

A Case Study of Parkinson's disease Diagnosis using Artificial Neural Networks

Farhad Soleimanian Gharehchopogh
Department of Computer Engineering, Urmia
Branch, Islamic Azad University, Iran

Peyman Mohammadi
Department of Computer Engineering, Science and
Research Branch, Islamic Azad University, West
Azerbaijan, Iran

ABSTRACT

Artificial Neural Network (ANN)-based diagnosis of medical diseases has been taken into great consideration in recent years. In this paper, two types of ANNs are used to classify effective diagnosis of Parkinson's disease. Multi-Layer Perceptron (MLP) with back-propagation learning algorithm and Radial Basis Function (RBF) ANNs were used to differentiate between clinical variables of samples ($N = 195$) who were suffering from Parkinson's disease and who were not. For this purpose, Parkinson's disease data set, taken from UCI machine learning database was used. Mean squared normalized error function was used to measure the usefulness of our networks during trainings and direct performance calculations. It was observed that MLP is the best classification with 93.22% accuracy for the data set. Also, we got 86.44% accuracy in RBF classification for the same data set. This technique can assist neurologists to make their ultimate decisions without hesitation and more astutely.

General Terms

Pattern Recognition, Machine Learning.

Keywords

Parkinson's disease, Artificial Neural Network, Multi-layer Perceptron, Radial Basis Function.

1. INTRODUCTION

The causes of Parkinson's disease are unknown; however researches have shown that the degradation of the dopaminergic neurons affects the dopamine production [1]. Dopamine is a chemical messenger that helps the transmission of signals in the brain and other vital areas of the body. Our body uses dopamine to control movements, so decreasing dopamine in the circulation makes it more difficult for the person to control the movements, actions, and motions, even leading to tremors and numbness in extremities. Neurologists classify Parkinson's disease as a neurodegenerative disorder, leading to difficulty in motor function, including, walking and balance.

The number of Parkinson's disease patients is estimated to be 120–180 per every 100,000 people, although the percentage (and hence the number of affected people) is increasing as the life expectancies increases. For decades, researchers have strived to understand more about this disease and therefore to

find some methods for successfully limiting its symptoms which are most commonly periodic muscle tremor and/or rigidity. Other symptoms such as akinesia, bradykinesia, and dysarthria may occur in later stages of Parkinson's disease [2]. Recent developments in artificial intelligence have led to the generation of expert systems and decision support systems for medical applications. ANNs are inspired by the attempts to simulate biological neural systems. Axons are used to transmit nerve impulses from one neuron to another whenever the neurons are stimulated. A neuron is connected to the axons of the other neurons via dendrites which are extensions from the cell body of the neurons. The contact point between a dendrite and an axon is called a synapse [3].

Clinical determinations are often focused on physician's sense and experience based on his knowledge, coming from large database of the hospitals. The high dimensions of data which are in datasets need some techniques to be discovered and it used in clinical decisions. This information must be evaluated for medical researches to be used in health centers [4]. We can consider ANN as a branch of Artificial Intelligence (AI) or Expert Systems which is related to Fuzzy Logic (FL). An ANN could be regarded as a black box able to report output patterns after specialization of input patterns. This tool helps us with processing clinical data and gaining some helpful information, applicable by neuro pathologists.

The remaining of this paper is structured as follows: Section 2 describes ANNs; Section 3 explains ANN's applications in healthcare domain; the proposed methods for discovering useful information in Parkinson and utilized dataset description are presented in Section 4. Section 5 contains experimental results and discussions about proposed methods. Finally, some concluding remarks and proofs out come as well as future paper lines are presented in Section 6.

2. ARTIFICIAL NEURAL NETWORKS

An ANN is a computational structure, inspired by the study of biological neural processing. There are many different types of ANNs, from relatively simple to very complex; just as there are so many theories on how biological neural processing works [5]. In other words, an ANN represents a highly parallelized dynamic system with a directed graph topology that can receive the output information by means of a reaction of its state on the input actions. Processor elements and directed channels are called nodes of the ANN [6].

