Detecting Blood Clots using Neural Networks

Leon Sanders\(^1\) and Y. B. Reddy\(^9\)

\( ^1\) Dept of Math and Computer Science, Grambling State University, Grambling, LA 71245, ybreddy@gram.edu
\( ^9\) Department of Biology, University of Pittsburgh, PA 15213, USA, Lss8@pitt.edu

Abstract

The purpose of this research is to determine a noninvasive user-friendly method that the general population can use to determine whether they may have a blood clot. One thing that has impeded detection of blood clots in the past is the threat of invasive surgery as well as the cost of the tests that are offered. Our goal is to create a software system that takes blood clot symptoms as user input and provides the type of blood clot disease they are suffering from so that the patient can go to the hospital for further diagnosis. Many times people feel pain and they do not know the cause nor do they see a doctor to determine the problem. By allowing people to check themselves first in the privacy of their own homes it is possible that more people will see doctors with better knowledge of their potential problem. To accomplish this goal, we used a neural network model to determine the disease based on symptom inputs of the patient. The `nftool` of MATLAB with appropriate training provides an accuracy of 99.99%.

Key Words

Blood clots, arterial embolism, stroke, thrombophlebitis, and neural networks

1. Introduction

Blood clots form as a result of the coagulation of blood. A blood clot that forms in a vessel or the heart and stays in the location it formed is called a thrombus. A blood clot that leaves the area of its formation is known as an embolus/embolism (disorder). Thrombi or emboli may attach to a blood vessel and partially or completely block the flow of blood in that vessel. Blockage that prevents normal blood flow and proper intake of oxygen is referred to as ischemia. If not treated quickly, damage or death (infarction or necrosis) of the tissues in that area occurs [25-27]. One of the most common forms of blood clot disorders is deep venous thrombosis. It is speculated that there are approximately 3 million cases per year in the United States [25-27]. While not necessarily a reliable number, this still indicates that blood clot disorders occur in a very large portion of the US population each year.

There are a number of disorders that may result from thrombi and can occur in various parts of the body (see Appendix A). There are a number of risk factors that increase the likelihood of developing one of these disorders as a result of a blood clot: prolonged sitting, disorders related to blood clotting, inactivity due to recent injury or surgery, pregnancy or recent childbirth, oral contraceptive and estrogen therapy, cancer, stroke or other diseases that limit movement, family history of clotting disorder, obesity, smoking, and central venous catheters (used for injection of medications or for imaging). As there are a number of different disorders that are caused by blood clots, there are also a range of different symptoms that correspond to each disease.

There are a diverse range of diagnostic techniques available to physicians when a blood clot is suspected. These include but are not limited to duplex ultrasound, magnetic resonance imaging, venography, computed tomography scans, magnetic resonance angiography, D-Dimer test, arteriography/angiography, and impedance plethysmography. These techniques differ in the method they employ to determine the existence of a blood clot, and are also designed specifically to detect different disorders.

Early detection of blood clots was studied using neural network models [8 -11, 21-23], statistical models [6, 12, 13], genetic algorithms [7], and other machine learning techniques including decision tree, fuzzy sets, and evolutionary algorithms [1-6, 14-16, 24]. The models used on practical study of patients [5], for the early diagnosis of myocardial infarction [1-2], chest pain, heart attack [3, 6, 7], develop the statistical models on available data [4], and on available practical issues associated with the realization of a cardiac functions. Most of the information on Web available are the physical symptoms for blood clots. But none of these symptoms are used to make a tool useful for the average citizen. For every suspected case, doctors run a number of tests and spend enormous amounts of money before they ascertain the real cause. We feel a simple device developed as an outcome of the research to predict the symptoms of blood clot will help many people to avoid the deaths happening today [25-27]. The current research is based on developing an inexpensive tool for the average citizen to detect blood clots.

In this paper, we used the neural network model to predict blood clots with the help of physical symptoms provided by the patient. This is an experimental model. Once the neural network model is trained, it will predict to the accuracy of 99% depending upon the training data provided (accuracy of training data). Section 2 provides the overview of neural networks and `nftool` of MATLAB, section 3 discusses methodology, and is followed by discussion of results and conclusions.

2. Neural Network Overview

In this study, neural networks were used to determine the feasibility of diagnosing blood clots using symptoms as the data. For this experiment, stroke and arterial embolism were
studied as the number of symptoms made the training of the network easier.

