Prediction of Cancer Behavior Based on Artificial Intelligence

Shayma M. Al-Ani, Maysam Abbod

Abstract**—** Cancer has been one of the most famous conditions discussed and researched about throughout the human history. Some of the earliest medical records regarding cancer are dated back to around 1600 BC. Cancer is a general condition which is subdivided into a group of conditions that are concerned with an abnormal growth in the cells within an organ or a tissue with the chance of spreading and invading other parts of the body. Nowadays, there is a growing number of cancer patients and with this increase arises the necessity for new techniques to accurately diagnose and predict cancer in its different forms and thus playing a huge part in improving the quality of life. Moreover, techniques that depend on the principle of intelligent systems and artificial neural networks are proven to be very efficient in the field of cancer research.

Keywords**—**ANN, Cancer prediction, Ensemble model.

I. Introduction

CANCER has been known all the way through human history. Other names for cancer are malignant tumor or malignant neoplasm [1]. Genetic heritage, tobacco and alcohol intake, obesity, radiation exposure as well as having a poor and inactive lifestyle are some causes for abnormal cell growth and thus providing higher risks of getting cancer. For such reasons, cancer is considered as one of the most dangerous and unpredictable diseases nowadays. Much of research is done on the prevention of cancer and cancer treatment [2]. Innovative and modern technology is being implemented in the goal of providing proper diagnosis, prediction and in some cases treatments for cancer. Artificial intelligence is one of the methods for approaching cancer and understanding its nature [3]. One of the most popular types of cancer being approached by artificial intelligence is breast cancer in females.

An Artificial Neural Network (ANN) is an imitation to the basic human brain operation and it is an interconnected neurons system that is capable of computing values using mathematical functions in which they determine the activation of the neuron[4][5]. To adapt to the environmental changes, a learning system has to change itself. In addition, Multi-layers ANNs are complex neural networks providing a nonlinear relationship of input-to-output results. Multi-layer ANNs comprise of an input layer, a hidden layer and an output layer as illustrated in Fig. 1. Basically, the input layer provides an input value to the network and each of the input cells has a weighting factor, which identifies the effect of the cell on the network[6]. As for the hidden and output cells, they represent a function where the hidden layer is first computed, and then the results are used in computing the output layer.

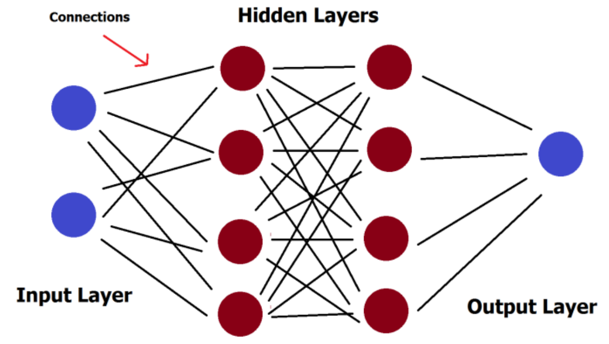

Fig. 1 Artificial neural network model.

The data presented in this paper is patients suffering from bladder cancer . This database is obtained for xx patient, each patient is represented with different information as input data such as the type of tumor, patient details (sex, age, tobacco consumption) in addition to protein expression (p53, msh2, mlh1) and DNA mutations (bat25, bat26, mfd15, apc, d2s123). Furthermore, the output data will be represented with the actual behavior of the tumor such as how long did and how many times the tumor went back to appear. In addition the time it took to advance to other stages, the time the cancer spread and lead to patient's death and whether or not the cancer was the cause of death or there might be other causes of death (e.g. complications).

II. Methodology

The proposed ANN based prediction algorithm accurately predicts the patient's cancer records output by employing the ensemble method shown in Table I. In In this method, the patient's record is equally divided into 10 window groups. In

S Alani and M Abbod are with the Department of Electronic and Computer Engineering, Brunel University, Uxbridge, UB8 3PH, UK, (email: Shayma.Al-Ani@brunel.ac.uk, maysam.abbod@brunel.ac.uk)

this method, the average of 10 ANN network functions under different combinations of 10 window groups have been used in order to find out the predicted output, in addition to improving the prediction performance of a model with more accurate results.

