PREDICTING THE FUNCTIONAL RESULT OF TYMPANOPLASTY WITH THE AID OF NEURONAL NETWORKS USING A REDUCED DATA SET

Oana SANDU [1], Daniel-Petru GHENCEA [2], Miron ZAPCIU [3]
Caius-Codruț SARAFOLEANU [4]

Rezumat. Aplicațiile inteligenței artificiale își fac simțite prezența, în ultimii ani, în foarte multe domenii, inclusiv în medicină. Întrucât medicina este un domeniu cu un nivel de imprevizibilitate mare, metodele tradiționale de statistică medicală sunt insuficiente. Aplicațiile inteligenței artificiale, așa cum sunt rețelele neuronale artificiale, logica fuzzy sau algoritmul genetic, permit crearea de modele inteligente care pot predicționa răspunsul pacientului la tratament, sau pot determina riscul la care pacientul este expus pentru o anumită patologie. Lucrarea are scopul de a prezenta un model de rețele neuronale artificiale pentru predicția rezultatului funcțional după o intervenție chirurgicală asupra membranei timpanice (timpanoplastie). Metodologia aplicată în această cercetare are la bază un set redus de date de intrare, una/două intrări în stratul de intrare și ieșiri multiple în stratul de ieșire.

Abstract. Applications of artificial intelligence have made their presence felt over the last few years in many areas, including medicine. Since medicine is a field of high unpredictability, traditional methods of medical statistics are insufficient. Artificial intelligence applications, such as artificial neural networks, fuzzy logic, or genetic algorithm, allow intelligent models to predict the patient's response to treatment or to determine the risk to which the patient is exposed for a particular pathology. Our study aims to present a model of artificial neural networks to predict the functional outcome after surgery on the tympanic membrane (tympanoplasty). The methodology applied in this research is based on a reduced set of input data, one / two entries in the input layer and multiple outputs to the output layer.

Keywords: *tympanoplasty, artificial neural network, prediction*

1. Introduction

Computer science changed completely our life in the last years. Computers are used in a diversity of disciplines such as: aerospace, biology, chemistry, physics,

[1] PhD Student, Carol Davila University of Medicine and Pharmacy, Bucharest, e-mail: oanastelea@yahoo.com.
[2] PhD. Eng.-Ec., University Politehnica of Bucharest, Romania, e-mail: daniel.ghencea@blackseasuppliers.ro
[3] Prof. PhD. Eng., University Politehnica of Bucharest, Corresponding member of Academy of Romanian Scientists, miron.zapciu@upb.ro
[4] Prof. PhD. Dr., Carol Davila University of Medicine and Pharmacy, Bucharest, e-mail: csarafoleanu@gmail.com

economics, meteorology. The applicability of artificial neural networks has proven efficient in many medical specialties. With the help of these, artificial neural models can be developed that can be useful in supporting a diagnosis, as it is well known that the diagnosis of a pathology involves more inaccurate data and some levels of uncertainty, absolutely inherent in medical practice. They can also be used to follow up a patient status, demonstrate the efficacy of a particular treatment or surgical technique and predict the risk of illness.

Artificial neural networks offer a number of advantages compared to classical medical statistics, including:

- Statistical training is less formal. Artificial neural networks have the ability to learn and model nonlinear and complex relationships, which is really important because in medicine, many of the relationships between inputs and outputs are both nonlinear and complex;
- The ability to detect complex nonlinear relationships between dependent and independent variables;
- An artificial neural network can generalize, since after learning from their initial inputs and relationships, it can deduce unseen relationships and unseen data, making the generalized model and predicted on non-identifiable data;
- The ability to detect all possible interactions between predictor variables;
- Availability of several training algorithms.

Unlike many other prediction techniques, the artificial neural network does not impose any restrictions on input variables. In addition, many studies have shown that artificial neural networks can better model data with high volatility and non-constant variance, given their ability to learn hidden relationships in data without imposing fixed relationships. This is very useful in predicting a particular pathology where data volatility is very high.

2. Current level of knowledge

The complexity of medical practice makes the traditional approaches to medical statistical analysis insufficient at present. In medicine there are a number of inaccurate or missing information and therefore the use of artificial intelligence is a real help. Multiple studies have explored the skill of artificial neural networks, fuzzy logic and genetic algorithms in the field of medicine.