Neural networks are inspired by nervous systems found in biological organisms. It is comprised of data processing units (neurons) connected via adjustable connection weights. Neurons are arranged in layers, an input layer, hidden layer(s), and an output layer. There is no specific rule that dictates the number of hidden layers. The function is largely established based on the connections between elements of the network. In the input layer each neuron is designated to one of the input parameters. The network learns by applying a back-propagation algorithm which compares the neural network simulated output values to the actual values and calculates a prediction error. The error is then back-propagated through the network and weights are adjusted as the network attempts to decrease the prediction error by optimizing the weights that contribute most to the error. The training or learning of the network occurs through the introduction of cycles of data patterns (epochs or iterations) to the network. One problem with neural network training is the tendency for the network to memorize the training data after an extended learning phase. If the network over learns the training data it is more difficult for the network to generalize to a data set that was not seen by the network during training. Therefore, it is common practice to divide the data set into a learning data set that is used to train the network and a validation data set that is used to test network performance. Fig 1 shows the representation of neural network diagram with inputs, weights, hidden layer, and an output [17 - 20].

Fig. 1 feed-forward neural network model

In the present study, we used neural network fitting tool (nftool) of MATLAB. Using ‘nftool’ it is possible to select data, create a network, train it and evaluate its performance through the use of mean square error and regression analysis. If the first network does not perform optimally, it is then possible to use the tool to make adjustments to the sample size, retrain, or increase the network size. This tool is most suitable for static fitting problems utilizing a standard two-layer-feed-forward neural network trained with Levenberg-Marquardt [17-19]. Fig 2 shows the input, hidden layer(s), output layer(s) and output of a ‘nftool’. The tool has a provision of selecting the number of hidden neurons, which leads to the accuracy and fast convergence of the problem.

Fig. 2 MATLAB Neural Network Model (nftool)

3. Methodology

An artificial neural network (ANN) system does not suffer from fatigue or psychological factors that can affect the reliability of the diagnosis procedure. The ANN system, once trained, can offer the expertise of an expert radiologist in interpreting the scans when an expert radiologist is not accessible. An ANN system has the promise for a more accurate diagnosis than is possible with human interpretation. In the clinical laboratory, backpropagation neural network have been applied successfully to multiple-predictor clinical tests to determine a diagnosis [21-23]. The ANN learns the prescribing behavior of expert physicians or alternatively learns the outcomes associated with such decisions. Both these possibilities are evaluated and we show how, through the prediction of outcomes that the prescribing practice of physicians may be improved. The ANN can learn to predict the outcome resulting from prescribing levels better than expert physicians. In evaluating the use of ANNs on these regression problems, the possibility of ensemble of these nets to improve performance is also considered.

The program model selected was backpropagation algorithm [17-20], a systematic method for training multilayer ANN. It has a mathematical foundation that is strong if not highly practical. Despite its limitations, backpropagation has dramatically expanded the range of problems to which ANN can be applied, and it has generated many successful demonstrations of its power. In the present research, backpropagation model was developed in MATLAB language with hidden layers, reset weights, learning rate, and error tolerance for termination of learning status. The sample data prepared [25-31] was provided to this model and trained the model. Depending upon the accuracy of data, number of training samples, and error for termination of training, the test results were above 80% to 95%. Sample results were shown in Table 1 for the case ‘Arterial Embolism’. The data used was based on the symptoms studied by many expert doctors [25-31]. The presence of symptom is 1 and absence is 0. For example, let us collect the data for Arterial Embolism. If some one feels numbness, pale color on arm or leg, body feels cold, weakness, and
The system was trained to 99.99% with validation 98.59%. The testing shows that diagnosis was close to 99.38% which is satisfactory. The mean squared error values indicate the average squared difference between (normalized) outputs and targets. Zero means no error, over 0.6667 means high error. Fig 6 displays the plots of the regression values for testing, validation and training. Regression R values measure the correlation between (unnormalized) outputs and targets. An R value of 1 means a close relationship, while 0 indicates a random relationship. The points lie directly on the best fit line demonstrating a strong percent correlation of the data.

Fig 7 and Fig 8 devoted to the stroke sample data. Since we found that ‘nftool’ was able to detect the disease 99.99% correctly, we recommend ‘nftool’ for the best results. Fig 7 depicts the second training routine of the stroke sample data and displays the same high percentage of detection found in the arterial embolism example. Again the network tool correctly detects the disease to 99.99% correct indicating the versatility of the tool for this kind of study. In Fig 7, the training routine was completed in 13 iterations.

In the second case, the neural network model ‘nftool’ of MATLAB was used to train the data (Fig. 2). To use ‘nftool’, first convert the data into MATLAB data file and provide as input to ‘nftool’. The system takes 60% of data for training, 20% for validation and 20% for testing, so that the same data will not use for testing. It requires many runs to converge or to get expected training. Once the system was trained then it tests with remaining symptoms data for testing. The results were discussed more in the section 4.

4. Discussion of Results

In this study, we used ‘nftool’ and neural network program written in MatLab language to determine the best possible test results. We found that ‘nftool’ is easy to use and is faster in producing diagnoses that are 99.99% correct. We provided two example tests in this study (Arterial Embolism and Stroke). We provided 214 samples with 10 symptoms for Arterial Embolism and 150 samples with 12 symptoms for stroke. The ‘nftool’ takes 60% of samples for training, 20% for validation, and 20% for testing. The test results are satisfactory and shown in Fig 4.