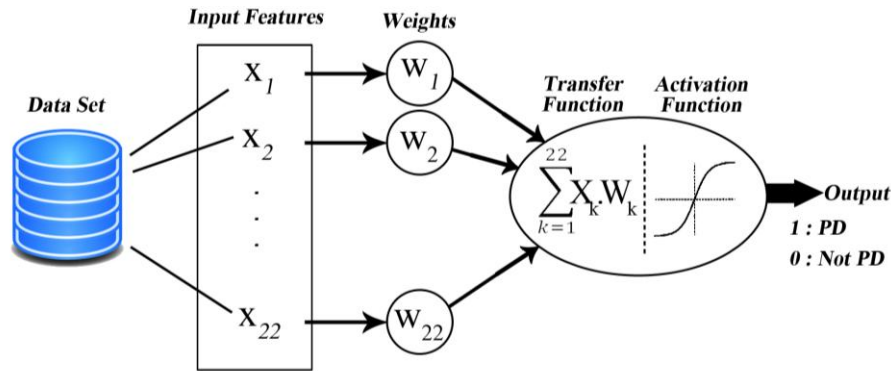


Fig 1: A simple topology of the ANN.

ANNs are computational tools for pattern classification that have been the subject of renewed research interest during the past 15 years [8, 9]. The works concerning threshold logic appeared in the 1960s and 1970s [7]. An ANN, a processing system sample, is inspired by biological neural systems such as brain. The key member of this novel structure is information processing system or numerous numbers of them, working interactively like as brain hormones in order to solve some special problems like pattern reorganization or data classifications through this training process.

Recent advances in the field of ANNs have made them attractive for analyzing signals. The application of ANNs has opened a new area for solving problems which are not resolvable by other signal processing techniques [8, 9]. The ANN node provides a variety of feed forward networks that are commonly called back-propagation networks. Back-propagation refers to the method for computing the error gradient for a feed forward network; a straightforward application of the chain rule of elementary calculus [10].

A simple and general topology of the ANN can be seen in Fig. 1. Each input neuron receives a numerical input from each of the variables in the dataset. All variables are normalized to a 0-1 scale. These values are then multiplied by factors, known as connection weights. The products of these multiplications are summed up and become the net inputs. Then these values enter into an activation function that calculates the outputs of the hidden neurons [11].

3. ANN IN HEALTHCARE

Equipping medical science with intelligent tools for the diagnosis and treatment of diseases purposes can reduce the deficiencies of physicians, decrease death toll and financial losses. In many medicine researches which need to be inferred from a huge dataset, we can't use manual execution methods because of their magnitude. In such cases, ANNs can help physicians with disease diagnosis. ANNs have been used in different clinical settings to predict the effectiveness of instrumental evaluation (echocardiography, brain single photon emission computed tomography, lung scans, prostate biopsy) in order to increasing diagnostic sensitivity and specificity and also used in laboratory medicine in general [13]. ANNs are inspired by the functioning processes of the human brain, which are adaptive models for the analysis of data [14].

For example some medicine applications of ANNs are:

- a) Patient Diagnose Systems: ANNs are widely used in disease detections. Those are able to diagnose and recognize cancers, cardiovascular diseases, tuberculosis, and sinus infections. One of the uses of ANNs is that some lateral factors such as tiredness, fatigue, emotional statuses, and work in specific conditions have no influences on them.
- b) Bio-chemical analysis: ANNs have various and extensive applications in analyzing blood and urine samples, tracking glucose levels in diabetics, and detecting ions' surface in body fluids.
- c) Analysis of Medical Imaging: ANNs are helpful tools in the analysis of tumors and MRI images.
- d) Drug Development: ANNs can be used as tools for the amplification of cancers and AIDS.

Also, in the design and manufacture of medical equipment science, ANNs are beneficiary auxiliary tools. We can test trail tools before implementing them as an actual case. The use of ANNs in biological systems has been proposed for different objects, which including studies on deoxyribonucleic acid sequencing and protein structure [12].