The theory of fuzzy logic was created and developed by L.A. Zadeh [1]. This logic makes it possible to define medical entities, which are often inaccurate, in fuzzy sets. The main purpose of using fuzzy logic is to formulate uncertainty for applications such as making certain decisions. To control hypertension during general anesthesia, Oshita and colleagues conducted a study using fuzzy logic [2]. The authors then developed an automated blood pressure control system based on

a fuzzy system and evaluated its efficacy and clinical safety. The main feature of a fuzzy logic controller is that its action is very close to human judgment. It has been concluded that the operating state diagram and automatic blood pressure controller, both made possible with the concept of Zadeh's fuzzy logical control rules, are satisfactory and safe methods of administration of nicardipine.

The genetic algorithm is a mathematical method, inspired by the laws of medical genetics (cross-over and mutations). A study led by Castiglione et al. used the genetic algorithm to determine the best antiviral therapy (HAART) in patients diagnosed with HIV. A virtual system of an immune system has been developed that has been used to evaluate the effects of medication [3]. Another study of the literature that used genetic algorithms was developed by Jaremko et al., who developed a genetic algorithm to estimate the angle of the spine deformation in patients diagnosed with scoliosis. The developed system was able to determine with an accuracy of 5% to more than two thirds of their patients [4]. Latkowski and Osowski used the genetic algorithm to select the most relevant genes associated with autism [5]. Normally, there are several genes associated with autism (RMI1, NRIP1, TOP1, ZFHX3, CEP350, NFYA, PSENEN, ANP32A, SEMA4C and SP1), but this study has introduced a set of classifications that recognized the presence of genes with a sensitivity of 96% and with a specificity of 83%.

Artificial neural networks are computational systems that, at least in theory, are similar to biological neural networks. The similitudes between the two types of networks are, in fact, extremely limited. Unlike previous classification systems, these systems are not governed by predefined set of rules (logical function within linear regression, kernel). Neural networks have been used in many disciplines (physics, chemistry, biology, economics, meteorology) and have all proved to be effective. This diversity of use suggests that they can also be applied in the field of medicine to help either in medical research or to improve medical care. Such an artificial neural network can help the doctor to plan and optimize a particular surgical technique, can be helpful in managing treatments by providing certain preconditions, facilitating the processes necessary for medical statistics.

There are many parameters that can be set for an artificial neural network with deep learning and the most important are:

- The number of hidden layers, a higher number of layers increases the complexity of the network and the computational cost;
- The number of neurons (nodes) in each layer;
- The way to quantify the relationship between input and output;
- The activating function, which determines how the information from inputs (predictors) is processed by hidden layers. There are several functions that can be used;

- "Loss" function, the function that determines how errors are calculated (differences between network and actual values);
- The number of epochs of the artificial neural network, i.e. how many times the artificial neural network data (similar to the number of repetitions of the matter before an exam) is analyzed. Generally, a greater number of epochs is associated with better network performance.

An interesting study by a Polish team led by Szalenic presented the results of tympanoplasty using artificial neural networks. Unfortunately, their study has not yet standardized a prognosis model for improving postoperative hearing. Observational studies have shown that only postoperative clinical examination and surgical technique itself are not the determinants of surgical success. In this study, the best developed artificial neural network model had an accuracy of 98%, 89% for learning and 84% for validation compared to classical statistical methods that produced predictions with an accuracy of 75%. The study has shown that hearing improvement after timing can be anticipated by using artificial neural network. Hearing disorders in small children can be difficult to recognize and often go undetected until older age[6], [7].Dunmde and his collaborators have developed an artificial neural network for the diagnosis and classification of hearing loss. This artificial neural network can support the ENT physician for a faster detection of hearing loss and can help children to reach their full educational potential [8], [9].

3. Methodology for simulation, prediction and interpretation of results

The initial set of data based on the characteristics of 50 (male and female) patients with tympanoplasty aged between 12 and 61 was divided into two subsets corresponding to the sex type.
Because the data set has both numerical and linguistic variables, it is necessary to transform linguistic variables into numerical variables.

The Visual Gene Developer 1.7 (VGD) software used for prediction has an RNA-BP module. Since VGD works with input data sets in the [-1, 1] range, the data sets being overrunning must be demultiplied. Finally, the prediction result will be multiplied by the size of the factor with which it was demultiplied.