The proposed model is trained by using three different ANN networks which are cascade-forward back propagation network (NEWCF), feed-forward input time-delay back propagation network (NEWFFTD), and fitting network (NEWFIT), each network is trained by using ensemble methods under different combination of groups by applying two methods which are the averaging and voting methods.

The averaging method is one of the major types of static committee machines. The network design for such method depends upon mean average of the networks. In addition, ensemble averaging depends on the mean average networks results. So in general, the whole idea of averaging method can be summarized by the following; generating N experts each having their initial values which are chosen from a random distribution. After that, each expert will be separately trained separately and finally, they are combined and their values are averaged.

As for the voting method, it does not consider the level of significance by each network. This as a result, allows simple integration of all different sorts of network architectures. Majority voting is a simple voting method in which a group of unlabeled instance are performed depending on the class with the most frequent votes. This technique has been widely used to compare newly proposed methods.

TABLE I. SLIDING WINDOW METHOD.

W1	W2	W3	W4	W5	W6	W7	W8	W9	W10
W2	W3	W4	W5	W6	W7	W8	W9	W10	W1
W3	W4	W5	W6	W7	W8	W9	W10	W1	W2
W4	W5	W6	W7	W8	W9	W10	W1	W2	W3
W5	W6	W7	W8	W9	W10	W1	W2	W3	W4
W6	W7	W8	W9	W10	W1	W2	W3	W4	W5
W7	W8	W9	W10	W1	W2	W3	W4	W5	W6
W8	W9	W10	W1	W2	W3	W4	W5	W6	W7
W9	W10	W1	W2	W3	W4	W5	W6	W7	W8
W10	W1	W2	W3	W4	W5	W6	W7	W8	W9

In the case of Artificial Neural Network model, the neuron behaves as an activation function *f*(.) producing an output $y = f(net)$, where net is the cumulative input stimuli to the neuron and *f* is typically a nonlinear function of net, where x_i indicates the inputs and w_i indicate the weighting parameters.

$$net = x_1w_1 + x_2w_2 + x_3w_3 = \sum_{i=1}^{3} x_iw_i \quad (1)$$

Output performances of the proposed algorithms are analysed using various parameters such as Sensitivity, Specificity, Accuracy, Receiver Operator Curve (ROC), Area Under the Curve (AUC) and Mean Square Error (MSE) value.

Regression model is a statistical model for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables and statistical models to compare it with the ANN model, output performances of them are analysed using various parameters like Sensitivity, Specificity, Accuracy, Roc, AUC and MSE value.

III. Findings

The proposed ANN models are trained using three different ANN networks, namely NEWCF, NEWFFTD and NEWFIT[7][8]. First of all, randomly dividing the networks into two groups called training records and testing records[9][10][11]. Training records group contains about 70% of the total records, which are used to train the ANN by using 80% for training and 20% for validation of the ANN networks. The trained ANN networks are used to predict the output parameter of testing records group which contain 30% of total records .

Moreover, three different methods have been used, average, voting, and regression model[12][13][14]. Table II shows the input variables used in the modeling analyses. Tables III and IV show the performance of three methods for various ANN training networks in which it follows the principle of 70% training and 30% testing, while Tables V and VI show the predicted patients records for three methods and three different ANN trained networks using ensemble method.

Sensitivity relates to the test's ability to identify positive results which measures the proportion of actual positives which are correctly identified as such.

While specificity relates to the test's ability to identify negative results**,** which measures the proportion of negatives which are correctly identified.

The accuracy is the proportion of true results (both true positive and true negative) in the population.

Sensitivity=TP/(TP+FN)

Specificity=TN/(TN+FP)

Accuracy = (TP+TN)/ (TP+FP+TN+FN

TABLE II. INPUT VARIABLES USED IN THE MODELING ANALUSES.

