:

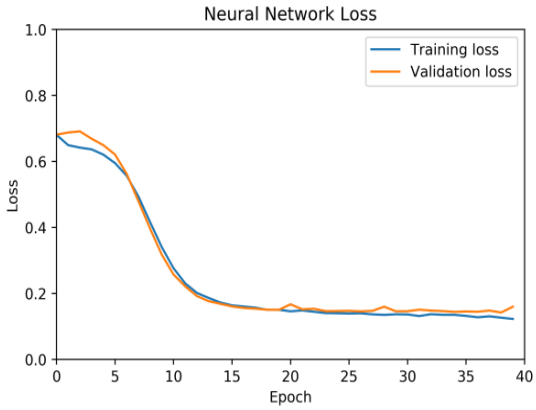


Fig.7 Neural Network Accuracy Loss Graph

D. Genetic Algorithm

Genetic algorithms have provided the optimized solutions for making decision and obtaining better results

The population size (an integer >= 1)

=> 3

The mutation rate (a decimal point '.', followed by a nonnegative integer, so that 0 <= mutation rate <= 1)

=> 2.4

The crossover rate (a decimal point '.', followed by a nonnegative integer, so that 0 <= crossover rate <= 1)

=> 1.3

The number of crossover points (either '1' or '2')

=> 2

The maximum number of generations (an integer >= 1)

=> 3

The number of times to run the GA with these parameters (an integer >= 1)

=> 3

For Run 1 and generation count 3, Population's initial best fitness values are 57, or 38.0 percent fit and

Population's final best fitness values are 51, or 34.0 percent fit. Its Average fitness delta is given by -1.333. Similarly for Run 2 and generation count 3, Population's initial best fitness value is evaluated as 53, or 35.3333 percent fit and Population's final best fitness value is given by 51, or 34.0 percent fit. Its Average fitness delta is given by -0.444. For Run 3 and generation count 3, Population's initial best fitness is given by 71, or 47.3333 percent fit and Population's final best fitness is given by 51, or 34.0 percent fit. Its Average fitness delta: -4.444.

Result for most fit individual with different Runs is given in Table 6.

TABLE VI. Most Fit Individual

Run	Generation count	Most Fit individual Genotype values
1	3	-108, -29, 88, 85, 104, -80,
2	3	-110, -84, -66, 60, -70, 6,
3	3	23, -23, 113, 4, 81, 24,

IX. CONCLUSION AND FUTURE WORK

To work in the field of medical sciences many methods are available. The most important and challenging part with machine learning techniques is to provide accuracy and computational efficiency. The proposed research, prediction of cancer is identified using different algorithm of machine learning in comparison with genetic algorithm. Genetic Algorithm optimizes the neural network by involving connection weights. The Breast Cancer Wisconsin (Original) Data Set is taken from UCI machine learning repository and it is trained and tested by applying different algorithm like Random forest, Naïve Bayes, Artificial neural network and Genetic algorithm. Random forest is used as an ensemble learning algorithm which depends on the combination of several decision trees, it is applied on the n tree bootstrap_sample , For predictive modeling Naïve Bayes algorithm is used to analyze the data within each attribute and for the events that evaluate conditional probability and back propagation algorithm considers different neurons in hidden layer to analyze output by calculating the error using weight adjustment method which reduces error between required and actual output. From experimental analysis genetic algorithm provides topmost accuracy of 97% where as other algorithms like artificial neural network provides accuracy of 96%, Naïve Bayes provides accuracy of 94% and Random forest provides accuracy of 95%. This research work can further be enhanced by incorporating fuzzy in selection function to improve the accuracy.

REFERENCE

[1] A. Krause, "New genetic testing technologies: Advantages and limitations", South African Medical Journal, vol. 109, no. 4, p. 207, 2019.

[2] Ronagh, Mehrnaz, and Mohammad Eshghi. "Hybrid Genetic Algorithm and Particle Swarm Optimization Based Microwave Tomography for Breast Cancer Detection." In 2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE), pp. 244-248. IEEE, 2019.

[3] P. Chauhan, Pragya, and Amit Swami. "Breast Cancer Prediction Using Genetic Algorithm Based Ensemble Approach." In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT),IEEE, pp. 1-8, 2018.