In 2002-2004 a research, following prostate radiotherapy was done on the application of ANN in predicting biological outcomes. An alternative method for predicting specific examples of tumor control and normal tissue complications is to use an ANN. It was shown that the ANNs were able to predict biochemical control and specific bladder and rectum complications with sensitivity and specificity of above 55% when the outcomes were dichotomized [15]. In one of the last-mentioned ascertainties in current year (2013), Daniel Ansari et al. enquired about ANNs as nonlinear pattern recognition techniques that can be used as a tool in medical decision making. Based on ANNs, a flexible nonlinear survival model was designed by using clinical and histopathological data from 84 patients. After the use of ANNs for predicting survival in pancreatic cancer, the C-index was 0.79 for the ANN and 0.67 for Cox regression [16].

Also, Suresh Kumar et al. obtained some beneficial information about the early detection of Parkinson's disease from blood plasma by FTIR micro-spectroscopy. A three-layer network implemented by NeuNet Pro Software (CorMac Technologies Inc., Canada) by choosing back-prop algorithm, which is used for the prediction of disease stages. For automatic detection, ANN was implemented on the variable regions, which shows 96.29% accuracy in the detection of disease progression [17].

4. PROPOSED METHODS AND IMPLEMENTATION

ANNs have wide and common uses in classification and recognition problems of health systems. In this section, we describe the materials and methods which are used in current research.

4.1 Date set

The Parkinson database used in this paper is taken from the University of California at Irvine (UCI) machine learning repository [18]. The dataset was created by Max little from the University of Oxford with collaboration of the National Centre for Voice and Speech, Denver, Colorado, who recorded the speech signals of patients. This dataset is composed of a range of biomedical voice measurements from 31 people, 23 of which were with Parkinson's disease. Each column in the table is a particular voice measure, and each row corresponds to one of 195 voice records from these individuals ("name" column). The main aim of the data is to discriminate healthy people from those with Parkinson's disease, according to "status" column which is set to 0 for healthy people and 1 for the people with Parkinson's disease.

This paper reviewed the literature and used the following 23 variables as explanatory variables in Table.1. Before building models of ANNs, the data set were randomly split into two subsets: 70% (n=137) of the data for training set and 30%

Table 1. Parkinson Patient Dataset Attributes

| Attribute | Description |
|------------------|---|
| MDVP : Fo (Hz) | Average vocal fundamental frequency |
| MDVP : Fhi (Hz) | Maximum vocal fundamental frequency |
| MDVP :Flo(Hz) | Minimum vocal fundamental frequency |
| MDVP: Jitter (%) | Several measures of variation in fundamental frequency |
| MDVP:Jitter(Abs) | |
| MDVP: RAP | |
| MDVP: PPQ | |
| RPDE | Two nonlinear dynamical complexity measures |
| D2 | |
| DFA | Signal fractal scaling exponent |
| spread1 | Three nonlinear measures of fundamental frequency variation |
| spread2 | |
| PPE | |
| NHR | Two measures of ratio of noise to tonal components in the voice |
| HNR | |
| Status | Healthy, Sick |

(n=58) of the data for validation set.

4.2 Multi-Layer Perceptron

In this section, we give a brief overview of the MLP_ANN as we are using it. This section can probably be overlooked for those who are familiar with the network. An

MLP is based on Rosenblatt's single perceptron [19]. In this paper, we are using conventional three-layer MLP network as classifier. The network contains an input layer, a hidden layer and an output layer. However there is merely a single neuron in the output layer. One architecture of the ANN, used in this paper, is the MLP network architecture with 22 input nodes, 5 hidden nodes and 1 output nodes in our approach. The number of input nodes is determined by the ultimate data; the number of hidden nodes is determined through trial and error; and the number of output nodes is represented as a rage, demonstrating the disease classification. Each neuron at one layer receives a weighted sum from all neurons in the previous layer and provides an input to all neurons of the later layer. The computation of the network's overall output, y , was based on Eq. (1):

$$y = g(\sum_{j=1}^m w_j^2 g(\sum_{i=1}^d w_{ji}^1 X_i + b_j^1) + b^2) \quad (1)$$