Input Variables
1. age
2. sex
3. chest pain type (4 values)
4. resting blood pressure
5. serum cholesterol in mg/dl
6. fasting blood sugar > 120 mg/dl
7. resting electrocardiographic results (values 0,1,2)
8. maximum heart rate achieved
9. exercise induced angina
10. oldpeak = ST depression induced by exercise relative to rest
11. the slope of the peak exercise ST segment **12**. number of major vessels (0-3) colored by flourosopy
13. thal: 3 = normal; 6 = fixed defect; 7 = reversible defect

TABLE III. PERFORMANCE OF ANN NETWORKS TRAIN RECORDS RESULTS ANALYSIS.

Methods	**Sensitivity**	**Specificity**	**Accuracy**
Average	78.1818	86.25	82.963
Voting	74.5455	88.75	82.963
Regression Model	63.6364	73.75	69.630

TABLE IV. PERFORMANCE OF ANN NETWORKS TRAIN RECORDS RESULTS ANALYSIS MSE AND AUC VALUES.

Methods	**MSE Value**	**AUC**
Average	0.1378	0.9009
Voting	0.1330	0.9048
Regression Model	0.1995	0.7398

TABLE V. PERFORMANCE OF ANN NETWORKS TEST RECORDS RESULTS ANALYSIS.

Methods	**Sensitivity**	**Specificity**	**Accuracy**
Average	65.2174	74.2857	70.6897
Voting	60.8696	77.1429	70.6897
Regression Model	52.609	50	51.034

TABLE VI. PERFORMANCE OF ANN NETWORKS TEST RECORDS RESULTS ANALYSIS MSE AND AUC VALUES.

Methods	**MSE Value**	**AUC**
Average	0.1908	0.7280
Voting	0.1956	0.7193
Regression Model	0.1335	0.5745

Figs. 2-5 show the ROC plot of bladder cancer train records of average method and regression model for NEWCF, NEWFFTD and NEWCF networks, while Figs. 6-9 show the ROC plot of bladder cancer test records of average method and regression model.

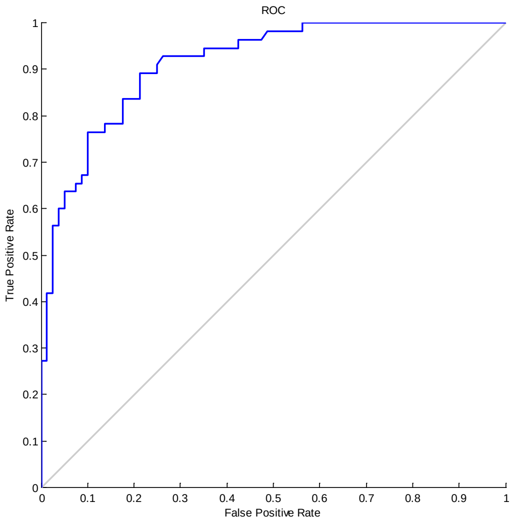

Fig.2 Average method train case.

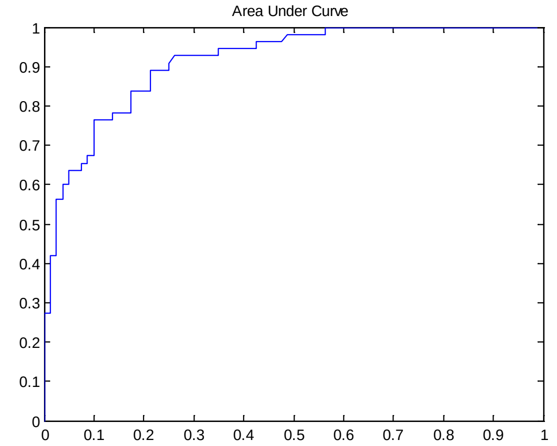

Fig.3. Average method train case.