[4] Y. Li, Y. Liu and L. Bai, "Genetic algorithm based feature selection for mass spectrometry data, 2008 8th IEEE International Conference on Bioinformatics and Bio Engineering, Athens", p. 1-6, 2008.

[5] Q. Chen, M. Huang, Q. Xu, H. Wang and J. Wang, "Reinforcement Learning-Based Genetic Algorithm in Optimizing Multidimensional Data Discretization Scheme, Mathematical Problems in Engineering" ,p.1-13,2020.

[6] P. Chauhan and A. Swami, "Breast Cancer Prediction Using Genetic Algorithm Based Ensemble Approach, 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bangalore",p. 1-8,2018.

[7] Genetics - Methods in genetics, Encyclopedia Britannica;2020. [Online].

[8] Chauhan, P., & Swami, A. Breast Cancer Prediction Using Genetic Algorithm Based Ensemble Approach, 9th International Conference on Computing, Communication and Networking Technologies, (ICCCNT), p.12-22, 2018

[9] Microarray-based Comparative Genomic Hybridization (aCGH) | Learn Science at Scitable", Nature.com;2020. https://www.nature.com/scitable/topicpage/microarray-based-comparative-genomic-hybridization-acgh-45432/

[10] J. Bridge, "Advantages and limitations of cytogenetic, molecular cytogenetic, and molecular diagnostic testing in mesenchymal neoplasms, Journal of Orthopaedic Science", vol. no.13; p. 273-282, 2008.

[11] Espejo, P. G., Ventura, S., & Herrera, F., "A survey on the application of genetic programming to classification, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)", vol. 40(2),p. 121-144,2009.

[12] Shapiro J., "Genetic Algorithms in Machine Learning", In: Paliouras G., Karkaletsis V., Spyropoulos C.D. (eds) Machine Learning and Its Applications. ACAI 1999. Lecture Notes in Computer Science", 2049, p. 111-121, 2009.

[13] Berlin, Heidelberg. Pereira, F., Amorim, A., & van Asch, B., "Genetic and DNA-Based Techniques. Comprehensive Analytical Chemistry", vol. no. 6, p. 195–220, 2013

[14] A. Durmaz, E. Karaca, U. Demkow, G. Toruner, J. Schoumans and O. Cogulu, "Evolution of Genetic Techniques: Past, Present, and Beyond, BioMed Research International" , p. 1-7,2015

[15] P. Thomson and J. F. Miller, "Optimisation techniques based on the use of genetic algorithms (GAs) for logic implementation on FPGAs, IEEE Colloquium on Software Support and CAD Techniques for FPGAs, London, UK" , vol.no. 4, p. 4/1-4/, 1994.

[16] M. Chen and Z. Yao, “Classification Techniques of Neural Networks Using Improved Genetic Algorithms,2008 Second International Conference on Genetic and Evolutionary Computing, Hubei” , p. 115-119,2008
[17] S. Phiromlap and S. Rimcharoen, “A frequency-based updating strategy in Compact Genetic Algorithm, 2013 International Computer Science and Engineering Conference (ICSEC), Nakorn Pathom”, p. 207-211, 2013.
[18] V. Rajput and A. Manjhvar, “A model for forecasting dengue disease using genetic based weighted FP-growth, 2017 International Conference on Trends in Electronics and Informatics (ICEI), Tirunelveli” , p. 944-948, 2017.
[19] M. Ronagh and M. Eshghi, “Hybrid Genetic Algorithm and Particle Swarm Optimization Based Microwave Tomography for Breast Cancer Detection, 2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Malaysia” , p. 244-248, 2019
[20] R. N. Waykole and A. D. Thakare, “Intelligent Classification of Clinically Actionable Genetic Mutations Based on Clinical Evidences, 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India”, p. 1-4, 2018.
[21] Singhal, P., & Pareek, S., “Artificial Neural Network for Prediction of Breast Cancer, 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference On”, 2018.
[22] B. Dai, R. Chen, S. Zhu and W. Zhang, Using Random Forest Algorithm for Breast Cancer Diagnosis, 2018 International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan; 2018; p. 449-452.
[23] P. Singhal and S. Pareek, “Artificial Neural Network for Prediction of Breast Cancer, 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on, Palla dam, India”, p. 464-468, 2018.