3.1. Simulation, prediction and interpretation of results for male patients

The dataset is divided into three age categories: 31-39, 40-49 and 50-58. Patients aged 12, 25 and 29 are eliminated because the minimum set of data required to create an age category is not met. RNA-BP training criteria is also not met and this could lead to high prediction errors.

3.1.1. Simulation, prediction and interpretation of age outcomes of missing patients in age data set

For the 3 age categories, there is a prediction for the missing ages within each age group as follows:

- for category 31-39 we want - 3 predictions with 7 measurements for training;
- for category 40-49 we want - 4 predictions with 6 measurements for training;
- for category 50-58 we want - 6 predictions with 4 measurements for training;

Based on 17 measurements representing (56.67%), we predict 13 (43.33%) possible outcomes.

The research focuses on identifying a RNA-BP architecture that can be used in general terms.

For Predictions 1 and 2, the datasets we want to predict are within the domain, and for Prediction 3 a set of data is out of the domain (59 years), so RNA-BP has reference values for training and validation, which leads to the possibility of accurately predicting the values outside the upper range. The data set is characterized by a single age pair (39 years) for which there are different results. This duplication does not influence RNA prediction because it is at the top of the domain.

The regression coefficient has a value of 0.93 to 1 in 16 cases (89%), so it is close to 1, which indicates that variation of the response variable can be explained by explanatory variables and in 2 cases (11 %) is due to unknown variables or variables inherent.

For Prediction 1 (Figure 1) a uniform distribution of predicted data set along the slope is observed, resulting in 2 cases at a low regression coefficient value (0.64 and 0.81). For Predictions 2 and 3 the values of the regression coefficient are over 0.93 being assigned to a distribution on the extreme poles of the predicted data set [10].

The RNA-BP architecture used in the three cases is complex, female [11], with 5 hidden layers and 10 knots per hidden layer.

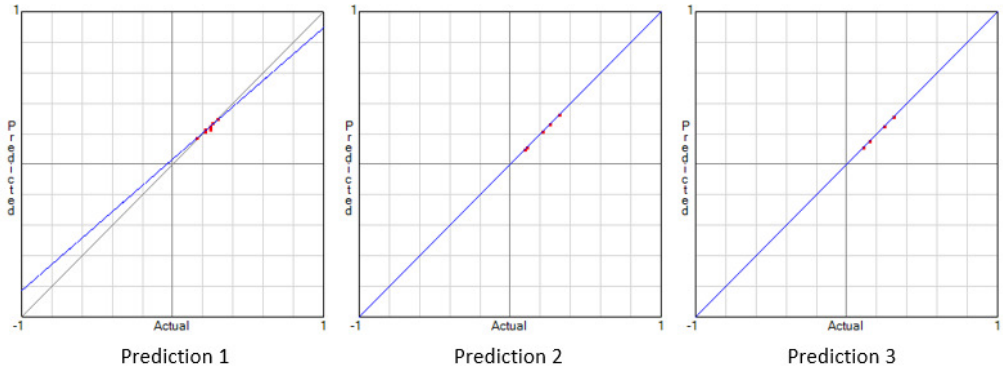


Fig. 1 The prediction data distribution pattern in 3 cases

In the Figure 2 we can see that in the case of Prediction 2 and Prediction 3 the information flows to cover the entire range ([-1, 1]) with which VGD works, which leads to ensuring values for the coefficient of regression close to 1 for all outputs from the exit layer. It is also noted that the importance of the connections between the input layer and the hidden first layer has high values, which are shown by the width of the connection.