Therefore in our case, for the results presented, $d=22$ and $m=5$. All 22 input neurons were fully connected to the hidden layer. The single output neuron was, in turn, fully connected with the 5 hidden neurons, each one being defined by Eq. (1). The transfer function g for both hidden and output layer neurons was a tan-sigmoid transfer function described by Eq. (2). This function is a hyperbolic tangent sigmoid transfer function and a good tradeoff for ANNs, where speed is important and the exact shape of the transfer function is not.

$$\text{tansig}(n) = \frac{2}{(1+e^{-2n})} - 1 \quad (2)$$

Where n is the overall summed input to that neuron, respectively.

For the Parkinson Disease's classification problem, the MLP was required to output 1 for Parkinson's disease and 0 for healthy (non- Parkinson's disease status) people.

The simulations were performed with MATLAB's version 7.12.0.635 (R2011a) installed. The whole system of ANNs, and dividing and transfer functions was implemented in MATLAB (The MathWorks, Inc., Natick, MA). We used the following function in details for training which creates a feed-forward back-propagation network, in details:

$$\text{net} = \text{newff}(P, T, S, TF)$$

where net is the trained MLP, P is the $N \times M$ matrix containing the values of the dataset variables for the N training peoples with M attributes, and T is a $N \times 1$ vector containing the target class values with 1 or 0 value, and S is the sizes of hidden layers, and finally TF is transfer function of the layer. It's Tansig for hidden layer, as showed by Eq. (2). In Fig. 2. the structure of MLP implemented in this paper is shown.

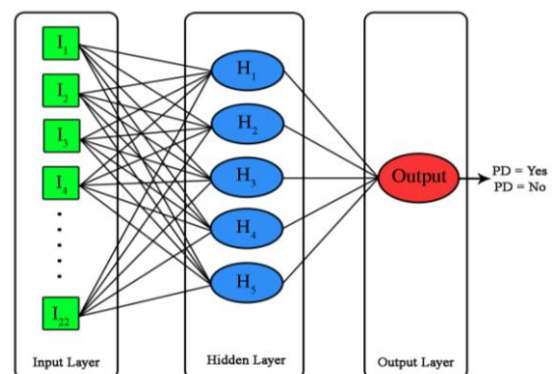


Fig 2: MLP structure that used.

4.3 Radial Basis Function

RBF networks was originally proposed by Lowe and Broomhead in 1988[20]. This network also consists of input, hidden and output layers. One of the principal differences between RBF and MLP is the neuron transfer function inside the hidden and output layer. The RBF generically (and certainly in our paper) uses a Gaussian function as shown in Eq. (3). The RBF consists of only two layers with a single neuron in the output layer. The output is computed according to Eq. (4); essentially it is merely a weighted linear combination of the hidden layer neuron outputs.

$$\varphi_i(x) = \exp\left(\frac{-\|x - u_i\|^2}{\sigma_i^2}\right) \quad (3)$$

$$y = \sum_{i=1}^N w_i \varphi_i(X) + w_0 \quad (4)$$

Where $X = [x_1, x_2, \dots, x_m]^T \in \mathbb{R}^m$ is the input vector, $u_i = [u_{i1}, u_{i2}, \dots, u_{im}]^T \in \mathbb{R}^m$ is the center vector, and σ is the corresponding width of the i_{th} node. N , meanwhile, indicates the number of neurons in the hidden layer. A similar type of RBF method has been previously demonstrated to work well with this type of data [21, 22]. Here, however, according to the comparison conducted, we have focused on the specific data considered in terms of signal frequency. There are two design aspects to an RBF network: (1) Determination of the parameters of the RBF neurons, which can also be translated into Gaussian function widths; and (2) Calculation of the connection weights between the hidden and output layers. In this paper, we used a fuzzy c means, clustering algorithm to determine the initial centers. For RBF implementation, we used the ANN tools which are available in Matlab. In fact, we used the following function to train designed radial basis network:

$Net = newrb(P, T, goal, spread)$

where net is the trained RBF, P is the $N \times M$ matrix containing the values of the dataset variables for the N training people with M attributes, and T is a $N \times 1$ vector containing the target class values, with 1 or 0 value; and $goal$ is Mean squared error goal that we set it to 0.0 value, and finally $spread$ is spread of RBFs whose default value is 1.0 and we set it in our paper to

10. The larger the spread is, the smoother the function approximation becomes. A too-large spread means a lot of neurons are required to fit a fast-changing function. Too small a spread means many neurons are required to fit a smooth function, and the network might not generalize well. The *newrb* adds neurons to the hidden layer of a RBF network until it meets the specified mean squared error goal.

5. RESULT AND DISCUSSION

In our paper, MLP and RBF models were implemented, using the neural network toolbox (v 7.0.1) provided with Matlab software (v R2011a) published by the Math Works Inc. The data set, used in the detection of Parkinson's disease and healthy status with two ANNs, was divided into two subsets as Training and Test data. There were a total of 136 pieces of data, exclusively used for training and 59 pieces of data exclusively used for testing. The ANN that we designed was used as a divider and function to divide dataset randomly into two subsets. We used default derivative function to choose the recommended derivative algorithm for the type of network whose derivatives are being calculated. This function is used to calculate error gradients and Jacobians when the network is trained to use a supervised algorithm such as back propagation.

In implementations of ANNs in this paper, we utilize Levenberg-Marquardt back-propagation to train our networks. This is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization. Levenberg-Marquardt is often the fastest backpropagation algorithm in the toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms. Fig. 3 presents a comparison of the real statuses and our MLP detected statuses; Fig. 4 shows the differences between real outputs and designed RBF for test data.

Fig. 3 and 4 show the differences between real outputs and our ANN's outputs. The X-axis shows the output, representing Parkinson's disease and healthy status while the Y-axis shows the cases. The red lines are real output of test