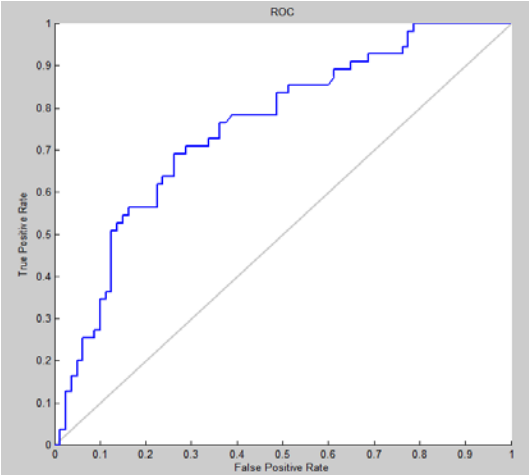

Fig.4 Regression model train case.

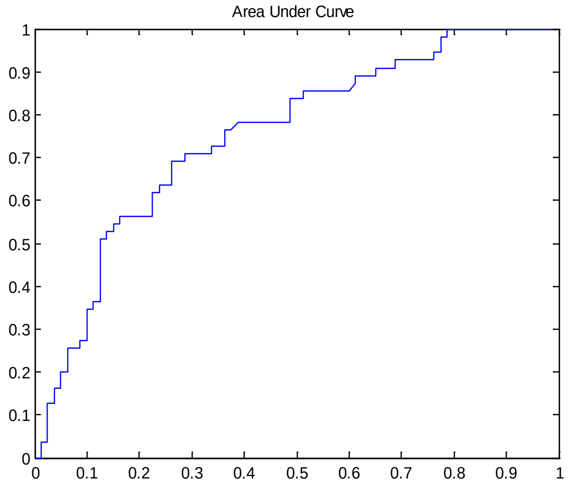

Fig.5 Regression model train case.

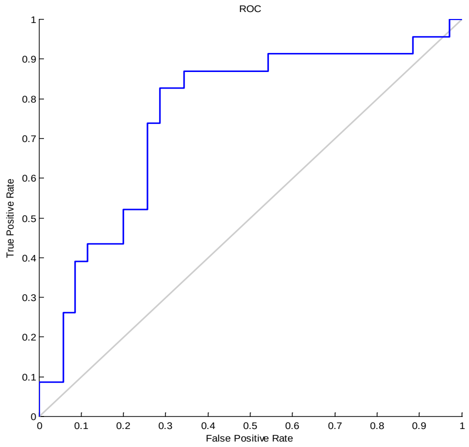

Fig.6. Average method test case.

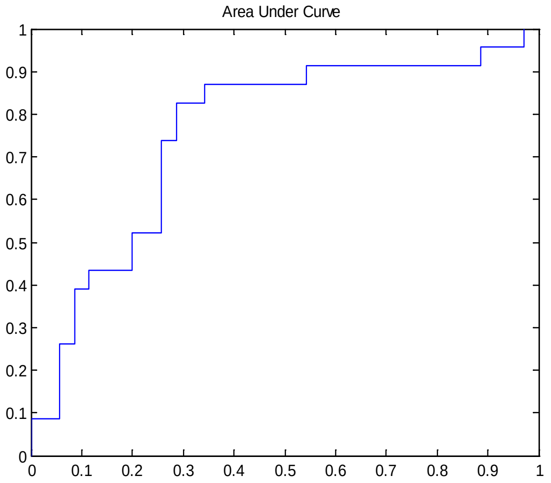

Fig.7 Average method test case.

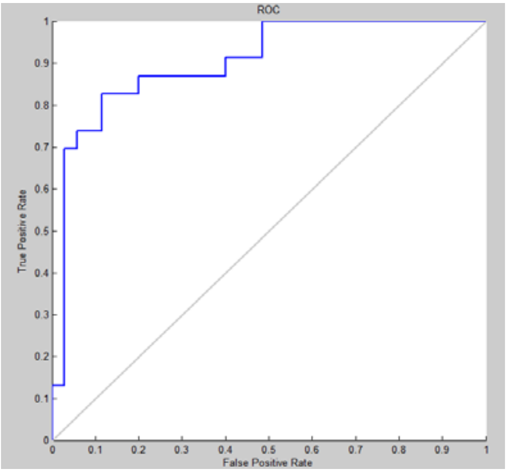

Fig.8 Regression model test case.