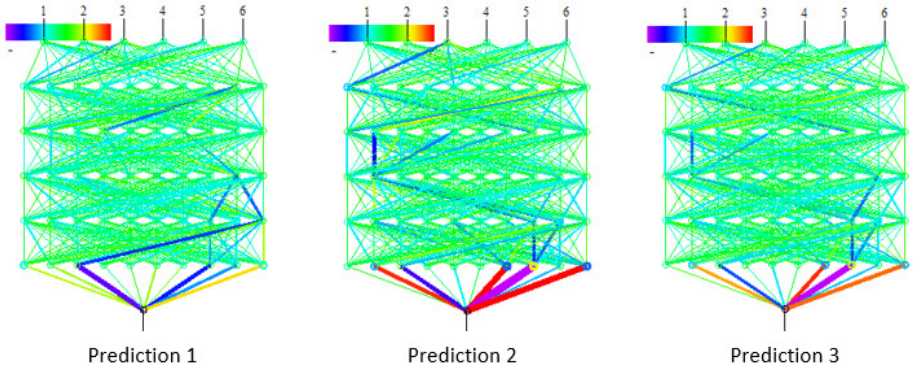


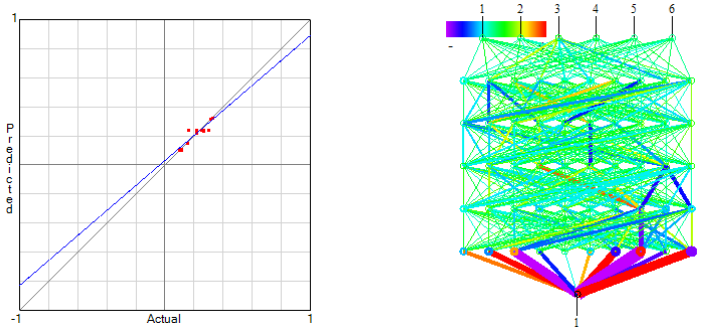
Fig. 2 RNA-BP architecture andthe values/importance of information flow connections in 3 cases

All 3 predictive architectures have been set for 100,000 training cycles. Note that for Predictions 2 and 3, although RNA-BP was set for 100,000 training cycles, it reached its maximum performance in 78,506 and 64,660 training cycles respectively.

Conclusion: Given that the predicted regression slope aligns with the regression slope almost totally in 2 cases (Prediction 2 and Prediction 3) and in 89% the regression coefficient has values above 0.93 with an average of 0.99643 and the predicted results are in the range of those measured by age groups, we can say that we have a prediction with a close trend to very good.

3.1.2. Simulation, prediction, and interpretation of age outcomes of missing Patients from the age data set

In this paragraph we analyze the prediction without interval division and the results and compare them with the ones in the previous paragraph (1.1.2) where we made the prediction for age parameter breakdown by intervals.



a) b)

Fig. 3 Predictive data set distribution module (a), RNA-BP architecture values/importance of information flow connections (b)

Figure 3 shows that the distribution of the predicted values is uniform on the regression slope and the regression coefficient obtains values well below 0.93 (0.4918-0.8995).

The RNA-BP attempt to identify a best algorithm is accomplished by exploring the entire range of values [-1, 1] and high values of the importance of the connections between the input layer and the first hidden layer manifested through the width of the connection.

Conclusion: Comparing the results for the two predictive modes, we can see the following:

- two errors below 1% (0.51% and 0.75%);
- error below 5% (2.85%);
- two errors below 10% (6.87% and 9.9%);
- maximum error - 12.73%;

An *average error of 5.6%* is obtained**.**

3.1.3. Simulation, prediction, and interpretation of results for the size of the eardrum perforation in the initial set of data

To perform this research, it is used from the original data set:

- Input variables: age and pre-operative result
- Output variable (prediction): perforation

From this analysis, we want to determine the minimal error for the hidden 5 layers architecture, 10 nodes per hidden layer, and 100,000 training cycles used in the predictions in paragraphs 3.1.1 and 3.1.2.

One output is shown for which the regression coefficient exceeds 0.93. Predictive values have a data set distribution on two poles on the regression slope.

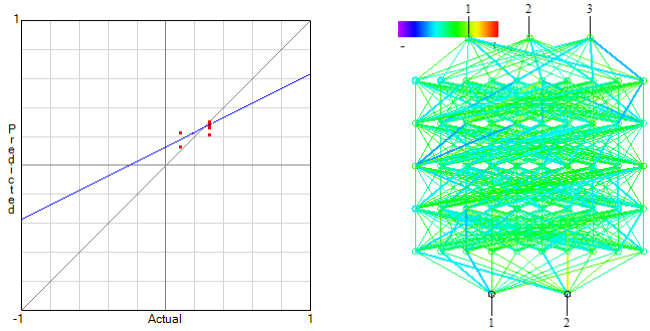


Fig. 4 Predictive data set distribution (a), RNA-BP architecture, and values/importance of information flow connections (b)

Conclusion: Under the above conditions we obtain:

- a 5.6% error with over-valuation;
- 2 errors (7,14% and 20%) with under-valuation;

In this case, an *average error of 10.9%* results.

