ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

Predicting Tumor Category Using Artificial Neural Networks

Ibrahim M. Nasser, Samy S. Abu-Naser
Department of & Information Technology ,
Faculty of Engineering & Information Technology,
Al-Azhar University - Gaza, Palestine
azhar.ibrahimn@gmail.com

Abstract: *In this paper an Artificial Neural Network (ANN) model, for predicting the category of a tumor was developed and tested.*

Taking patients' tests, a number of information gained that influence the classification of the tumor. Such information as age, sex, histologic-type, degree-of-diffe, status of bone, bone-marrow, lung, pleura, peritoneum, liver, brain, skin, neck, supraclavicular, axillar, mediastinum, and abdominal. They were used as input variables for the ANN model. A model based on the Multilayer Perceptron Topology was established and trained using data set which its title is "primary tumor" and was obtained from the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia

Test data evaluation shows that the ANN model is able to correctly predict the tumor category with 76.67 % accuracy.

Keywords: Artificial Neural Networks, Tumor category, Cancer, ANN, Medicine, Predictive Model.

1. INTRODUCTION

Obstructive sleep apnea (OSA) (or apnea) is the common method ofsleep apneaand is caused by complete or partial collapse of the pharyngeal upper airway tract.

The main objective of this study is to determine tumor category for patents based on attributes which are set of tests for the patient body.

Specifically the study seeks to explore the possibility of using an Artificial Neural Network model to predict the category of a tumor.

The category of a tumor maybe certain type of function with a number of factors. However, it seems that it will be difficult to find a mathematical model that effectively models these factors relationship.

A useful approach to deal with this type of problem is to apply common regression analysis in which historical data are the best fitted to some function. The result is an equation in which each of the inputs xj is multiplied by a weight wj; the sum of all such products and a constant θ, gives an output $y = \Sigma\, wj\, xj + \theta$, where j=0..n.

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

Such studies face problems with the complexity of selecting an appropriate function fit to capture all forms of data associations as well as automatically adjusts output in case of additional information, because of the performance of a candidate is controlled by a number of factors, and this control is not going to be any straightforward well-known regression model.

Artificial neural network emulates humans' brain in solving problems; it is a common approach that can tackle that kind

of problems. Therefore, the attempt to build an adaptive system such as Artificial Neural Network to predict a tumor's category based on the consequence of such factors.

The objectives of this study are:

- To identify some suitable factors that affects tumor classification,
- To convert these factors into forms appropriate for an adaptive system coding, and
- To model an Artificial Neural Network that can be used to predict the tumor category based on some predetermined data for a given patient.

2. THE ARTIFICIAL NEURAL NETWORKS

An Artificial Neural Network (ANN) is a branch of Artifical Intelligence [8-28]. It is a mathematical model that is encouraged by the organization and/or functional feature of biological neural networks. A neural network has a connected set of artificial neurons, and it processes information using a connectionist form to computation. Generally, an ANN is an adaptive system that fine-tunes its organization based on external or internal information that runs through the network during the learning process.

Latest neural networks are non-linear numerical data modeling tools. They usually used to model sophisticated relationships among inputs and outputs or to uncover patterns in data. ANN has been applied in various applications with considerable fulfillment [1-2]. For example, ANN has been applied effectively in the area of prediction, handwritten character recognition, evaluating prices of housing [3-4].

Neurons often grouped into layers. Layers are groups of neurons that implement similar tasks. There are three types of layers. The input layer is the layer of neurons that receive input from the user program. The output layer is the layer of neurons that send data to the user program. And Between of them there are hidden layers. The Hidden layer neurons are connected only to other neurons and never directly interact with the user program. Every neuron in a neural network has the opportunity to affect processing which can occur at any layer in

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

the neural network. In neural networks, the hidden layers are optional. The input and output layers are essential, however it is possible to have on layer that act as an input and output layer [4].

ANN learning can be directed or undirected. Directed training means giving the neural network a set of sample data alongside the predicted outputs from each of these samples. Directed training is the most common form of neural network training. As directed training continues, the neural network goes through several iterations, or epochs, until the actual output of the neural network equals the predicted output, with a reasonably small error rate. Each iteration is one pass through the training samples. Undirected training is similar to the directed one but no predicted outputs are provided. Undirected training usually occurs when the neural network tends to classify the inputs into several groups. The training progresses through many epochs, just as in directed training. As training progresses, the neural network discovers the classification groups [3].

Training is the process by which these connection weights are assigned. Most training algorithms begin by assigning random numbers to the weight matrix. Then the validity of the neural network is inspected. Next, the weights are tuned based on how valid the neural network done. This process is repetitive until the validation error is within an acceptable limit [2].