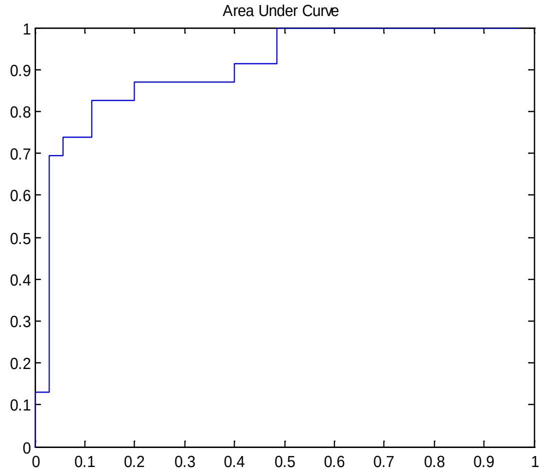

Fig.9 Regression model test case.

IV. Conclusions

The proposed ensemble model, the artificial neural network algorithm using two methods averaging, voting and for various artificial neural network functions such as feed-forward input time-delay back-propagation network, cascade-forward back-propagation network and radial basis network, successfully and accurately predicted patient's records. Output performances of records are analyzed using various parameters such as Sensitivity, Specificity, Accuracy, ROC, AUC and MSE value and the results show that artificial neural network methods obtain better predictive performance than could be obtained from regression models and that was all based on the different validations of the artificial neural networks.

References

[1] (2014). "What is Cancer". *National Cancer Institute*. [Online]. Available: http://www.cancer.gov/cancertopics/cancerlibrary/what-is-cancer

[2] (2014). "Definition of Bladder Cancer". *National Cancer Institute*. [Online]. Available http://www.cancer.gov/cancertopics/types/bladder.

[3] J.A. Cruz and D. S. Wishart. (Feb, 2007). Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Informatics*. [Online]. 2(2006), 59-77. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2675494/.

[4] Z. Chi, Z. Lu and F. Chan, “Multi-channel handwritten digit recognition using neural networks”, Circuit and Systems ISCAS’97 proceeding of 1997 IEEE International Symposium, Vol.1, 1997, pp. 625-628.

[5] K. Tadashi; U. Junji; T. Shoichiro, “Hybrid GMDH-type neural network using artificial intelligence and its application to medical image diagnosis of liver cancer”, System Integration (SII), 2011 IEEE/SICE International Symposium on 2011, pp. 1101-1106.

[6] P.J.G. Lisboa, T.A. Etchells, I.H. Jarman and M.S.H. Aung, “Time-to-event analysis with artificial neural networks: An integrated analytical and rule-based study for breast cancer”, Neural Networks, 2007. IJCNN 2007. International Joint Conference on 2007, pp: 2533-2538.

[7] P. Melville, “Creating Diverse Ensemble Classifiers,” PhD proposal, Texas Univ., Texas, United States, 2003.

[8] Robi Polikar (2009) Ensemble learning. Scholarpedia, 4(1): 2776.

[9] Opitz, D.; Maclin, R. (1999). "Popular ensemble methods: An empirical study".*Journal of Artificial Intelligence Research* **11**: 169–198.

[10] U, Naftaly, N. Intrator, and D. Horn. "Optimal ensemble averaging of neural networks." Network: Computation in Neural Systems 8, no. 3 (1997): 283–296.

[11] L. Rokach. (Nov, 2009). Ensemble-based classifiers. *Springer Science+Business Media*. [Online]. 33 (2010), 1-39. Available: http://www.ise.bgu.ac.il/faculty/liorr/AI.pdf .

[12] (2014). “Regression analysis”. Wikipedia. [Online]. Available http://en.wikipedia.org/wiki/Regression_analysis#Regression_models.

[13] C.Chen, C. Hsu, H.Chiu and H.Rau, “Prediction of survival in patients with livercancer using artificial neural networks and classification and regression trees”, Natural Computation (ICNC), 2011 Seventh International Conference on 2011, pp. 811-815.

[14] Janghel, R. R.; Shukla, Anupam; Tiwari, Ritu; Kala, Rahul, 2010. Breast cancer diagnosis using Artificial Neural Network models. Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on 2010, pp. 89-94.