3.1.4. Conclusions on prediction for male patients

For the 3 age groups analyzed (male patients), RNA-BP with the predicted characteristics obtained the following performances:

a) RNA-BP that used an unbiased data set (paragraph 2.1.1 vs 2.1.2):
   1. Category 31-39 - for the numerical part of the predicted 3 values of 6 i.e. 50% and for the linguistic part was predicted 4 values from 12 ie 33%;
   2. Category 40-49 - for the numerical part, 3 values of 8, ie 37.5%, were preached and for the linguistic part, 8 values from 16 to 50% were predicted;
   3. Category 50-58 - for the numerical part of the predicted 10 values of 12 ie 83,34% and for the linguistic part 22 of 24 values, meaning 91,67%;

compared to the age group predicted data set.

b) RNA-BP that predicted domain data of measured values:
   - for the linseed part predicted 17 out of 24, i.e. 70.84%, of which 3 full rows;

compared to the age group predicted data set.

3.2. Simulation, prediction and interpretation of results for female patients

The data set is used to simulate, predict and interpret the results for female patients. The data set is divided into four age categories: 20-29, 30-39, 40-49 and

50-59. Patients aged 15, 16, 18 and 61 are eliminated because the minimum set of data required to create an age category is not met. RNA-BP training criteria is also not met and this could lead to high prediction errors.

3.2.1. Simulation, prediction and interpretation of age outcomes of missing patients in age data set

For the 4 age categories, the prediction for the missing ages within each age category is as follows:

- for category 20-29 we want - 6 predictions with 8 measurements for training;
- for category 30-39 we want - 4 predictions with 6 measurements for training;
- for category 40-49 we want - 5 predictions with 6 measurements for training;
- for category 50-59 we want - 4 predictions with 6 measurements for training;

Based on 26 measurements representing (57.78%), we predict 19 (42.22%) possible outcomes.

For Predictions 2 and 4, the data sets we want to predict are inside the domain, and Predictions 1 and 3 have data sets outside the domain. Thus, for Prediction 1, RNA-BP has no reference values for the lower domain (20, 21 and 22 years) and Prediction 3, RNA-BP has reference values for training and validation for the upper domain (48 and 49 years).

For the data set we have the following non-homogeneous characteristics that affect RNA-BP:

- Prediction 1 - the set contains multiple results for ages 27 (preoperative 2 x 25.00 and 1 x 32.50) and 29 years (preoperative 1 x 16.30 and 2 x 26.30);
- Prediction 3 - the set contains multiple results for age 42 (preoperative 1 x 12,50 and 1 x 31,30);

Values that in both cases are within the domain contribute negatively to the quality of the prediction. If the data set contains multiple values for multiple ages within the analyzed domain, RNA-BP manages to make the correct predictions.

The data set inconsistency also leads to regression coefficient values well below 0.93. For one age range (Prediction 2, 30-39 years) RNA-BP achieves a regression coefficient of over 0.99, averaging 0.997. The overall average for the regression coefficient is 0.378.

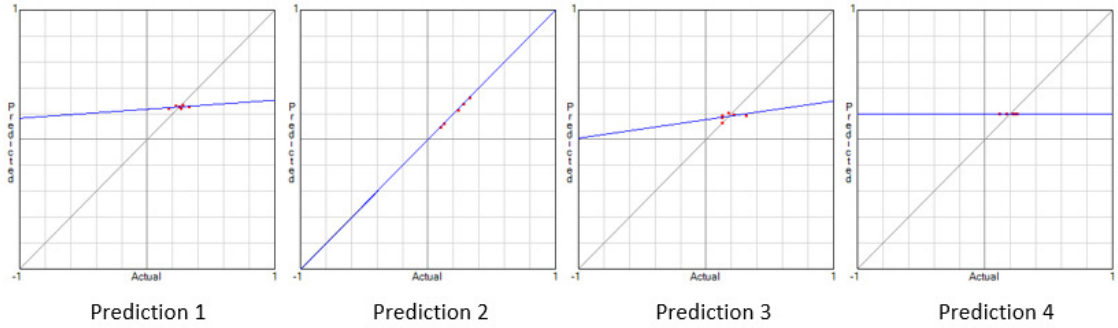


Fig. 5 The prediction data distribution pattern in 4 cases.