Validation of the system is done once a neural network has been trained and it must be assessed to tell if it is ready for actual use. This final step is important so that it can be determined if additional training is required. To correctly validate a neural network, validation data records must be completely separated from the training data records [4].

About 80% of the total sample data was used for network training in this paper. About 20% of the total sample data used for validation of the system.

3. METHODOLOGY

A data set refer to Igor Kononenko, and Bojan Cestnik [5] was used, it contains a number of factors that are considered to have an effect on the classification of a tumor. These factors were carefully studied and synchronized into a convenient number appropriate for computer coding within the environment of the ANN modeling. These factors were classified as input variables. The output variables represent the predicted tumor classification based on those inputs.

3.1. The Input Variables

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

Table 1: Input Data Transformation

/N	Input	Domain	Transformed domain
.	age	<30, 30-59, >=60	1, 2, 3
.	sex	male, female	0,1
.	histologic-type	epidermoid, adefalse, anaplastic	1,2,3
.	degree-of-DIFFE	well, fairly, poorly	1,2,3
.	bone	yes, no	1,0
.	bone-marrow	yes, no	1,0
.	lung	yes, no	1,0
.	pleura	yes, no	1,0
.	peritoneum	yes, no	1,0
0.	liver	yes, no	1,0
1.	brain	yes, no	1,0
2.	skin	yes, no	1,0
3.	neck	yes, no	1,0
4.	supraclavicular	yes, no	1,0
5.	axilla	yes, no	1,0
6.	mediastinum	yes, no	1,0

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

7.	abdominal	yes, no	1,0

These factors were converted into a format suitable for neural network analysis as shown in Table1.

3.2. The Output Variable

The output variable is the Tumor Class, and its domain is:

Lung, Head & neck, Esophagus, Thyroid, Stomach, Duodenum & sm.int, Colon, Rectum, Anus, Salivary glands, Pancreas, Gallbladder, Liver, Kidney, Bladder, Testis, Prostate, Ovary, Corpus uteri, Cervix uteri, Vagina, Breast

4. THE NEURAL NETWORK

4.1. Network Architecture

Humans and other animals process information with neural networks. These are formed from trillions of neurons (nerve cells) exchanging brief electrical pulses called action potentials. Computer algorithms that mimic these biological structures are formally called artificial neural networks to distinguish them from the squishy things inside of animals. However, most scientists and engineers are not this formal and use the term neural network to include both biological and nonbiological systems [29-49].

Neural network research is motivated by two desires: to obtain a better understanding of the human brain, and to develop computers that can deal with abstract and poorly defined problems. For example, conventional computers have trouble understanding speech and recognizing people’s faces. In comparison, humans do extremely well at these tasks [50-68].

The network is a multilayer perceptron neural network using the linear sigmoid activation function as seen in Figure 1.

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

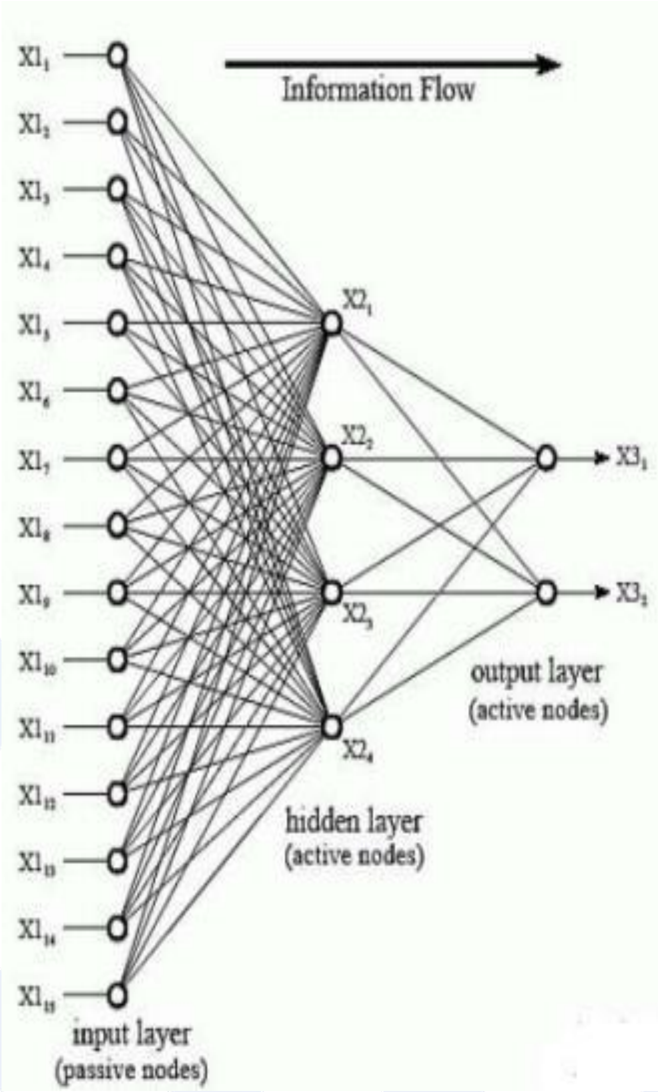

Figure 1: Artificial Neural System Architecture

4.2. The Back-propagation Training Algorithm

Algorithm 1 The basic backpropagation algorithm
[11]