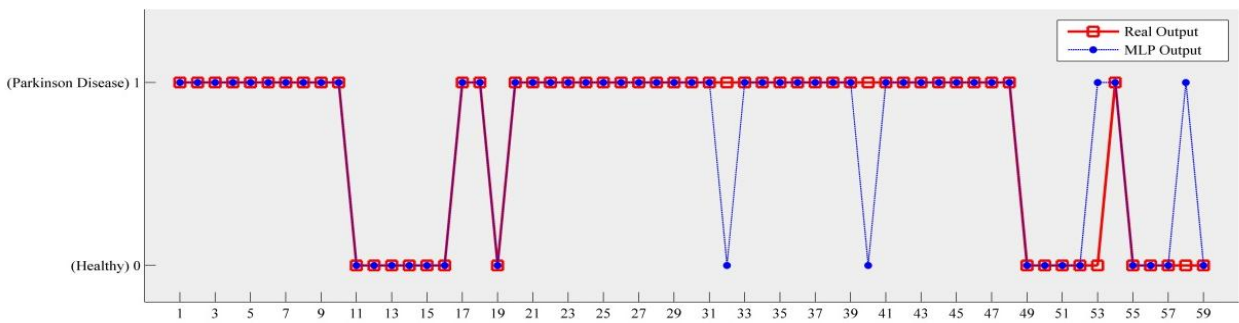


Fig 3: MLP classifier test output

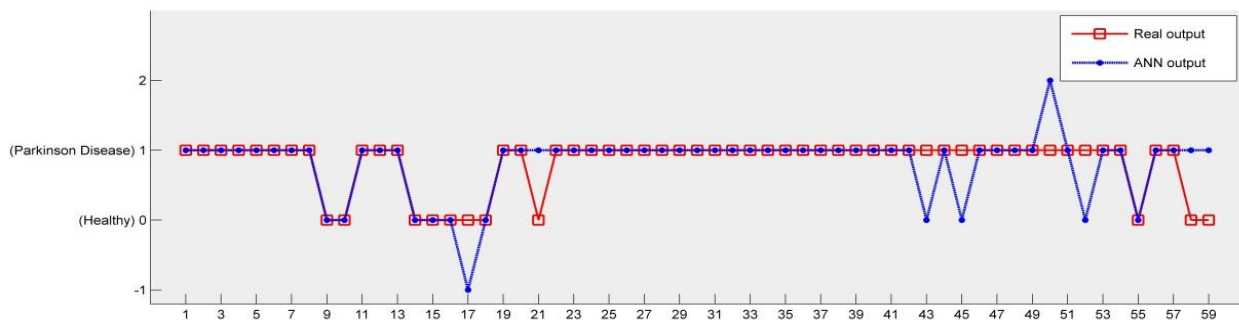


Fig 4: RBF classifier test output

data, and the red squares display Parkinson's Disease or Healthy status for each test case; the blue lines are our ANNs output of the same test data, and blue circles display Parkinson's disease or Healthy for each test case, detected by designed ANNs. Identical position of a square and circle for each instance indicates that our ANN correctly classified that case and vice versa.

Experiments were calculated information, obtained from Matlab Neural Networks Toolbox. The classification accuracies obtained by this paper for Parkinson's disease dataset are presented in Table 2. We first used MLP to classify data set; the accuracy of this model in test dataset was 93.22%; so our MLP network classified 55 instances out of 59 correctly. Furthermore, the accuracy in training data set was 96.32%. Meanwhile the RBF was implemented with the accuracy of 86.44% for test data, while displaying less accuracy than MLP. As seen in Table 2, it could correctly classify 51 instances out of 59. Khemphila and Boonjing in 2012 reported 91.453% classification accuracy for training set with the same 22 attributes by means of using Multilayer is feed-forward ANNs trained with the standard back-propagation algorithm [3]. But we obtained classification accuracy of 96.323% for training set with MLP method.

6. CONCLUSION AND FUTURE WORKS

The aim of this paper was to compare the efficiency of MLP and RBF on a specific medical application, namely

Table 2. MLP and RBF ANNs Test Results Accuracy

| Classifier | Hidden Neurons | Accuracy (%) | Correctly Classified Instances | | Incorrectly Classified Instances | |
|------------|----------------|--------------|--------------------------------|------------|----------------------------------|------------|
| | | | Num | Percentage | Num | Percentage |
| MLP | 5 | 93.2203 | 55 | 93.22 | 4 | 6.78 |
| RBF | 136 | 86.4407 | 51 | 86.44 | 8 | 13.56 |

Parkinson's disease diagnosis. The studied disease was Parkinson's disease and we used ANNs to classify patients and healthy peoples. The results showed that we can widely use ANNs to diagnose Parkinson's disease, so neurologists and medicine centers can use ANNs in automatic disease diagnosis systems.

In future studies, researchers can exert other types of ANN or machine learning algorithms for Parkinson's disease diagnoses, to change instance attributes into finding momentous attributes. So many attributes cause to increase correct classification accuracy and also increase training dataset is a good idea to test the accuracy rate.

7. REFERENCES

- [1] Hoglinger GU, Rizk P, Muriel MP, Duyckaerts C, Oertel WH, Caille I, et al. Dopamine depletion impairs precursor cell proliferation in Parkinson disease. *Nat Neurosci*, Vol: 7, pp: 726–735, 2004.
- [2] Aziz, T. Z., Peggs, D., Sambrook, M. A., & Crossman, A. R., Lesion of the sub thalamic nucleus for the alleviation of 1-methyl-4-phenyl-1,2,3,6-