The RNA-BP architecture used in the four cases does not have the same features; the masculine analysis is centered on two configurations:

- 5 hidden layers and 10 knots per hidden layer (Predictions 2 and 4);
- 5 hidden layers and 7/7/5/5/5 nodes per hidden layer (Predictions 1 and 3);

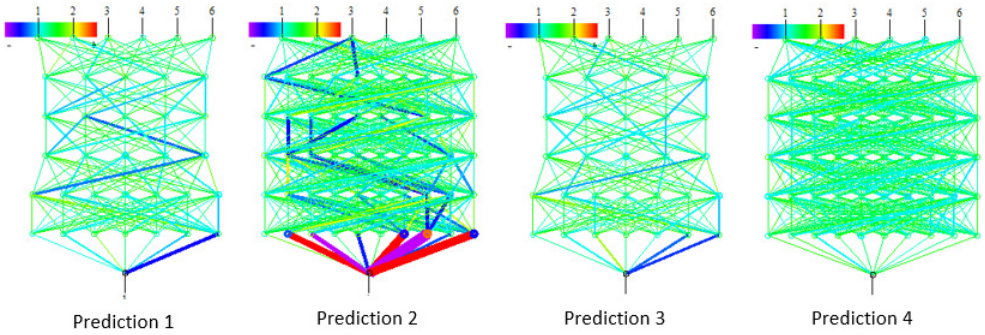


Fig. 6 RNA-BP architecture and the values/importance of information flow connections in four cases.

From Figure 6 we can see that in Prediction 2, for which the best regression coefficients were obtained, the flow of information values covers the whole range ([-1, 1]) with which VGD works, which leads to the provision of values for the coefficient of regression close to 1 for all outputs in the output layer. It is also noted that the importance of the connections between the input layer and the hidden first layer has high values, which are shown by the width of the connection. Prediction 1 and Predication 3 are highlighted by the use of negative information flows between the input layer and the output layer.

Prediction 2 also shows very good values in terms of the amount of errors (0.002175) and the average error per 0.00006 output set. The other 3 predictions have an error rate of over 0.1 and an average error per 0.003 output set.

3.2.2. Simulation, prediction, and interpretation of age outcomes of missing patients from the age data set

In this paragraph we analyze the prediction without interval division and compare the results with those in the previous paragraph (3.2.1) where we made the prediction for age parameter breakdown by intervals.
Given that in the previous paragraph for Predictions 1, 3 and 4 poor results were obtained for the regression coefficient, we focus on two cases:

1. Obtaining prediction results with a coefficient of error (Prediction C) as close as possible to the ones in the previous paragraph;
2. Obtaining prediction results for the best possible regression coefficient (CR prediction) (approaching as much as 0.93).

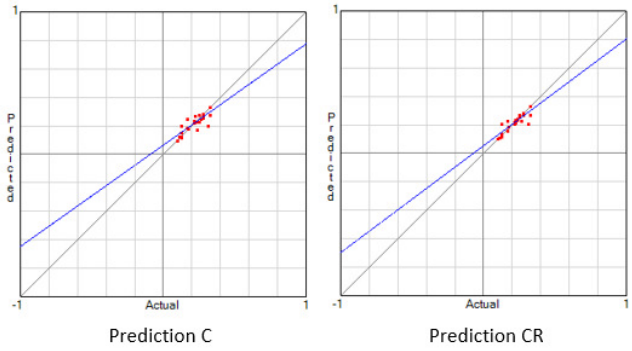


Fig. 7 The prediction data distribution pattern for the 2 cases.

In both cases an approximately uniform distribution on the regression slope is obtained.