1: Initialize weights randomly

2: Initialize *err*, *threshold*, and *maxEpochs*

3: **while** *epoch* < *maxEpoch* **and** *err* > *threshold* **do**

4: **for** each example (*x*,*y*) in the training set **do**

5: /* *Propagate the inputsforward to compute the outputs* */

6: **for** each node *i*in the input layer **do**

7: $a_i \leftarrow x_i$

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

```
    end for
    for £ = 2 to L do
    for each node j in layer £ do
  inj ← Σ i wi,j ai
  aj ← g(inj)
    end for
    end for
    /* Propagate deltas backward from output layer to input layer */
    for each node j in the output layer do
    Δ[j] ← g′(inj) × (yj − aj)
    end for
    for £ = L − 1 to 1 do
    for each node i in layer £ do
  Δ[i] ← g′(inj) Σj wi,j Δ[j]
    end for
    end for
    /* Update each weight using deltas */
    for each weight wi,j do
    wi,j ← wi,j + α × ai × Δ[j]
    end for
    end for
end while
```

4.3. The Design of The Neural Network

The ANN model consists of three layers: 1 input layer, 1 hidden layer, and 1 output layer (as seen in fig.

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

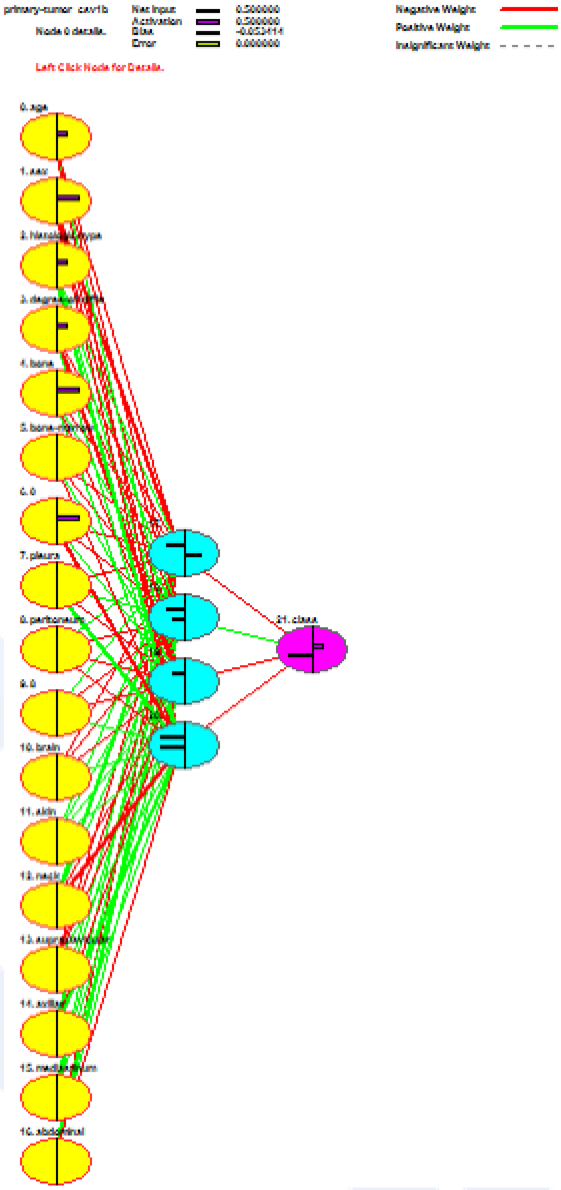

Figure 2: Shows the Design of the Neural Networks

4.4. Training and Error Rates

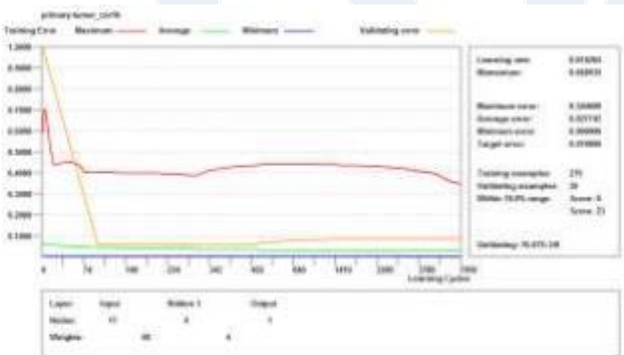

Figure 3: Shows the Training, error, and validation of the data set.