- tetrahydropyridine (MPTP)-induced parkinsonism in the primate. *Movement Disorders*, Vol: 6, No:4, Pp:228–292, 1992.
- [3] Khemphila A, Boonjing V. Parkinsons Disease Classification using Neural Network and Feature selection. *World Academy of Science, Engineering and Technology*, Vol: 64, pp:15-18, 2012.
- [4] Farhad Soleimani Gharehchopogh, Peyman Mohammadi and Parvin Hakimi. Application of Decision Tree Algorithm for Data Mining in Healthcare Operations: A Case Study. *International Journal of Computer Applications*, Vol: 52, No:6, PP:21-26, August 2012.
- [5] Valluru B. Rao and Hayagriva Rao. 1995. C++, Neural Networks and Fuzzy Logic (2nd Ed.). MIS: Press, New York, NY, USA.
- [6] Alexander I. Galushkin. 2007. Neural Network Theory. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- [7] Shigeo Oyagi, Ryoichi Mori, Noriaki Sanechika Realization of a Boolean function using an extended threshold logic. *Bulletin of the Electro technical Laboratory*, Vol: 42, PP: 9–74, 1978.
- [8] I.A.Basheer, M.Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, *J.Microbiol. Meth.* Vol:43, PP: 3–31, 2000.
- [9] B.B.Chaudhuri, U.Bhattacharya, ENcient training and improved performance of multilayer perceptron in pattern classification, *Neurocomputing*, Vol:34, pp: 11-27, 2000.
- [10] Hanbay, D., Turkoglu, I., & Demir, Y. (2008). An expert system based on wavelet decomposition and neural network for modeling Chua's circuit. *Expert Systems with Applications*, Vol: 34, No:4, Pp: 2278–2283.
- [11] Tran Nguyen, Richard Malley, Stanley H. Inkelis, Nathan Kuppermann, Comparison of prediction models for adverse outcome in pediatric meningococcal disease using artificial neural network and logistic regression analyses, *Journal of Clinical Epidemiology*, Vol: 55, No: 7, Pp: 687-695, 2002
- [12] Parbhane, R.V., Tambe, S.S., Kulkarni, B.D., ANN modeling of DNA sequences: New strategies using DNA shape code, *Computer. Chem.* Vol: 24, Pp: 699-711, 2000.
- [13] Tafeit, E., Reibnegger, G., Artificial neural networks in laboratory medicine and medical outcome prediction. *Clin. Chem. Lab. Med.* Vol: 37, Pp: 845-853, 1999.
- [14] Kenji S., Artificial Neural Networks - Methodological Advances and Biomedical Applications. IN-TECHJaneza Trdine 9, 51000 Rijeka, Croatia, April, 2011.
- [15] Sarah L. Gulliford, Steve Webb, Carl G. Rowbottom, David W. Corne, David P. Dearnaley, Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate, *Radiotherapy and Oncology*, Vol: 71, No:1, PP 3-12, 2004.
- [16] Daniel Ansari, Johan Nilsson, Roland Andersson, Sara Regnér, Bobby Tingstedt, Bodil Andersson, Artificial neural networks predict survival from pancreatic cancer after radical surgery, *The American Journal of Surgery*, Vol: 205, No: 1, Pp: 1-7, 2013.

- [17] Shiek S.S.J. Ahmed, Winkins Santosh, Suresh Kumar, T. Hema Thanka Christlet, Neural network algorithm for the early detection of Parkinson's disease from blood
- [18] Parkinson's Data Set, UCI repository of machine learning databases available from <ftp://ftp.ics.uci.edu/pub/machine-learning-databases/parkinsons/parkinsons.data>, [last accessed: 27 April 2013].
- [19] K., Warwick, Artificial intelligence: The basics. Taylor & Francis, 2011.
- [20] D.S. Broomhead and D. Lowe. Multivariate functional interpolation and adaptive networks. *Complex Systems*, Vol: 2, Pp: 321-355, 1988.
- plasma by FTIR micro-spectroscopy, *Vibrational Spectroscopy*, Vol: 53, No:2, Pp: 181-188, 2010.
- [21] Wu, D., Warwick, K., Ma, Z., Burgess, J. G., Pan, S., & Aziz, T. Z., Prediction of Parkinson's disease tremor onset using radial basis function neural networks. *Expert Systems with Applications*, Vol: 37, No: 4, Pp: 2923–2928, 2010.
- [22] Song Pan, Serdar Iplikci, Kevin Warwick, Tipu Z. Aziz, Parkinson disease tremor 's classification – A comparison between Support Vector Machines and neural networks, *Expert Systems with Applications*, Vol: 39, No: 12, Pp: 10764-10771, 2012