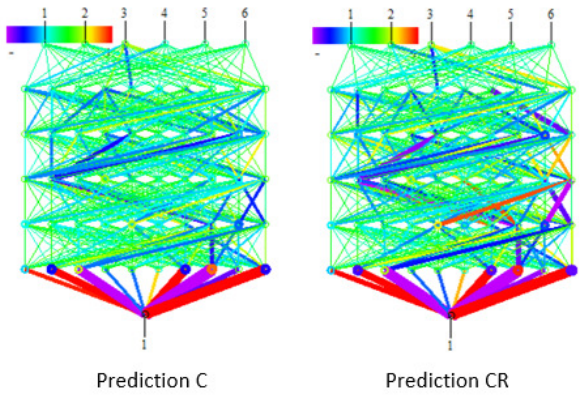


Fig. 8 RNA-BP architecture and the values/importance of information flow connections for 2 cases

The mean of the regression coefficient improves substantially, reaching the CR Prediction to double of the Media CR (0,378) with the division of the data set by age category.
RNA-BP predicts from a total of 114 values:

- for Predictive C (107,000 training cycles) 76 values;
- for CR Predictive (500,000 training cycles) 47 values;

These values are identical or very close (marked in green) than those obtained when the data set is divisible.
For 29 values (marked with yellow) (the Location and Size of tympanic membrane perforation Indicator), these are the same in all three cases of a total of 57 cases, which is 50.88%.

Conclusion: Not in all cases a regression coefficient of over 0.93 leads to accurate predictions. Table 16 shows that the average error rate for 107k training cycles is 1.5 times lower than for 500k training cycles due to the lower regression coefficient (0.60 vs. 0.75).

3.2.3. Simulation, prediction, and interpretation of results for the size of the tympanic membrane perforation in the initial set of data

In this case, the best predictions were made of the recorded data set (the order in which female patients were recorded), so random set of the set of data. The prediction values have a data set distribution on two poles on the regression slope, preserving the initial distribution.

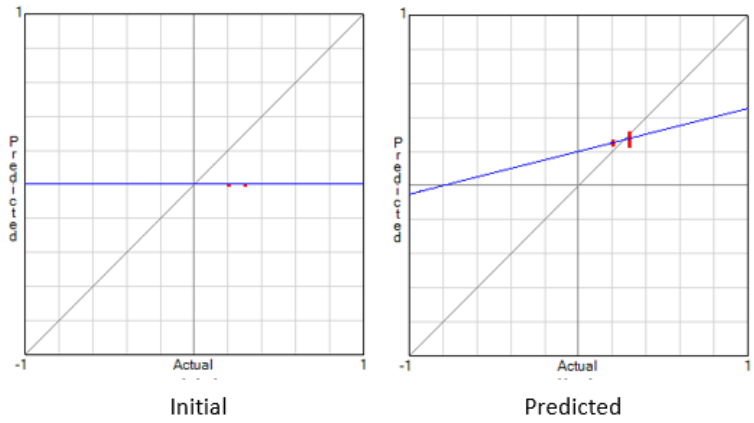


Fig. 9 The distribution of the original and predicted data set.

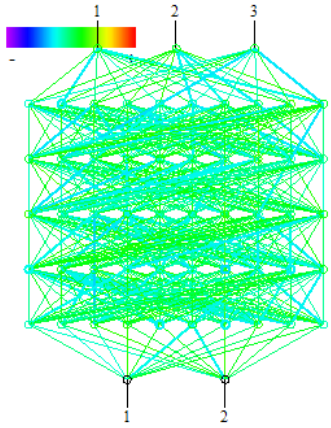


Fig. 10 RNA-BP Architecture and Values/Importance of Information Flow Connections

Conclusion: Under the above conditions, we obtain:

- 3.3% error with under-valuation;
- two errors (21.88% and 30%) with over-valuation;

resulting in an average error of 18.3%.

3.2.4. Conclusions on prediction for female patients

For the 4 age groups analyzed (female patients) RNA-BP having the predicted characteristics obtained the following performances:

a) RNA-BP using the dataset (Paragraph 3.1.1 vs 3.1.2):
   1. Category 20-29 - For the numerical part, 10 values from 12, i.e. 83,37%, were predicted, and for the linguistic part 21 of 24 values, i.e. 87,5%, out of which 4 full rows;
   2. Category 31-39 - for the numerical part it was predicted 7 values of 8 i.e. 87.50% and for the linguistic part were predicted 13 values from 16 i.e. 81.25%, of which 1 full row;
   3. Category 40-49 - for the numerical part of the predicted 3 values from 10 i.e. 30% and for the linguistic part were predicted 16 values from 20 i.e. 80%, of which 3 full rows;
   4. Category 50-58 - for the numerical part of the predicted 4 values of 8, i.e. 50% and for the linguistic part, 10 values from 16 i.e. 62.5%, of which 1 whole line, were predicted;

b) RNA-BP that predicted data from the range of measured values (Paragraph 3.1.1 vs 3.1.3):
   - For the measured values were predicted 30 of 39, i.e. 76.92%, of which 7 rows in full;

compared to the age group predicted data set.