We studied different cases using the neural network program. The program converges after 4,000 iterations as shown in Fig 3. The present study includes 5,000 and 6,000 iterations as shown in Table 1 and Fig 3. The results indicate that increasing the iterations did not appreciably change the percent correlation. In this test, 5,000 iterations proved optimum for achieving a percent correlation above 80% as indicated in Table 1. In Fig 3, the number of iterations is indicated for the 214 sample training data set versus the error rate. We conducted several experiments and provided one sample case study each for Arterial Embolism and Stroke.

Fig 4 concludes that the system converges much faster (after 14 epochs). Fig 5 depicts the testing and training results.
5. Conclusion

In this study, artificial neural networks were used to develop a system in which patients would be able to self-diagnose themselves for blood clots before undergoing a more thorough examination. The ‘nftool’ and neural network program developed in MatLab language were used to diagnose the symptoms. The ‘nftool’ proved to be the most effective due to its higher percent correlation. The program was only able to achieve 95% correlation. While a high percentage, the inconsistency of the neural network program remains a problem. One possible reason for the percent correlations observed from the program is over training the network can affect the results as the network is unable to properly generalize to a new data set. The results indicates the success of this study, however a tool must be developed to deliver the program to patients using these results as an experimental study. This tool makes it possible for the patient to save money by indicating to the doctor what the disease most likely is and allowing the doctor to choose the best diagnostic tool for the job instead of doing numerous tests that may not uncover the problem. In this study, the diseases of stroke and arterial embolism were selected due to the high number of symptoms and their effect on the performance of the neural network. Follow up studies examining other diseases caused by blood clots is desired and shall be undertaken.

Acknowledgement: The authors are thankful to Dr. Connie Walton-Clement, Dean College of Arts and Sciences for the continuous support.

References


Appendix-A
The symptoms for the disorders

Symptoms of Arterial Embolism [1]
- Muscle pain the extremity
- Numbness and tingling in the extremity
- Pale color of arm or leg
- Decreased or absent pulse in the extremity
- Decreased extremity temperature; extremity feels cold to the touch
- Lack of movement of the extremity
- Weakness of the extremity
- Muscle spasm in the extremity
- Body feels cold (fingers or hands)
- Muscle function loss

Symptoms of Pulmonary Embolism [19]
- Sudden feeling of apprehension
- Shortness of breath
- Sharp chest pain
- Rapid pulse
- Sweating
- Cough with bloody sputum
- Fainting

Symptoms of Thrombophlebitis [2,10]
- Pain
- Sudden swelling in the affected limb
- Enlargement of the superficial(close to the skin) veins
- Reddish-blue discoloration
- Skin that is warm to the touch
- Vein may feel like a cord
- Darkened, discolored or stained skin or non-healing ulcers
- Veins distended, sudden appearance of new vessels

Symptoms of Deep Venous Thrombosis [18]
- Sometimes asymptomatic
- Pain in the leg
- Tenderness in the calf
- Leg tenderness
- Swelling of the leg
- Increased warmth of the leg
- Redness in the leg
- Bluish skin discoloration
- Discomfort when the foot is pulled upward

Symptoms of Renal Vein Thrombosis [14]
- Flank or low back pain; may be severe
- Urine bloody
- Urine output decreased

Symptoms of Atheroembolic Renal Disease [8]
- Maybe asymptomatic
- Foot pain, ulcers on feet, or “blue toes”(due to decreased blood flow to toes and feet)
- Pain in the abdomen, nausea or vomiting(due to decreased blood flow to the intestine)
- Strokes or blindness
- Flank pain and blood in the urine(rare)
- Uncontrolled high blood pressure

Symptoms of Stroke [12]
- Weakness or paralysis of an arm, leg, side of the face, or any part of the body
- Numbness, tingling, decreased sensation
- Vision changes
- Slurred speech, inability to speak or understand speech, difficulty reading or writing
- Swallowing difficulties or drooling
- Loss of memory
- Vertigo (spinning sensation)
- Loss of balance or coordination
- Personality changes
- Mood changes (depression, apathy)
- Drowsiness, lethargy, or loss of consciousness
- Uncontrollable eye movements or eyelid drooping

Symptoms of Angina [13]
- Chest pain
- Is not clearly focused in one location
- May spread to the arms, shoulder, neck, back and other areas
- May feel like gas or indigestion

The pain can be intense and severe or quite subtle and confusing. It can feel like:
- Squeezing or heavy pressure
- A tight band on the chest
- “An elephant sitting on [your] chest”
- Bad indigestion

The symptoms for the disorders include various conditions such as arterial embolism, pulmonary embolism, thrombophlebitis, deep venous thrombosis, renal vein thrombosis, atheroembolic renal disease, stroke, and angina. Each condition has specific symptoms that can help in diagnosis and treatment. Understanding these symptoms is crucial for prompt medical intervention and effective management of these disorders.