TECHNICAL SCIENCE
RESEARCH IN UZBEKISTAN
ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

Figure 4: Shows the relative importance of the input attributes.

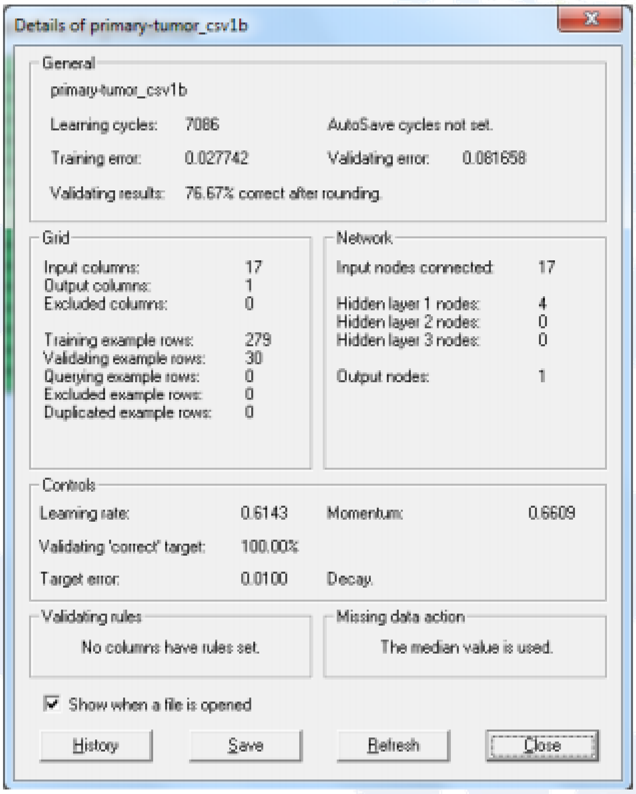

Figure 5: Shows the detail of the ANN model.

5. EVALUATION OF NEURAL NETWORK

As said, the purpose of this study was to predict the tumor type. Where we used patients test results, which provides the possibility to implement and test the neural network and itslearning algorithm. Our neural network is designed to classify the tumor based on those test results.

After training and validation, the network was tested using test records and the following results were obtained. This involves inputting variable input data into the grid without output variable results. The output from the grid is then compared with the actual variable data.

The neural network was successfully able to accurately classify 76.67 % of the data

HTTP://UNIVERSALPUBLISHINGS.COM

ISSN (E): 2992-9148 SJIF 2024 = 5.333
ResearchBib Impact Factor: 9.576 / 2024
VOLUME-2, ISSUE-3

REFERENCES

[1] A. Lotfi and A. Benyettou, "Using Probabilistic Neural Networks for Handwritten Digit Recognition", Journal of Artificial Intelligence, vol. 4,no. 4, (2011).

[2] P. Khanale and S. Chitnis, "Handwritten Devanagari Character Recognition using Artificial Neural Network", Journal of Artificial Intelligence, vol. 4, no. 1, (2011).

[3] P. Eriki and R. Udegbunam, "Application of neural network in evaluating prices of housing units in Nigeria: Apreliminary investigation", J. of Artificial Intelligence, vol. 3,no. 1, (2010).

[4] H. Martin and D. Howard, "Neural Network Design", 2nd Edition, Martin Hagan (2014).

[5] Dua, D. and Karra Taniskidou, E. (2017). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[6] Al-Shawwa, M., Al-Absi, A., Abu Hassanein, S., Abu Baraka, K., & Abu-Naser, S. S. (2018). Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network. International Journal of Academic Pedagogical Research (IJAPR), 2(9), 1-6.

[7] Al-Shawwa, M., & Abu-Naser, S. S. (2019). Predicting Birth Weight Using Artificial Neural Network. International Journal of Academic Health and Medical Research (IJAHMR), 3(1), 9-14.

[8] Abu-Naser, S., Al-Masri, A., Sultan, Y. A., & Zaqout, I. (2011). A prototype decision support system for optimizing the effectiveness of elearning in educational institutions. International Journal of Data Mining & Knowledge Management Process (IJDKP), 1, 1-13.

[9] Elzamly, A., Abu Naser, S. S., Hussin, B., & Doheir, M. (2015). Predicting Software Analysis Process Risks Using Linear Stepwise Discriminant Analysis: Statistical Methods. Int. J. Adv. Inf. Sci. Technol, 38(38), 108-115.