4. Conclusions

The main results for this research could be synthesized as follows:

- For both categories of patients assigned to sex the proportion of the set of measurements (57%) and prediction (43%) was kept.
- The regression slopes for male patients have a general average of 0.964 compared to female patients that is 0.378.
- RNA-BP architecture varies only in terms of the number of nodes per hidden layer, namely 5, 7 and 10 and is not influenced by the number of hidden layers.
- The information flow between the input layer and the first hidden layer, in 5 cases out of 10, comprises the domain [-1, 1] and the importance (weight) of the links is very large. These RNA-BPs are characterized by a negative flow between the inlet layer and the exit layer. In three cases between the input and the exit layer, a central information flow is established with the importance (weight) of the moderate and homogeneous links.
- The number of training cycles needed is between 65,000 and 130,000 with a centering at 100,000, i.e. 57%.
- The sum of prediction errors is 0.00001-0.3198, male RNA-BP being centered on low value errors and females to high value errors. For RNA-BP for male patients an average error of 10.9% is obtained and for female patients of 18.3%.

REFERENCES

[1] Zadeh LA. *Fuzzy sets*. Information and Control 1965; 3:338±53.

[2] Oshita S, Nakakimura K, Sakabe T. *Hypertension control during anesthesia. Fuzzy logic regulation of nicardipine infusion*. IEEE Engineering in Medicine and Biology Magazine. 1994; 13(5):667–670.

[3] Castiglione F, Poccia F, D'Offizi G, Bernaschi M. *Mutation, fitness, viral diversity, and predictive markers of disease progression in a computational model of HIV type 1 infection*. AIDS Res Hum Retroviruses 2004 Dec; 20(12):1314-1323.

[4] Jaremko JL, Poncet P, Ronsky J, Harder J, Dansereau J, Labelle H, et al. *Genetic algorithm-neural network estimation of cobb angle from torso asymmetry in scoliosis*. J Biomech Eng 2002 Oct; 124(5):496-503.

[5] Latkowski T, Osowski S. *Computerized system for recognition of autism on the basis of gene expression microarray data*. Comput Biol Med 2015 Jan; 56:82-88.

[6] Szaleniec J, Skladzien J, Tadeusiewicz R, Oles K, Konior M, Przeklasa R. *How can an otolaryngologist benefit from artificial neural networks?,* Otolaryngol Pol 2012:66:241-8.

[7] Szaleniec J, Wiatr M, Skladzien J, Skladzien J, Tomik J, Oles K, et al. *Artificial neural network modeling of the results of tympanoplasty in chronic suppurative otitis media patients*. Comput Biol Med 2013:43:16-22

[8] Dunmade, A. O., A. D. Dunmade, O. A. Taiwo, A. R. Tomori, and T. M. Komolafe, *A software system for diagnosis and classification of deafness*, European Journal of Scientific Research 25, no. 4 (2009): 597-605

[9] Oana Alexandra Stelea (căsătorit Sandu), Codrut Sarafoleanu, Gabriela Muşat, *Hearing loss in children-literature review and two clinical cases*, Romania Journal of Rhinology, Volume 3, No.11, July-September 2013/ ISSN 2069-6523

[10] Daniel-Petru Ghencea, Miron Zapciu, Claudiu Florinel Bişu, *Prediction the evolution of temperature and vibrations on spindle using artificial neural networks and fuzzy logic*, Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information, Vol. 8, No. 1/2016, ISSN 2066-6950, Ed. Academia Oamenilor de Ştiinţă din România, Bucureşti,

http://www.aosr.ro/wp-content/uploads/2016/08/TVol7Nr2Art.4.Abs_.pdf

[11] Daniel-Petru Ghencea, *Modelare-simulare şi predicția datelor experimentale specifice sistemelor de fabricație utilizând tehnici hibride bazate pe inteligența artificială*, Teză de doctorat, Universitatea ”POLITEHNICA” din Bucureşti - Facultatea de Ingineria şi Managementul Sistemelor Tehnologice, Bucureşti, 2018.